Large deviation approach to the generalized random energy model

2002 J. Phys. A: Math. Gen. 35 4385

(http://iopscience.iop.org/0305-4470/35/20/301)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **35** (2002) 4385–4394 PII: S0305-4470(02)33237-2

Large deviation approach to the generalized random energy model

T C Dorlas and W M B Dukes

School of Theoretical Physics, Dublin Institute for Advanced Studies, 10 Burlington Road, Dublin 4, Republic of Ireland

Received 25 January 2002, in final form 7 March 2002
Published 10 May 2002
Online at stacks.iop.org/JPhysA/35/4385

Abstract

The generalized random energy model is a generalization of the random energy model introduced by Derrida to mimic the ultrametric structure of the Parisi solution of the Sherrington–Kirkpatrick model of a spin glass. It was solved exactly in two special cases by Derrida and Gardner. A complete solution for the thermodynamics in the general case was given by Capocaccia *et al*. Here we use large deviation theory to analyse the model in a very straightforward way. We also show that the variational expression for the free energy can be evaluated easily using the Cauchy–Schwarz inequality.

PACS numbers: 75.10.Nf, 05.50.+q

Introduction

The generalized random energy model (GREM) was introduced by Derrida [1] as a generalization of his random energy model (see Derrida [2]) of a spin glass in order to incorporate some correlations between energy levels. The GREM is one of the few spin glass models which lends itself to rigorous analysis, not only for the free energy density (Capocaccia *et al* [3]), but also for the fluctuations (Galves *et al* [4]) and dynamical properties (Ben Arous *et al* [5]).

The purpose of the present paper is to show how the theory of large deviations is the most natural avenue of attack when solving these models (if they are exactly solvable). This was done for the REM by Dorlas and Wedagedera [6]. An introduction to large deviations can be found in the books [7–9].

In section 1 we describe the model. In section 2 we give an expression for the rate function of the measures associated with the energies on different levels of the tree. In section 3 we give an explicit form for the free energy by applying Varadhan's lemma and then solving the associated variational problem.

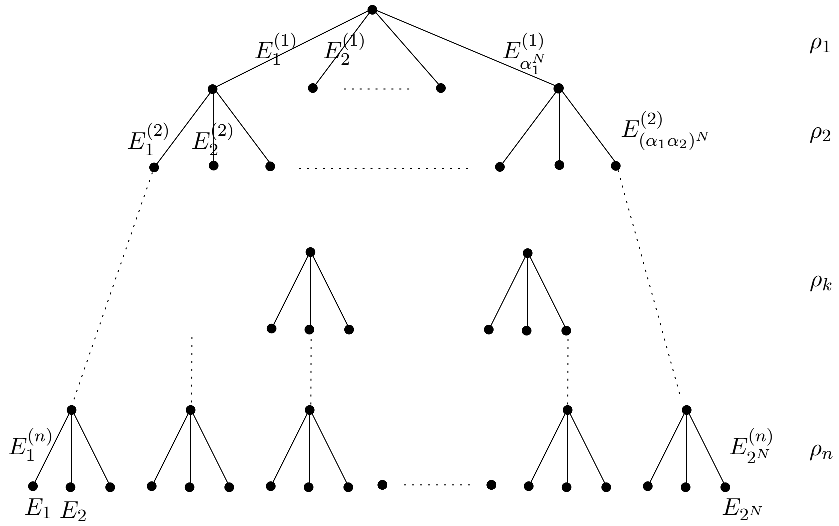

Figure 1. The tree-like structure of the GREM. The nodes on the *n*th layer represent the configurations. The energy of any configuration is the sum of the energies on the branches up to the source node.

1. Definition of the GREM

Whereas in the random energy model all energy levels E_i are i.i.d. random variables, and the partition function is given by

$$Z_N(\beta) = \sum_{i=1}^{2^N} \mathrm{e}^{-\beta E_i}$$

the energy levels of the generalized model have a tree-like structure. The tree is defined by a number of levels n and for each level $k = 1, \ldots, n$, a number $\alpha_k \in (1, 2)$ determines the number of branches per node (see figure 1). To make the total number of highest-level branches in the tree add up to 2^N as before, it is assumed that $\prod_{i=1}^n \alpha_k = 2$. For each $k = 1, \ldots, n$ there are $(\alpha_1 \cdots \alpha_k)^N$ independent random variables $\{E_j^{(k)}\}$, distributed according to $\rho_N^{(k)}$ with density

$$\rho_N^{(k)}(E) = \frac{1}{\sqrt{a_k \pi N J^2}} \mathrm{e}^{-E^2/a_k N J^2} \tag{1.1}$$

where the positive numbers a_k satisfy $\sum_{k=1}^n a_k = 1$. (Obviously, in general α_k^N is not an integer, but we can take its integer part which is very nearly the same for large *N*. We shall disregard the difference in the following.)

The partition function of the GREM is defined by

$$Z_N(\beta) = \sum_{i_1=1}^{\alpha_1^N} \sum_{i_2=(i_1-1)\alpha_2^N+1}^{i_1\alpha_2^N} \cdots \sum_{i_n=(i_{n-1}-1)\alpha_n^N+1}^{i_{n-1}\alpha_n^N} \exp\left[-\beta\left(\sum_{k=1}^n E_{i_k}^{(k)}\right)\right]. \tag{1.2}$$

The energy levels of the GREM are sums of energies corresponding to the different levels of the tree. This introduces a hierarchical dependence between energy levels similar to that in Parisi's solution of the Sherrington–Kirkpatrick model [10–12]. The previous formula is best understood by referring to figure 1.

As usual the free energy is defined by

$$f(\beta) = -\frac{1}{\beta} \lim_{N\to\infty} \frac{1}{N} \ln Z_N(\beta). \tag{1.3}$$

We shall prove that this limit exists almost surely w.r.t. the distribution of the energies $\{E_i^{(k)}\}$. To do this, we introduce the random distribution functions $F_N(x_1,\ldots,x_n)$ and $\bar{F}_N(x_1,\ldots,x_n)$ as follows:

$$\begin{aligned} F_N(x_1,\ldots,x_n) &:= \frac{1}{2^N} \sum_{i_1=1}^{\alpha_1^N} \sum_{i_2=(i_1-1)\alpha_2^N+1}^{i_1\alpha_2^N} \cdots \sum_{i_n=(i_{n-1}-1)\alpha_n^N+1}^{i_{n-1}\alpha_n^N} \bar{\mathbb{1}}_{i_1}^{(1)} \bar{\mathbb{1}}_{i_2}^{(2)} \cdots \bar{\mathbb{1}}_{i_n}^{(n)} \\ \bar{F}_N(x_1,\ldots,x_n) &:= \frac{1}{2^N} \sum_{i_1=1}^{\alpha_1^N} \sum_{i_2=(i_1-1)\alpha_2^N+1}^{i_1\alpha_2^N} \cdots \sum_{i_n=(i_{n-1}-1)\alpha_n^N+1}^{i_{n-1}\alpha_n^N} \mathbb{1}_{i_1}^{(1)} \mathbb{1}_{i_2}^{(2)} \cdots \mathbb{1}_{i_n}^{(n)} \end{aligned} \tag{1.4}$$

where we use the notation $\mathbb{1}_i^{(k)}$, $\bar{\mathbb{1}}_i^{(k)}$for the indicator functions of the sets $\{E_i^{(k)} > Nx_k\}$ and $\{E_i^{(k)} \leqslant Nx_k\}$, respectively. We also define G_N and $\bar{G}_N$ as

$$G_N(x_1,\ldots,x_n) = \int_{-\infty}^{Nx_1} \cdots \int_{-\infty}^{Nx_n} \rho_N^{(1)}(E_1)\cdots\rho_N^{(n)}(E_n)\,\mathrm{d}E_n\cdots\mathrm{d}E_1$$

$$\bar{G}_N(x_1,\ldots,x_n) := \int_{Nx_1}^{+\infty} \cdots \int_{Nx_n}^{+\infty} \rho_N^{(1)}(E_1)\cdots\rho_N^{(n)}(E_n)\,\mathrm{d}E_n\cdots\mathrm{d}E_1.$$

We will abbreviate $\bar{G}_N(x_1,\ldots,x_n)$ to $\bar{G}_N$ and $\bar{F}_N(x_1,\ldots,x_n)$ to $\bar{F}_N$. Let us also use, as short-hand,

$$p_k := \mathbb{P}\left(E^{(k)} > Nx_k\right)$$

where $\mathbb{P}$ denotes the probability w.r.t. the distribution (1.1). Note that $\bar{G}_N = p_1 p_2 \cdots p_n$. In the following section we prove a large deviation principle (LDP) for the distribution functions F_N analogous to that of Dorlas and Wedagedera [6].

2. The rate function

Theorem 2.1. *The sequence of measures $\mu_N(x_1,\ldots,x_n)$ with distribution function $F_N(x_1,\ldots,x_n)$ satisfies a LDP, almost surely with respect to the randomness, with rate function $I(x_1,\ldots,x_n)$ given by*

$$I(x_1,\ldots,x_n) = \begin{cases} \dfrac{1}{J^2} \displaystyle\sum_{1\leqslant i\leqslant n} \frac{x_i^2}{a_i} & \text{if} \quad (x_1,\ldots,x_n) \in \Psi(J; a_1,\ldots,a_n; \alpha_1,\ldots,\alpha_n) \\ +\infty & \text{otherwise} \end{cases}$$

where the region $\Psi(J; a_1,\ldots,a_n; \alpha_1,\ldots,\alpha_n)$ is given by

$$\left\{ (x_1,\ldots,x_n) \in \mathbb{R}^n \,\middle|\, \sum_{i=1}^{k} \frac{x_i^2}{a_i} \leqslant J^2 \sum_{i=1}^{k} \ln\alpha_i \text{ for all } 1 \leqslant k \leqslant n \right\}.$$

Proof. First we do the case for $(x_1,\ldots,x_n) \in \Psi(J; a_1,\ldots,a_n; \alpha_1,\ldots,\alpha_n)$. By Chebyshev's inequality, for all $\epsilon \in (0,1)$,

$$\mathbb{P}(|\bar{F}_N - \bar{G}_N| > \epsilon\bar{G}_N) \leqslant \frac{1}{\epsilon^2\bar{G}_N^2}\mathbb{E}\left(|\bar{F}_N - \bar{G}_N|^2\right).$$

Now

$$\mathbb{E}\big(|\bar{F}_N - \bar{G}_N|^2\big) = \mathbb{E}\big(\bar{F}_N^2\big) - 2\bar{G}_N\mathbb{E}(\bar{F}_N) + \bar{G}_N^2$$

and

$$\begin{aligned}\mathbb{E}(\bar{F}_N) &= \frac{1}{2^N}\mathbb{E}\left(\sum_{i_1=1}^{\alpha_1^N}\sum_{i_2=(i_1-1)\alpha_2^N+1}^{i_1\alpha_2^N}\cdots\sum_{i_n=(i_{n-1}-1)\alpha_n^N+1}^{i_{n-1}\alpha_n^N}\mathbb{1}_{i_1}^{(1)}\mathbb{1}_{i_2}^{(2)}\cdots\mathbb{1}_{i_n}^{(n)}\right)\\ &= \frac{1}{2^N}\alpha_1^N p_1\alpha_2^N p_2\cdots\alpha_n^N p_n\\ &= p_1p_2\cdots p_n\\ &= \bar{G}_N.\end{aligned}$$

To obtain $\mathbb{E}\big(\bar{F}_N^2\big)$ we introduce some new notation. Let

$$\mathcal{B}_k := \mathbb{E}\left\{\sum_{\substack{i_k=1\\ j_k=1}}^{\substack{i_k=\alpha_k^N\\ j_k=\alpha_k^N}}\mathbb{1}_{i_k}^{(k)}\mathbb{1}_{j_k}^{(k)}\sum_{\substack{i_{k+1}=(i_k-1)\alpha_{k+1}^N+1\\ j_{k+1}=(j_k-1)\alpha_{k+1}^N+1}}^{\substack{i_{k+1}=i_k\alpha_{k+1}^N\\ j_{k+1}=j_k\alpha_{k+1}^N}}\mathbb{1}_{i_{k+1}}^{(k+1)}\mathbb{1}_{j_{k+1}}^{(k+1)}\cdots\sum_{\substack{i_n=(i_{n-1}-1)\alpha_n^N+1\\ j_n=(j_{n-1}-1)\alpha_n^N+1}}^{\substack{i_n=i_{n-1}\alpha_n^N\\ j_k=j_{n-1}\alpha_n^N}}\mathbb{1}_{i_n}^{(n)}\mathbb{1}_{j_n}^{(n)}\right\}.$$

Now notice that the following recursion holds:

$$\mathcal{B}_k = \alpha_k^N p_k\left(\mathcal{B}_{k+1} + \left(\alpha_k^N - 1\right)p_k\left(\alpha_{k+1}^N p_{k+1}\cdots\alpha_n^N p_n\right)^2\right)$$

for all $1 \leqslant k < n$. The initial value is $\mathcal{B}_n = \alpha_n^N p_n + \left(\alpha_n^{2N} - \alpha_n^N\right)p_n^2$ but this may be obtained by defining $\mathcal{B}_{n+1} := 1$ and applying the above recursion for $k = n$. Notice also that $\mathbb{E}\big(\bar{F}_N^2\big) = \frac{1}{2^{2N}}\mathcal{B}_1$.

Alongside the above recursion, let us define a sequence $\mathcal{D}_k$ by which we upper bound $\mathcal{B}_k$. Let $\mathcal{D}_{n+1} := 1$ and define

$$\mathcal{D}_k = y_k\left(\mathcal{D}_{k+1} + y_k(y_{k+1}\cdots y_n)^2\right).$$

This gives rise to

$$\mathcal{D}_1 = y_1y_2\cdots y_n(1 + y_n + y_{n-1}y_n + \cdots + y_1y_2\cdots y_n).$$

If we now take $y_k = \alpha_k^N p_k$, then it is clear that $D_k \geqslant \mathcal{B}_k$ for all $1 \leqslant k \leqslant n$, hence the following bound:

$$\begin{aligned}\mathbb{E}\big(\bar{F}_N^2\big) &= \frac{1}{2^{2N}}\mathcal{B}_1\\ &\leqslant \frac{1}{2^{2N}}\mathcal{D}_1\\ &= \left(\prod_{1\leqslant k\leqslant n}\alpha_k^N p_k\right)\left(1 + \sum_{k=1}^{n}\alpha_k^N p_k\cdots\alpha_n^N p_n\right)\\ &= 2^N\bar{G}_N\left(1 + \sum_{k=1}^{n}\alpha_k^N p_k\cdots\alpha_n^N p_n\right).\end{aligned}$$

Thus

$$\begin{aligned}
\frac{1}{\epsilon^2\bar{G}_N^2}\mathbb{E}\left(|\bar{F}_N-\bar{G}_N|^2\right) &\leqslant \frac{1}{\epsilon^2\bar{G}_N^2}\left\{\frac{1}{2^{2N}}2^N\bar{G}_N\left(1+\sum_{k=1}^{n}\alpha_k^N p_k\cdots\alpha_n^N p_n\right)-\bar{G}_N^2\right\}\\
&= \frac{1}{\epsilon^2\bar{G}_N^2}\left\{\frac{1}{2^N}\bar{G}_N\left(1+\sum_{k=1}^{n}\alpha_k^N p_k\cdots\alpha_n^N p_n\right)-\bar{G}_N^2\right\}\\
&= \frac{1}{\epsilon^2\bar{G}_N^2}\left\{\frac{1}{2^N}\bar{G}_N\left(1+2^N\bar{G}_N+\sum_{k=2}^{n}\alpha_k^N p_k\cdots\alpha_n^N p_n\right)-\bar{G}_N^2\right\}\\
&= \frac{1}{\epsilon^2 2^N\bar{G}_N}\left\{1+\sum_{k=2}^{n}\alpha_k^N p_k\cdots\alpha_n^N p_n\right\}\\
&= \frac{1}{\epsilon^2 2^N\bar{G}_N}+\frac{1}{\epsilon^2}\sum_{k=2}^{n}\frac{\alpha_k^N p_k\cdots\alpha_n^N p_n}{2^N p_1 p_2\cdots p_n}\\
&= \frac{1}{\epsilon^2 2^N\bar{G}_N}+\frac{1}{\epsilon^2}\sum_{k=2}^{n}\frac{1}{\alpha_1^N p_1\cdots\alpha_{k-1}^N p_{k-1}}\\
&= \frac{1}{\epsilon^2 2^N\bar{G}_N}+\frac{1}{\epsilon^2}\sum_{k=1}^{n-1}\frac{1}{\alpha_1^N p_1\cdots\alpha_k^N p_k}\\
&= \frac{1}{\epsilon^2}\sum_{k=1}^{n}\frac{1}{\alpha_1^N p_1\cdots\alpha_k^N p_k}.
\end{aligned}\tag{2.1}$$

Using the inequality $\int_a^\infty \mathrm{e}^{-u^2/2}\,\mathrm{d}u > \frac{1}{a+a^{-1}}\,\mathrm{e}^{-a^2/2}$ (see McKean [13]), the kth term in this sum is bounded above by

$$\frac{1}{\epsilon^2}\left\{\prod_{1\leqslant i\leqslant k}\frac{\sqrt{\pi}\left(2Nx_i^2+a_iJ^2\right)}{x_iJ\sqrt{a_iN}}\right\}\exp\left[\frac{N}{J^2}\left(-J^2\sum_{i=1}^{k}\ln\alpha_i+\sum_{1\leqslant i\leqslant k}\frac{x_i^2}{a_i}\right)\right].$$

which will converge if and only if $\sum_{1\leqslant i\leqslant k}\frac{x_i^2}{a_i} < J^2\sum_{1\leqslant i\leqslant k}\ln\alpha_i$. Thus it is seen that equation (2.1) converges if all the sums of its individual terms converge. The values for which this happens are precisely those which define the region $\Psi(J; a_1,\ldots,a_n;\alpha_1,\ldots,\alpha_n)$ as stated in the theorem. Introducing the events

$$\mathcal{A}_N = \left\{\left\{E_{i_1}^{(1)},\ldots,E_{i_n}^{(n)}\right\}\big||\bar{G}_N-\bar{F}_N| > \epsilon\bar{G}_N\right\}$$

we see that $\sum_N \mathbb{P}(\mathcal{A}_N) < +\infty$. Hence by the Borel–Cantelli lemma,

$$\mathbb{P}\left(\bigcap_{v=1}^{\infty}\bigcup_{N=v}^{\infty}\mathcal{A}_N\right) = 0.$$

This means that with probability 1,

$$\left\{E_{i_1}^{(1)},\ldots,E_{i_n}^{(n)}\right\}\in\left(\bigcap_{v=1}^{\infty}\bigcup_{N=v}^{\infty}\mathcal{A}_N\right)^C = \bigcup_{v=1}^{\infty}\bigcap_{N=v}^{\infty}\mathcal{A}_N^C.$$

In other words, for almost all $\left\{E_{i_1}^{(1)},\ldots,E_{i_n}^{(n)}\right\}$ there exists a $v\in\mathbb{N}$ such that for all $N\geqslant v$, $\left\{E_{i_1}^{(1)},\ldots,E_{i_n}^{(n)}\right\}\in\mathcal{A}_N^C$. Hence $\bar{F}_N$ converges to $\bar{G}_N$ with probability 1 for all $N\geqslant v$.

For the case $(x_1, \ldots, x_n) \notin \Psi(J; a_1, \ldots, a_n; \alpha_1, \ldots, \alpha_n)$, it must hold that

$$\sum_{1\leqslant i\leqslant k} \frac{x_i^2}{a_i} > J^2 \sum_{1\leqslant i\leqslant k} \ln \alpha_i$$

for some k with $1 \leqslant k \leqslant n$. We may now upper bound the function $\bar{F}_N(x_1, \ldots, x_n)$ by

$$\begin{aligned}\bar{F}_N(x_1, \ldots, x_n) &\leqslant \frac{1}{2^N} \alpha_{k+1}^N \cdots \alpha_n^N \sum_{i_1=1}^{\alpha_1^N} \sum_{i_2=(i_1-1)\alpha_2^N+1}^{i_1\alpha_2^N} \cdots \sum_{i_k=(i_{k-1}-1)\alpha_k^N+1}^{i_{k-1}\alpha_k^N} \mathbb{1}_{i_1}^{(1)} \mathbb{1}_{i_2}^{(2)} \cdots \mathbb{1}_{i_k}^{(k)} \\ &= \frac{1}{\alpha_1^N \cdots \alpha_k^N} \sum_{i_1=1}^{\alpha_1^N} \sum_{i_2=(i_1-1)\alpha_2^N+1}^{i_1\alpha_2^N} \cdots \sum_{i_k=(i_{k-1}-1)\alpha_k^N+1}^{i_{k-1}\alpha_k^N} \mathbb{1}_{i_1}^{(1)} \mathbb{1}_{i_2}^{(2)} \cdots \mathbb{1}_{i_k}^{(k)} \\ &=: H_N(x_1, \ldots, x_k).\end{aligned}$$

We will show that $H_N(x_1, \ldots, x_k) = 0$ with probability 1 if N is large enough. We have

$$\begin{aligned}&\left\{\left\{E_{i_1}^{(1)}, \ldots, E_{i_k}^{(k)}\right\} \middle| H_N(x_1, \ldots, x_k) = 0\right\} \\ &\qquad = \left\{\left\{E_{i_1}^{(1)}, \ldots, E_{i_k}^{(k)}\right\} \middle| \sum_{i_1=1}^{\alpha_1^N} \sum_{i_2=(i_1-1)\alpha_2^N+1}^{i_1\alpha_2^N} \right. \\ &\qquad\qquad \left. \cdots \sum_{i_k=(i_{k-1}-1)\alpha_k^N+1}^{i_{k-1}\alpha_k^N} \mathbb{1}_{i_1}^{(1)} \mathbb{1}_{i_2}^{(2)} \cdots \mathbb{1}_{i_k}^{(k)} < 1 \right\}.\end{aligned}$$

By Chebyshev's inequality,

$$\begin{aligned}&\mathbb{P}\left(\sum_{i_1=1}^{\alpha_1^N} \sum_{i_2=(i_1-1)\alpha_2^N+1}^{i_1\alpha_2^N} \cdots \sum_{i_k=(i_{k-1}-1)\alpha_k^N+1}^{i_{k-1}\alpha_k^N} \mathbb{1}_{i_1}^{(1)} \mathbb{1}_{i_2}^{(2)} \cdots \mathbb{1}_{i_k}^{(k)} \geqslant 1\right) \\ &\qquad \leqslant \mathbb{E}\left(\sum_{i_1=1}^{\alpha_1^N} \sum_{i_2=(i_1-1)\alpha_2^N+1}^{i_1\alpha_2^N} \cdots \sum_{i_k=(i_{k-1}-1)\alpha_k^N+1}^{i_{k-1}\alpha_k^N} \mathbb{1}_{i_1}^{(1)} \mathbb{1}_{i_2}^{(2)} \cdots \mathbb{1}_{i_k}^{(k)}\right) \\ &\qquad = \alpha_1^N \cdots \alpha_k^N \mathbb{P}\big(E^{(1)} > N x_1\big) \cdots \mathbb{P}\big(E^{(k)} > N x_k\big) \\ &\qquad \leqslant \alpha_1^N \cdots \alpha_k^N \prod_{1\leqslant i\leqslant k} \frac{J\sqrt{a_i}}{2x_i\sqrt{\pi N}} \exp\left(-\frac{N x_i^2}{a_i J^2}\right) \\ &\qquad = \left(\prod_{1\leqslant i\leqslant k} \frac{J\sqrt{a_i}}{2x_i\sqrt{\pi N}}\right) \exp\left\{N \sum_{1\leqslant i\leqslant k} \left(\ln \alpha_i - \frac{x_i^2}{a_i J^2}\right)\right\}.\end{aligned}$$

Since

$$\sum_{1\leqslant i\leqslant k} \frac{x_i^2}{a_i} > J^2 \sum_{1\leqslant i\leqslant k} \ln \alpha_i$$

the series

$$\sum_{N=1}^{\infty} \left(\prod_{1\leqslant i\leqslant k} \frac{J\sqrt{a_i}}{2x_i\sqrt{\pi N}}\right) \exp\left\{N \sum_{1\leqslant i\leqslant k} \left(\ln \alpha_i - \frac{x_i^2}{a_i J^2}\right)\right\}$$

converges. Introducing the events

$$\mathcal{A}_N = \left\{ \left\{E^{(1)}_{i_1}, \ldots, E^{(k)}_{i_k}\right\} \middle| \sum_{i_1=1}^{\alpha_1^N} \sum_{i_2=(i_1-1)\alpha_2^N+1}^{i_1\alpha_2^N} \cdots \sum_{i_k=(i_{k-1}-1)\alpha_k^N+1}^{i_{k-1}\alpha_k^N} \mathbb{1}^{(1)}_{i_1} \mathbb{1}^{(2)}_{i_2} \cdots \mathbb{1}^{(k)}_{i_k} \geqslant 1 \right\}$$

we see again by the Borel–Cantelli lemma, for almost all $\left\{E^{(1)}_{i_1}, \ldots, E^{(k)}_{i_k}\right\}$ there exists a $v \in \mathbb{N}$ such that for all $N \geqslant v$, $\left\{E^{(1)}_{i_1}, \ldots, E^{(k)}_{i_k}\right\} \in \mathcal{A}_N^C$ and hence $H_N(x_1, \ldots, x_k) = 0$. Thus we have:

$$\begin{aligned}\limsup_N \frac{1}{N} \ln \bar{F}_N(x_1, \ldots, x_n) &\leqslant \limsup_N \frac{1}{N} \ln H_N(x_1, \ldots, x_k) \\ &= -\infty.\end{aligned}$$

□

3. The variational problem

We may re-write the partition function in (1.2) as

$$\mathcal{Z}_N(\beta) = 2^N \int_{\mathbb{R}^n} \exp\{-N\beta(x_1 + \cdots + x_n)\}\, \mathrm{d}F_N(x_1, \ldots, x_n)$$

where $F_N(x_1, \ldots, x_n)$ is given in (1.4). Using Varadhan's lemma, we may evaluate $-\beta f(\beta)$ almost surely as follows:

$$\begin{aligned}-\beta f(\beta) &= \lim_{N\to\infty} \frac{1}{N}\{\ln \mathcal{Z}_N(\beta)\} \\ &= \ln 2 + \sup_{(x_1,\ldots,x_n)\in\mathbb{R}^n} \{-\beta(x_1 + \cdots + x_n) - I(x_1, \ldots, x_n)\} \\ &= \ln 2 - \inf_{\vec{x}\in\Psi} \left\{ \sum_{i=1}^n \frac{x_i^2}{a_i J^2} + \beta x_i \right\} \\ &= \ln 2 + \frac{1}{4}\beta^2 J^2 - \frac{1}{J^2} \inf_{\vec{x}\in\Psi} \left\{ \sum_{i=1}^n \frac{1}{a_i} \left(x_i + \frac{1}{2} a_i \beta J^2 \right)^2 \right\}.\end{aligned}$$

Performing the change of variables: $x_i = J y_i \sqrt{a_i}$, $\beta' = \frac{1}{2}\beta J$ and $\gamma_i = \ln \alpha_i$, the above expression becomes

$$= \ln 2 + \frac{1}{4}\beta^2 J^2 - \inf_{\vec{y}\in\Psi'} \left\{ \sum_{i=1}^n \left(y_i - \sqrt{a_i}\beta' \right)^2 \right\}$$

where

$$\Psi' = \left\{ (y_1, \ldots, y_n) \in \mathbb{R}^n \middle| \sum_{i=1}^k y_i^2 \leqslant \sum_{i=1}^k \gamma_i \text{ for all } 1 \leqslant k \leqslant n \right\}.$$

3.1. Evaluation of the infimum in $\mathbb{R}^n$

Define the numbers $B(j, k)$ for all $1 \leqslant j \leqslant k \leqslant n$:

$$B(j, k) := \sqrt{\frac{\gamma_j + \cdots + \gamma_k}{a_j + \cdots + a_k}}.$$

Let $m_0 := 0$ and recursively define the numbers m_i as

$$m_i := \inf\{m > m_{i-1} | B(m_{i-1} + 1, m) \leqslant B(m_{i-1} + 1, l) \text{ for all } m_{i-1} + 1 \leqslant l \leqslant n\}$$

terminating at the value K such that $m_K = n$. A crucial property of algebraic expressions such as $B(j,k)$ is the following: if a, b, c and d are positive reals, then $\frac{a}{b} < \frac{c}{d}$ if and only if $\frac{a}{b} < \frac{a+c}{b+d}$. Define the sequence of inverse temperatures β_i $(i = 0, \ldots, K+1)$ by

$$\beta_i := B(m_{i-1}+1, m_i) \qquad i = 1, \ldots, K$$

and $\beta_0 := 0$, $\beta_{K+1} := +\infty$. Note that this sequence is increasing by the above property.

Lemma 3.1 . *If $\beta_j < \beta' < \beta_{j+1}$ for some $0 \leqslant j \leqslant K$, then the infimum is attained at $\vec{x}$ given by*

$$x_i = \begin{cases} \beta_l \sqrt{a_i} & \textit{if} \quad i \in [m_{l-1}+1, \ldots, m_l] \qquad \text{for some} \quad 1 \leqslant l \leqslant j \\ \beta' \sqrt{a_i} & \textit{if} \quad i \in [m_j+1, \ldots, n] \end{cases}$$

for all $1 \leqslant i \leqslant n$.

Proof. Let $p_i = \sqrt{a_i}$ for all $1 \leqslant i \leqslant n$. We will show that the point $\vec{x}$ with coordinates given above is the point such that for all $y \in \Psi'$, $\| \vec{y} - \beta'\vec{p} \| \geqslant \| \vec{x} - \beta'\vec{p} \|$. First, let us note two trivial inequalities,

$$\sum_{l=1}^{j} \sum_{i=m_{l-1}+1}^{m_l} (y_i - \beta_l p_i)^2 \geqslant 0 \tag{3.1}$$

$$\sum_{i=m_j+1}^{n} (y_i - \beta' p_i)^2 \geqslant 0. \tag{3.2}$$

Note that for all $1 \leqslant l \leqslant j$, $\left(\frac{\beta'}{\beta_l} - 1\right) > 0$. By the Cauchy–Schwarz inequality we have, for all $1 \leqslant j' \leqslant j$,

$$\sum_{l=1}^{j'} \sum_{i=m_{l-1}+1}^{m_l} \beta_l p_i y_i \leqslant \left(\sum_{i=1}^{m_{j'}} y_i^2 \right)^{1/2} \left(\sum_{l=1}^{j'} \sum_{i=m_{l-1}+1}^{m_l} \beta_l^2 p_i^2 \right)^{1/2}$$

$$\leqslant \left(\sum_{i=1}^{m_{j'}} \gamma_i \right)^{1/2} \left(\sum_{l=1}^{j'} \beta_l^2 \sum_{i=m_{l-1}+1}^{m_l} p_i^2 \right)^{1/2}.$$

Notice that $\sum_{l=1}^{j'} \beta_l^2 \sum_{i=m_{l-1}+1}^{m_l} p_i^2 = \sum_{l=1}^{j'} \sum_{i=m_{l-1}+1}^{m_l} \gamma_i = \sum_{i=1}^{m_{j'}} \gamma_i$ and so the above expression becomes

$$= \sum_{l=1}^{j'} \beta_l^2 \sum_{i=m_{l-1}+1}^{m_l} p_i^2.$$

Thus we have

$$\sum_{l=1}^{j'} \sum_{i=m_{l-1}+1}^{m_l} \beta_l p_i (\beta_l p_i - y_i) \geqslant 0$$

for all $1 \leqslant j' \leqslant j$ and $\vec{y} \in \Psi'$. Introducing the numbers $\left(\frac{\beta'}{\beta_l} - 1\right)$ into the sum, it is shown by a recursive argument (see appendix) that

$$\sum_{l=1}^{j} \left(\frac{\beta'}{\beta_l} - 1 \right) \sum_{i=m_{l-1}+1}^{m_l} \beta_l p_i (\beta_l p_i - y_i) \geqslant 0.$$

Multiplying the above inequality by 2 and re-writing we have

$$2\sum_{l=1}^{j}\sum_{i=m_{l-1}+1}^{m_l}(\beta' p_i - \beta_l p_i)(\beta_l p_i - y_i) \geqslant 0. \tag{3.3}$$

Combining equations (3.1), (3.2) and (3.3) while noting that $\beta' p_i =: x_i$ (for $m_j + 1 \leqslant i \leqslant n$) and $\beta_l p_i =: x_i$ (for $1 \leqslant i \leqslant m_j$), we have

$$(\vec{y} - \vec{x}) \cdot (\vec{y} + \vec{x} - 2\beta' \vec{p}) \geqslant 0.$$

Re-writing gives

$$\| \vec{x} - \beta' \vec{p} \| \geqslant \| \vec{y} - \beta' \vec{p} \|$$

for all $\vec{y} \in \Psi'$. □

3.2. Expression for the free energy and conclusion

Applying the coordinates of our point of infimum to the expression for the free energy gives the required expressions. Recalling $\beta' := \frac{1}{2}\beta J$, $p_i = \sqrt{a_i}$ and $\gamma_i = \ln \alpha_i$ gives

Corollary 3.2. *The free energy is given by*

$$-\beta f(\beta) = \begin{cases} \ln 2 + \dfrac{1}{4}\beta^2 J^2 & \textit{if} \quad \beta < \dfrac{2}{J}\beta_1 \\ \displaystyle\sum_{i=m_j+1}^{n}\left(\ln\alpha_i + \frac{\beta^2 J^2 a_i}{4}\right) + \beta J \sum_{l=1}^{j}\sqrt{\left(\sum_{i=m_{l-1}+1}^{m_l} a_i\right)\left(\sum_{i=m_{l-1}+1}^{m_l}\ln\alpha_i\right)} & \\ & \textit{if} \quad \dfrac{2}{J}\beta_j < \beta < \dfrac{2}{J}\beta_{j+1} \\ \beta J \displaystyle\sum_{l=1}^{K}\sqrt{\left(\sum_{i=m_{l-1}+1}^{m_l} a_i\right)\left(\sum_{i=m_{l-1}+1}^{m_l}\ln\alpha_i\right)} & \textit{if} \quad \dfrac{2}{J}\beta_K < \beta. \end{cases}$$

Applying $n = 2$ to the above expression yields the same answer as Derrida [1]. In this case the answer depends on whether $a_1/\ln\alpha_1 \gtreqless a_2/\ln\alpha_2$. If $a_1/\ln\alpha_1 > a_2/\ln\alpha_2$, then

$$-\beta f(\beta) = \begin{cases} \ln 2 + \dfrac{J^2\beta^2}{4} & \text{if} \quad \beta < \dfrac{2}{J}\sqrt{\dfrac{\ln\alpha_1}{a_1}} \\ \ln\alpha_2 + \dfrac{1}{4}a_2\beta^2 J^2 + \beta J\sqrt{a_1 \ln\alpha_1} & \text{if} \quad \dfrac{2}{J}\sqrt{\dfrac{\ln\alpha_1}{a_1}} < \beta < \dfrac{2}{J}\sqrt{\dfrac{\ln\alpha_2}{a_2}} \\ \beta J\sqrt{a_1\ln\alpha_1} + \beta J\sqrt{a_2\ln\alpha_2} & \text{if} \quad \dfrac{2}{J}\sqrt{\dfrac{\ln\alpha_2}{a_2}} < \beta. \end{cases}$$

Otherwise,

$$-\beta f(\beta) = \begin{cases} \ln 2 + \dfrac{J^2\beta^2}{4} & \text{if} \quad \beta < \dfrac{2\sqrt{\ln 2}}{J} \\ \beta J\sqrt{\ln 2} & \text{if} \quad \beta > \dfrac{2\sqrt{\ln 2}}{J}. \end{cases}$$

It is an easy exercise to see the solutions also concur for cases A and B in Derrida and Gardner [14]. Notice that the Capocaccia *et al* [3] solution to the variational problem contains a few minor flaws: in their notation, they should have $J_0^\star = 0$ and the definition of β_k should be $\beta_k^2 := B_{J_{k-1}^\star+1, J_k^\star}$.

An added benefit of our approach is that in theorem 2.1 we have proved a LDP for the measures $\mu_N(x_1, \ldots, x_n)$. This result contains much more information than is needed for deriving the variational expression for the free energy. This leaves open the possibility of calculating other aspects of the GREM. Moreover, variants of the model, where one for example replaces the energies $\sum E_{i_k}^{(k)}$ by $\frac{1}{n}\big(\sum E_{i_k}^{(k)}\big)^2$ can also be solved.

Acknowledgments

The authors would like to thank the referees for their helpful comments.

Appendix

Lemma A. *Let* $x_1, x_2, \ldots, x_n > 0$ *and* $\{y_i\}_{i=1}^n$ *be a sequence of reals. Let* $G_m := \sum_{i=1}^m y_i$ *be such that* $G_m \geqslant 0$ *for all* $1 \leqslant m \leqslant n$. *Then*

$$F(n) := \sum_{i=1}^n x_i y_i \geqslant 0.$$

Proof. *Let us define* $G_0 = 0$. Notice that $y_i = G_i - G_{i-1}$ for all $1 \leqslant i \leqslant n$. Then

$$\begin{aligned} F(n) &= \sum_{i=1}^n x_i\,(G_i - G_{i-1}) \\ &= x_n G(n) + \sum_{i=1}^n (x_i - x_{i-1}) G(i) \\ &\geqslant 0 \end{aligned}$$

since $x_i - x_{i-1} > 0$ for all i. □

References

[1] Derrida B 1985 A generalization of the random energy model which includes correlations between energies *J. Phys. Lett.* **46** 401–7
[2] Derrida B 1980 *Phys. Rev. Lett.* **45** 79
[3] Cassandro M, Capocaccia D and Picco P 1987 On the existence of thermodynamics for the generalized random energy model *J. Stat. Phys.* **46** 493–505
[4] Martinez S, Galves A and Picco P 1989 *J. Stat. Phys.* **54** 515–29
[5] Bovier A, Ben Arous G and Gayrard V 2001 Aging in the random energy model *Preprint* cond-mat/0110223
[6] Dorlas T C and Wedagedera J R 2001 Large deviation and the random energy model *Int. J. Mod. Phys.* B **15** 1–15
[7] Ellis R S 1985 *Entropy, Large Deviations and Statistical Mechanics* (Berlin: Springer)
[8] Varadhan S R S 1984 *Large Deviations and Applications* (Philadelphia, PA: SIAM)
[9] Dorlas T C 1999 *Statistical Mechanics, Fundamentals and Model Solutions* (Bristol: Institute of Physics) p l
[10] Parisi G 1980 *J. Phys. A: Math. Gen.* **13** 1101
[11] Parisi G 1980 *J. Phys. A: Math. Gen.* **13** 1887
[12] Parisi G 1980 *Phil. Mag.* B **41** 677
[13] McKean H 1969 *Stochastic Integrals* (New York: Academic)
[14] Derrida B and Gardner E 1986 Solution of the generalized random energy model *J. Phys. C: Solid State Phys.* **19** 2253–74