Markov Processes Relat. Fields **13**, 251–266 (2007)

Markov Processes and **Related Fields**

On the Free Energy of a Directed Polymer in a Brownian Environment

J. Moriarty and N. O'Connell

Department of Mathematics and BCRI, University College Cork, Ireland

Dedicated to the memory of J.T. Lewis

Abstract. We prove a formula conjectured in [14] for the free energy density of a directed polymer in a Brownian environment in $1+1$ dimensions.

KEYWORDS: directed polymers, exponential functionals of Brownian motion, quasi-reversibility, Brownian queues, large deviations

AMS SUBJECT CLASSIFICATION: 60K37, 82D30, 60K25, 60J65

1. Introduction

Let $B^{(1)}, B^{(2)}, \ldots$ be independent standard one-dimensional Brownian motions. Denote the increments of $B^{(i)}$ by $B^{(i)}_{(s,t)} = B^{(i)}_t - B^{(i)}_s$.

For $\beta \in \mathbf{R}$ set

$$Z_n(\beta) = \int\limits_{0<s_1<\ldots<s_{n-1}<n} ds_1 \ldots ds_{n-1} \, \exp\left\{\beta(B^{(1)}_{(0,s_1)} + \ldots + B^{(n)}_{(s_{n-1},n)})\right\}. \tag{1.1}$$

This is the partition function for a continuous model of a directed polymer in a Brownian environment in $1+1$ dimensions. In the paper [14], using queueing-theoretic ideas in the context of geometric functionals of Brownian motion, certain limiting results were obtained which led the authors to conjecture an explicit formula for the free energy density

$$\lim_{n\to\infty} \frac{1}{n} \log Z_n(\beta), \tag{1.2}$$

namely that it should be given by, almost surely,

$$f(\beta) = \begin{cases} -(-\Psi)^*(-\beta^2) - 2\log|\beta| & : \quad \beta \neq 0, \\ 1 & : \quad \beta = 0, \end{cases} \tag{1.3}$$

where $\Psi(m) \equiv \Gamma'(m)/\Gamma(m)$ is the restriction of the digamma function to $(0, \infty)$, and $(-\Psi)^*$ is the convex dual of the function $-\Psi$. The aim of this paper is to give a rigorous proof of this conjecture. The proof uses tools from large deviation theory. As a corollary we give a new proof that $c \geq 2$, where

$$\lim_{n\to\infty} \frac{1}{n} L_n(n) = c \quad \text{a.s.} \tag{1.4}$$

and

$$L_n(t) = \sup_{0 \leq s_1 \leq \ldots \leq s_{n-1} \leq t} B^{(1)}_{(0,s_1)} + \ldots + B^{(n)}_{(s_{n-1},t)}.$$

It was proved using direct methods in [11] that $c = 2$, where the authors also describe how the result may be deduced from the theory of random matrices.

The directed polymer model we have discussed here, and for which we have computed the free energy density, is a continuous version of the classical two-dimensional directed polymer, where it is not known how to compute the free energy density (see, for example, [4, 6]). Recent work on a continuous model different to this can be found in [5].

In the next section we recall the framework which was developed in [14] to extend some standard constructions from queueing theory to the context of geometric functionals of Brownian motion, and explain how this leads to the conjectured formula for the free energy density. Section 3 is devoted to the proof of the main result. In Section 4 we prove that f is analytic and strictly convex, and record a large deviation principle that we will use in Section 5 to prove (1.4).

2. Generalised Brownian queues

In this section we recall the framework which was developed in [14] to extend some standard constructions from queueing theory to the context of geometric functionals of Brownian motion, and explain how this leads to the conjectured formula for the free energy density.

The *generalised Brownian queue* is characterised as follows. Let B and C be two independent standard Brownian motions indexed by the entire real line, and write

$$B_{(s,t)} = B_t - B_s, \quad C_{(s,t)} = C_t - C_s.$$

Fix $m > 0$ and, for $t \in \mathbf{R}$, set

$$
\begin{aligned}
r(t) &= \log \int_{-\infty}^{t} ds \, \exp\{B_{(s,t)} + C_{(s,t)} - m(t-s)\}, \\
f(s,t) &= B_{(s,t)} + r(s) - r(t), \\
g(s,t) &= C_{(s,t)} + r(s) - r(t),
\end{aligned}
$$

and define $f : \mathbf{R} \to \mathbf{R}$ by $f(t) = f(0,t)$.

To put this in context, the 'Brownian queue' is defined similarly but with 'log $\int$ exp' replaced by 'sup' (see, for example, [14]). The usual $M/M/1$ queue is defined similarly to the Brownian queue but with Brownian motions replaced by Poisson counting processes. Thus the Brownian motions B_t and $mt - C_t$ can be thought of, respectively, as the *arrivals* and *service* processes, r as the *queue-length process* and f as the *output*, or *departure*, process.

In [14] it is shown, using results of Matsumoto and Yor [12, 13], that the generalised Brownian queue is *quasi-reversible*, that is: f is a standard Brownian motion, and $\{f(s), s \le t\}$ is independent of $r(t)$. We can thus consider a sequence of generalised Brownian queues in tandem and expect this 'queueing network' to have nice properties (analogous to the 'product-form solutions' of classical queueing theory).

Let $B, B^{(1)}, B^{(2)}, \ldots$ be a sequence of independent standard Brownian motions, each indexed by $\mathbf{R}$, and let $m > 0$ be a fixed constant. For $-\infty < s \le t < \infty$, set

$$
\begin{aligned}
r_1(t) &= \log \int_{-\infty}^{t} ds \, \exp\{B_{(s,t)} + B^{(1)}_{(s,t)} - m(t-s)\}, \\
f_1(s,t) &= B_{(s,t)} + r_1(s) - r_1(t), \\
g(s,t) &= B^{(1)}_{(s,t)} + r_1(s) - r_1(t),
\end{aligned}
$$

for each $k = 2, 3, \ldots$ set

$$
\begin{aligned}
r_k(t) &= \log \int_{-\infty}^{t} ds \, \exp\{f_{k-1}(s,t) + B^{(k)}_{(s,t)} - m(t-s)\}, \\
f_k(s,t) &= f_{k-1}(s,t) + r_k(s) - r_k(t),
\end{aligned}
$$

and for all k define $f_k : \mathbf{R} \to \mathbf{R}$ by $f_k(t) = f_k(0,t)$.

Note that $r_1(t)$ is clearly stationary in t; to see that $r_1(0) < \infty$ almost surely simply note that, with probability one, $B_{(s,0)} + B^{(1)}_{(s,0)} + ms < ms/2$ for all s sufficiently negative (by Strassen's law of the iterated logarithm, for example).

In fact, $r_1(0)$ has the same law as $-\log Z_m$, where Z_m is gamma-distributed with parameter m: this is Dufresne's identity [7,8].

We first state the quasi-reversibility property, as presented in [14].

Theorem 2.1.

1. *f_1 and g are independent standard Brownian motions indexed by* $\mathbf{R}$.
2. *For each $t \in \mathbf{R}$, $\{(f_1(s), g(s)), -\infty < s \le t\}$ is independent of $\{r_1(s), s \ge t\}$.*

It follows from Theorem 2.1 that $r_1(0), r_2(0), \ldots$ is a sequence of i.i.d. random variables, each distributed as $-\log Z_m$. By construction, we have

$$\sum_{k=1}^{n} r_k(0) = \log \Big[\int_{-\infty}^{0} du \, \exp(B_{(u,0)} + mu) \times \int_{u<s_1<\ldots<s_{n-1}<0} ds_1 \ldots ds_{n-1} \, \exp\{B^{(1)}_{(u,s_1)} + \ldots + B^{(n)}_{(s_{n-1},0)}\} \Big]. \tag{2.1}$$

Applying the strong law of large numbers, it can be deduced (see [14] for details) that:

Theorem 2.2. *For each $m > 0$:*

$$\lim_{n\to\infty} \frac{1}{n} \log \int_{-\infty<u<s_1<\ldots<s_{n-1}<0} du \, ds_1 \ldots ds_{n-1} \times \exp\{mu + B^{(1)}_{(u,s_1)} + \ldots + B^{(n)}_{(s_{n-1},0)}\} = -\Psi(m)$$

almost surely, where

$$\Psi(m) = \mathsf{E} \log Z_m = \Gamma'(m)/\Gamma(m)$$

is the digamma function (and Γ is the Gamma function).

Theorem 2.2 can be interpreted as follows. Let $\mathcal{B}$ denote the σ-field generated by the Brownian motions $B^{(1)}$, $B^{(2)}, \ldots$, and let $\tau_1, \tau_2, \ldots$ be the points of a unit-rate Poisson process on $\mathbf{R}_+$, independent of $\mathcal{B}$. For $t_0, t_1, \ldots, t_n \in \mathbf{R}$ define

$$\begin{aligned} E_n(t_0, t_1, \ldots, t_n) &= B^{(1)}_{(t_0,t_1)} + \ldots + B^{(n)}_{(t_{n-1},t_n)}, \\ F_n^m(t_0, t_1, \ldots, t_{n-m+1}) &= \exp\big(B^{(n)}_{(t_0,t_1)} + \ldots + B^{(m)}_{(t_{n-m},t_{n-m+1})}\big), \\ F_n &= F_n^1. \end{aligned}$$

By Brownian scaling, Theorem 2.2 is equivalent to:

Theorem 2.3. *For $\theta \neq 0$,*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathsf{E}\left[\exp(\theta E_n(0,\tau_1,\dots,\tau_n)) \mid \mathcal{B}\right] = -2\log|\theta| - \Psi(1/\theta^2),$$

almost surely.

Thus, if we set

$$\Lambda(\theta) = \begin{cases} -2\log|\theta| - \Psi(1/\theta^2), & \theta \neq 0, \\ 0, & \theta = 0, \end{cases}$$

we have

$$\lim_{n\to\infty} \frac{1}{n} \log \mathsf{E}\left[\exp(\theta E_n(0,\tau_1,\dots,\tau_n)) \mid \mathcal{B}\right] = \Lambda(\theta),$$

almost surely. From the asymptotic expansion

$$\Psi(x) \sim \log x - \frac{1}{2x} - \sum_{k=1}^{\infty} \frac{B_{2k}}{2kx^{2k}} \tag{2.2}$$

as $x \to \infty$ (see, for example, [1]), we have that Λ is finite and differentiable everywhere, with $\Lambda(0)=\Lambda'(0)=0$. It follows that the sequence $(1/n)E_n(0,\tau_1,\dots,\tau_n)$ satisfies the following conditional large deviation principle:

Theorem 2.4. *Given $\mathcal{B}$, $(1/n)E_n(0,\tau_1,\dots,\tau_n)$ satisfies a large deviation principle with good rate function*

$$\Lambda^*(x) = \sup_{\theta\in\mathbf{R}} [x\theta - \Lambda(\theta)]$$

almost surely.

This is a *quenched large deviation principle*, associated with the conditional law of large numbers. For example, Theorem 2.4 implies that given $\mathcal{B}$, $(1/n)E_n(0,\tau_1,\dots,\tau_n) \to 0$ almost surely. Another implication is that for any $x > 0$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathsf{P}\left(E_n(0,\tau_1,\dots,\tau_n) > xn \mid \mathcal{B}\right) = -\Lambda^*(x)$$

almost surely. For two other related large deviation principles see [14].

We will now describe how this relates to the Brownian directed polymer model. It is shown in Lemma 3.1 using Kingman's subadditive ergodic theorem that there exists a function $\gamma : \mathbf{R} \to \mathbf{R}$ such that given $x < 0$,

$$\lim_{n\to\infty} \frac{1}{n} \log \int_{xn<s_1<\dots<s_{n-1}<0} F_n(xn, s_1,\dots,s_{n-1},0)\, ds_1 \dots ds_{n-1} = \gamma(x) \tag{2.3}$$

almost surely, and it is shown in Lemma 3.4 that γ is a concave function on $(-\infty, 0)$. Therefore, by (2.1), Theorem 2.2 and Laplace's method, we would expect

$$\begin{aligned}-\Psi(m) &= \sup_{x<0}\,[mx + \gamma(x)] \\ &= (-\gamma)^*(m)\end{aligned}$$

and hence by inversion, $\gamma = -(-\Psi)^*$. The free energy density for our model of a directed polymer in a Brownian environment in $1+1$ dimensions, defined in (1.2), can then be expressed in terms of the digamma function by first using the Brownian scaling property:

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n} \log Z_n(\beta) &= \gamma(-\beta^2) - 2\log|\beta| \\ &= -(-\Psi)^*(-\beta^2) - 2\log|\beta|.\end{aligned}$$

The heuristic argument above is made rigorous by the following theorem, which is the main result of this paper.

Theorem 2.5. *Almost surely,*

$$\lim_{n\to\infty} \frac{1}{n} \log Z_n(\beta) = f(\beta)$$

where f is defined by (1.3), $\Psi(m) \equiv \Gamma'(m)/\Gamma(m)$ *is the restriction of the digamma function to* $(0,\infty)$, *and* $(-\Psi)^*$ *is the convex dual of the function* $-\Psi$.

3. Proofs

We begin by defining for $t \geq 0$ and $x < 0$

$$L_n(t) = \sup_{0\leq s_1 \leq \ldots \leq s_{n-1} \leq t} E_n(0, s_1, \ldots, s_{n-1}, t), \tag{3.1}$$

$$k_{m,n}(x) = \log \int\limits_{\substack{xn<s_1<\ldots \\ \ldots<s_{n-m}<xm}} F_{n+1}^{m+1}(xn, s_1, \ldots, s_{n-m}, xm)\, ds_1 \ldots ds_{n-m},$$

$$\mathcal{Z}_n(x) = \int\limits_{xn<s_1<\ldots<s_{n-1}<0} F_n(xn, s_1, \ldots, s_{n-1}, 0)\, ds_1 \ldots ds_{n-1}, \tag{3.2}$$

$$\gamma_n(x) = \frac{1}{n} \log \mathcal{Z}_n(x). \tag{3.3}$$

and recording the following lemma:

Lemma 3.1.

1. *There exists a constant* $c \in \mathbf{R}$ *such that* $(1/n)L_n(nt) \to c\sqrt{t}$ *almost surely and in expectation.*
2. *There exists a function* $\gamma : (-\infty, 0) \to \mathbf{R}$ *such that given* $x < 0$,
$$k_{0,n}(x)/n \to \gamma(x) \tag{3.4}$$
almost surely.
3. *The function* γ *is continuous on* $(-\infty, 0)$.
4. *Given* $x < 0$, $\lim_{n\to\infty} \gamma_n(x) = \gamma(x)$ *almost surely.*

Proof. The first part of the lemma follows from Brownian scaling and Kingman's subadditive ergodic theorem (see [11] for details). For the second part, observe that

$$\begin{aligned} k_{0,m}(x) + k_{m,n}(x) &= \log \int\limits_{\substack{xn<s_1<\ldots<s_n<0 \\ s_{n-m}<xm<s_{n-m+1}}} F_{n+1}(xn, s_1, \ldots, s_n, 0) ds_1 \ldots ds_n \\ &\le k_{0,n}(x) \end{aligned}$$

and k is therefore superadditive for fixed x. By construction we have the required conditions for Kingman's subadditive ergodic theorem, and so we may define a function $\gamma : (-\infty, 0) \to \mathbf{R}$ by (3.4). For the third part we have by Brownian scaling, for $x < 0$ and $\delta > x$

$$\begin{aligned} k_{0,n}(x-\delta) &=_d \log \int\limits_{\substack{0<s_1<\ldots \\ \ldots<s_n<n}} \exp\left(\sqrt{-x+\delta} E_{n+1}(0, s_1, \ldots, s_n, n)\right) ds_1 \ldots ds_n \\ &\qquad + n\log(-x+\delta) \\ &\le \log \int\limits_{\substack{0<s_1<\ldots \\ \ldots<s_n<n}} \exp\left(\sqrt{-x} E_{n+1}(0, s_1, \ldots, s_n, n)\right) ds_1 \ldots ds_n \\ &\qquad + (\sqrt{-x+\delta} - \sqrt{-x}) L_{n+1}(n) + n\log(-x+\delta) \\ &=_d k_{0,n}(x) + (\sqrt{-x+\delta} - \sqrt{-x}) L_{n+1}(n) \\ &\qquad - n\log(-x) + n\log(-x+\delta), \end{aligned} \tag{3.5}$$

where $=_d$ denotes equality in distribution. Let c be the constant in the first part of this lemma; then by Brownian scaling and Slutsky's Lemma (see for example [9]), $(1/n)L_{n+1}(n) \Rightarrow c$. By the second part of this lemma and (3.5),

$$\gamma(x-\delta) \le \gamma(x) + c(\sqrt{-x+\delta} - \sqrt{-x}) - \log(-x) + \log(-x+\delta).$$

By symmetry,

$$V_n := \inf_{0\le s_1\le\ldots\le s_{n-1}\le n} E_n(0,s_1,\ldots,s_{n-1},n) =_d -L_n(n) \tag{3.6}$$

and so we have similarly

$$\gamma(x-\delta) \ge \gamma(x) - c(\sqrt{-x+\delta}-\sqrt{-x}) - \log(-x) + \log(-x+\delta).$$

So

$$|\gamma(x-\delta)-\gamma(x)| \le c|\sqrt{-x+\delta}-\sqrt{-x}| + |\log(-x+\delta)-\log(-x)|. \tag{3.7}$$

For the fourth part we have

$$\begin{aligned} k_{0,n}(x) &\le \log\left(e^{2\overline{B}(n,0,x)} \int\limits_{xn<s_1<\ldots<s_n<0} F_n(xn,s_2,\ldots,s_n,0)\,ds_1\ldots ds_n\right) \\ &\le \log\left(xne^{2\overline{B}(n,0,x)} \int\limits_{xn<s_2<\ldots<s_n<0} F_n(xn,s_2,\ldots,s_n,0)\,ds_2\ldots ds_n\right) \\ &= \log\left(xne^{2\overline{B}(n,0,x)}\mathcal{Z}_n(x)\right) \end{aligned}$$

where for $y\le x\le 0$,

$$\overline{B}(n,x,y) = \max_{i=n-1,n,n+1} \sup_{yn\le r<s\le xn} |B^{(i)}(r,s)|. \tag{3.8}$$

From Borell's inequality and the Borel−Cantelli Lemma, there exists a null set $\mathcal{N}$ such that on its complement $\mathcal{N}^c$, $\overline{B}(n,x,y)/n\to 0$ for all x,y, so

$$\gamma(x) \le \liminf_{n\to\infty}\gamma_n(x) \quad \text{a.s.}$$

Now for $\varepsilon>0$,

$$\begin{aligned} k_{0,n}(x+\varepsilon) &\ge \log\left(e^{-2\overline{B}(n,0,x+\varepsilon)} \int\limits_{\substack{(x+\varepsilon)n<s_1<xn \\ xn<s_2<\ldots<s_n<0}} F_n(xn,s_2,\ldots,s_n,0)\,ds_1\ldots ds_n\right) \\ &= \log\left(\varepsilon ne^{-2\overline{B}(n,0,x+\varepsilon)}\mathcal{Z}_n(x)\right) \end{aligned}$$

so

$$\gamma(x+\varepsilon) \ge \limsup_{n\to\infty}\gamma_n(x) \quad \text{a.s.}$$

and the result follows by the third part of this lemma. □

Let $\mathbf{Q}_+ = \mathbf{Q}\cap(0,\infty)$, $\mathbf{Q}_- = -\mathbf{Q}_+$.

Lemma 3.2. *There exists a null set $\mathcal{M}$ such that the following statement holds on its complement $\mathcal{M}^c$: $\lim_{n\to\infty}\gamma_n(x)=\gamma(x)$ and $\liminf_{n\to\infty}\gamma_n(y)\ge\gamma(x)$ for every $x\in\mathbf{Q}_-$ and $y<x$.*

Proof. Let $x\in\mathbf{Q}_-$ and $y<x$. From (3.2) we have

$$\begin{aligned}\mathcal{Z}_n(y) &\ge \int\limits_{xn<s_1<\ldots<s_{n-1}<0} F_n(yn,s_1,\ldots,s_{n-1},0)\,ds_1\ldots ds_{n-1}\\ &=\exp\{B^{(n)}_{(yn,xn)}\}\mathcal{Z}_n(x).\end{aligned}\tag{3.9}$$

By Lemma 3.1 there exists a null set $\mathcal{N}_x$ such that on $\mathcal{N}_x^c$,

$$\lim_{n\to\infty}\gamma_n(x)=\gamma(x).$$

With $\mathcal{N}$ as in the proof of Lemma 3.1, let $\mathcal{M}=\mathcal{N}\cup\bigcup_{x\in\mathbf{Q}_-}\mathcal{N}_x$. □

Lemma 3.3. *Almost surely, $\lim_{n\to\infty}\gamma_n(x)=\gamma(x)$ for all $x<0$.*

Proof. Choose $x,y\in\mathbf{Q}_-$ with $y<x$. Then if $z\in[y,x]$,

$$\begin{aligned}\mathcal{Z}_n(z) &= \int\limits_{zn<s_1<\ldots<s_{n-1}<0} F_n(zn,s_1,\ldots,s_{n-1},0)\,ds_1\ldots ds_{n-1}\\ &=I_1(n,z)+I_2(n,z)\end{aligned}$$

where

$$\begin{aligned}I_1(n,z)&=\exp\{B^{(n)}_{(zn,xn)}\}\mathcal{Z}_n(x)\\ I_2(n,z)&=\int\limits_{s_1=zn}^{xn}\ \int\limits_{s_1<\ldots<s_{n-1}<0} F_{n-1}(s_1,\ldots,s_{n-1},0)\,ds_2\ldots ds_{n-1}\\ &\quad\times\exp\{B^{(n)}_{(zn,s_1)}\}\,ds_1.\end{aligned}$$

Now $I_1(n,z)\le e^{\overline{B}(n,x,y)}\mathcal{Z}_n(x)$ and

$$I_2(n,z)\le(x-z)ne^{2\overline{B}(n,x,y)}\mathcal{Z}_{n-1}\Big(\frac{ny}{n-1}\Big),$$

therefore

$$\sup_{z\in[y,x]}\mathcal{Z}_n(z)\le e^{\overline{B}(n,x,y)}\mathcal{Z}_n(x)+(x-y)ne^{2\overline{B}(n,x,y)}\mathcal{Z}_{n-1}\Big(\frac{ny}{n-1}\Big).\tag{3.10}$$

Take $\mathcal{M}$ as in the proof of Lemma 3.2. Now

$$\mathcal{Z}_{n-1}\Big(\frac{ny}{n-1}\Big) \geq e^{-\overline{B}(n,0,y)}\mathcal{Z}_{n-1}(y),$$

so $\liminf_{n\to\infty}(1/n)\log\mathcal{Z}_{n-1}(ny/n-1) \geq \gamma(y)$ on $\mathcal{M}^c$; and if $\varepsilon \in \mathbf{Q}_-$ then

$$\begin{aligned}\mathcal{Z}_n(y+\varepsilon) &\geq \int\limits_{\substack{(y+\varepsilon)n<s_1<yn\\ yn<s_2<\ldots<s_{n-1}<0}} F_{n-1}(yn, s_2, \ldots, s_{n-1}, 0)\, ds_1 \ldots ds_{n-1}\, e^{-2\overline{B}(n,0,y+\varepsilon)}\\ &= (-\varepsilon)ne^{-2\overline{B}(n,0,y+\varepsilon)}\mathcal{Z}_{n-1}\Big(\frac{ny}{n-1}\Big)\end{aligned}$$

hence

$$\gamma(y+\varepsilon) \geq \limsup_{n\to\infty}\frac{1}{n}\log\mathcal{Z}_{n-1}\Big(\frac{ny}{n-1}\Big)$$

on $\mathcal{M}^c$, and letting $\varepsilon \to 0$ in $\mathbf{Q}$ gives $\lim_{n\to\infty}(1/n)\log\mathcal{Z}_{n-1}(ny/n-1) = \gamma(y)$ on $\mathcal{M}^c$. Then by (3.10) and the proof of Lemma 3.2, on $\mathcal{M}^c$

$$\begin{aligned}\gamma(x) &\leq \liminf_{n\to\infty}\frac{1}{n}\log\inf_{z\in[y,x]}\mathcal{Z}_n(z) \leq \limsup_{n\to\infty}\frac{1}{n}\log\sup_{z\in[y,x]}\mathcal{Z}_n(z)\\ &\leq \limsup_{n\to\infty}\frac{1}{n}\log e^{\overline{B}(n,x,y)}\mathcal{Z}_n(x) \vee \limsup_{n\to\infty}\frac{1}{n}\log(x-y)\\ &\quad\times ne^{2\overline{B}(n,x,y)}\mathcal{Z}_{n-1}\Big(\frac{ny}{n-1}\Big)\\ &= \gamma(x) \vee \gamma(y) = \gamma(y).\end{aligned} \tag{3.11}$$

The result now follows from (3.7). □

Lemma 3.4. *The function γ is concave.*

Proof. For $x, y < 0$ and $\alpha \in (0,1)$,

$$\gamma_n(\alpha y + (1-\alpha)x) \geq \frac{[\alpha n]}{n}G_n + \frac{k_n}{n}H_n \tag{3.12}$$

where

$$\begin{aligned}G_n &= \frac{1}{[\alpha n]}\log\int\limits_{\substack{(\alpha y+(1-\alpha)x)n<s_1<\ldots\\ \ldots<s_{[\alpha n]-1}<(1-\alpha)xn}} F_n^{k_n}\big((\alpha y+(1-\alpha)x)n, s_1, \ldots\\ &\qquad \ldots, s_{[\alpha n]-1}, (1-\alpha)xn\big)\, ds_1 \ldots ds_{[\alpha n]-1}\\ k_n &= n - [\alpha n] + 1\\ H_n &= \frac{1}{k_n}\log\int\limits_{\substack{(1-\alpha)xn<s_{[\alpha n]}<\ldots\\ \ldots<s_{n-1}<0}} F_{k_n}\big((1-\alpha)xn, s_{[\alpha n]}, \ldots, s_{n-1}, 0\big)\, ds_{[\alpha n]} \ldots ds_{n-1}.\end{aligned}$$

Now

$$G_n =_d \gamma_{[\alpha n]}\Big(\frac{\alpha n}{[\alpha n]}y\Big),$$
$$H_n = \gamma_{k_n}\Big(\frac{(1-\alpha)n}{k_n}x\Big).$$

Choose $w, u \in \mathbf{Q}_-$ with $w < x < u$, and choose $\varepsilon, \delta > 0$. Then there exists $n_0 \in \mathbf{N}$ such that for all $n \geq n_0$, $w < ((1-\alpha)n/k_n)x < u$. Also, by (3.11) there exists $n_1 \in \mathbf{N}$ such that for any $n \geq n_1$,

$$\mathsf{P}\left(\gamma(u) - \varepsilon \leq \frac{1}{n}\log \inf_{z\in[w,u]} \mathcal{Z}_n(z) \leq \frac{1}{n}\log \sup_{z\in[w,u]} \mathcal{Z}_n(z) \leq \gamma(w) + \varepsilon\right) > 1-\delta.$$

Therefore if $k_n \geq n_0 \vee n_1$,

$$\mathsf{P}\left(\gamma(u) - \varepsilon \leq H_n \leq \gamma(w) + \varepsilon\right) > 1-\delta$$

and therefore $H_n \Rightarrow \gamma(x)$. Similarly $G_n \Rightarrow \gamma(y)$, and hence by Slutsky's Theorem

$$\frac{[\alpha n]}{n}G_n + \frac{k_n}{n}H_n \Longrightarrow \alpha\gamma(y) + (1-\alpha)\gamma(x).$$

Hence by (3.12),

$$\gamma(\alpha y + (1-\alpha)x) \geq \alpha\gamma(y) + (1-\alpha)\gamma(x).$$

□

Proof of Theorem 2.5. Choose $m > 0$. Define a probability density function on $(-\infty, 0)$ by

$$\mathsf{K}_n^{(m)}(dx) = \frac{1}{\Xi_n(m)} \exp n(mx + \gamma_n(x))\, dx$$

where

$$\Xi_n(m) = \int_{-\infty}^{0} \exp n(mx + \gamma_n(x))\, dx.$$

In the nomenclature of statistical physics, $\mathsf{K}_n^{(m)}$ is the *Kac density* and $\Xi_n(m)$ is the *grandcanonical partition function*. Theorem 2.2 says that given $\theta > -m$, the convergences

$$\frac{1}{n}\log \Xi_n(m) \longrightarrow -\Psi(m), \tag{3.13}$$

$$\frac{1}{n}\log \int_{-\infty}^{0} e^{n\theta x}\, \mathsf{K}_n^{(m)}(dx) \longrightarrow \Psi(m) - \Psi(m+\theta) =: \Lambda_m(\theta)$$

hold almost surely as $n \to \infty$. Choose $\varepsilon > 0$; then since $\Lambda_m^*(\Lambda_m'(0)+\varepsilon)$ and $\Lambda_m^*(\Lambda_m'(0)-\varepsilon)$ are strictly positive, we can apply the Chernoff bound to give $\mathsf{K}_n^{(m)}((\Lambda_m'(0)-\varepsilon, \Lambda_m'(0)+\varepsilon)) \to 1$ almost surely as $n \to \infty$. Letting $\varepsilon \to 0$ in $\mathbf{Q}$ gives that $\mathsf{K}_n^{(m)}$ is almost surely concentrated on $\Lambda_m'(0) = -\Psi'(m)$ as $n \to \infty$. Therefore for any $x \in \mathbf{Q}_-$ and $\varepsilon > 0$, using (3.9) we have

$$\begin{aligned}
\Xi_n(m) &\geq \int_{x-\varepsilon}^{x} \exp n(my + \gamma_n(y))\, dy \\
&\geq \exp\{n(m(x-\varepsilon) + \gamma_n(x))\} \int_{x-\varepsilon}^{x} \exp\{B_{(yn,xn)}^{(n)}\}\, dy \\
&\geq \varepsilon \exp\left\{n(m(x-\varepsilon) + \gamma_n(x)) - \overline{B}(n, x, x-\varepsilon)\right\}.
\end{aligned}$$

Therefore by Lemma 3.3 and (3.13),

$$-\Psi(m) \geq m(x-\varepsilon) + \gamma(x)$$

and we may let $\varepsilon \to 0$ and appeal to the regularity of Ψ to conclude that for all $x < 0$

$$\begin{aligned}
\gamma(x) &\leq \inf_{m \in \mathrm{Q}_+} (-mx - \Psi(m)) \\
&= -(-\Psi)^*(x).
\end{aligned}$$

For the reverse inequality we note that for $x \in \mathbf{Q}_-$ and $\varepsilon \in (0, -x)$ we may choose $m > 0$ such that $-\Psi'(m) = x + \varepsilon$ (see for example [1,2]). Then by (3.9),

$$\begin{aligned}
\mathsf{K}_n^{(m)}((x, x+2\varepsilon)) &= \frac{1}{\Xi_n(m)} \int_{x}^{x+2\varepsilon} \exp\{n(my + \gamma_n(y))\}\, dy \\
&\leq \frac{2\varepsilon}{\Xi_n(m)} \exp\left\{n(m(x+2\varepsilon) + \gamma_n(x)) - \overline{B}(n, x+2\varepsilon, x)\right\}.
\end{aligned}$$

Therefore using Lemma 3.3,

$$\begin{aligned}
0 &= \lim_{n\to\infty} \frac{1}{n} \log \mathsf{K}_n^{(m)}((x, x+2\varepsilon)) \\
&\leq m(x+2\varepsilon) + \gamma(x) + \Psi(m)
\end{aligned}$$

almost surely, in which case we may let $\varepsilon \to 0$ to conclude that

$$\begin{aligned}
\gamma(x) &\geq \inf_{m>0} (-mx - \Psi(m)) \\
&= -(-\Psi)^*(x).
\end{aligned}$$

Therefore $\gamma(x) = -(-\Psi)^*(x)$. Finally (1.1), the last part of Lemma 3.1 and Brownian scaling allow us to conclude that almost surely,

$$\lim_{n\to\infty} \frac{1}{n} \log Z_n(\beta) = \begin{cases} \gamma(-\beta^2) - 2\log|\beta| & : \quad \beta \neq 0 \\ 1 & : \quad \beta = 0 \end{cases}$$

as required. □

4. Analyticity of the free energy density and a large deviation principle

Theorem 4.1. *The function f defined in* (1.3) *is analytic and strictly convex on* $\mathbf{R}$, $f'(0) = 0$, *and* $\lim_{\beta\to\infty} f(\beta)/\beta = 2$.

Proof. For $x < 0$ we have

$$(-\Psi)^*(x) = \sup_{\theta>0} [x\theta + \Psi(\theta)]$$

Denoting by Ψ_n the nth derivative of the function Ψ, and noting that Ψ_2 is strictly negative everywhere (see for example [2]), we have

$$(-\Psi)^*(x) = x\Psi_1^{-1}(-x) + \Psi(\Psi_1^{-1}(-x))$$

and since Ψ_1 is an invertible analytic function with nonzero derivative, its inverse is analytic. Therefore f is analytic everywhere except possibly at 0.

To investigate the behaviour of f near 0, let $a = \Psi_1^{-1}(\beta^2)$. Then

$$f(\beta) = a\Psi_1(a) - \Psi(a) - \log \Psi_1(a). \tag{4.1}$$

Now $a \to \infty$ as $\beta \to 0$, and from [1] we have

$$\Psi_1(x) \sim \frac{1}{x} + \frac{1}{2x^2} + \frac{1}{6x^3} \quad (x \to \infty), \tag{4.2}$$

therefore recalling (2.2),

$$f(\beta) = 1 + O(a^{-1}) \quad (\beta \to 0).$$

One further application of (4.2) to this expression gives $f'(0) = 0$.

Since $f - 1$ is an asymptotic logarithmic moment generating function (see the proof of Lemma 4.1 below), f is convex. The regularity of Ψ implies further that f is strictly convex.

The power series

$$\log \Gamma(1+z) = -\log(1+z) + z(1-\xi) + \sum_{n=2}^{\infty} (-1)^n [\zeta(n) - 1] z^n / n$$

where ξ is Euler's constant and ζ is the Riemann Zeta Function, is valid for $|z| < 2$ [1]; therefore, since $\Psi(z) = d \log \Gamma(z)/dz$, it may be differentiated to give power series for Ψ and Ψ_1 in a neighbourhood of the origin. Substituting in (4.1) and letting $a \to 0$, we obtain $\lim_{\beta\to\infty} f(\beta)/\beta = 2$. □

Lemma 4.1. *Let $\tau_1 \le \ldots \le \tau_{n-1}$ be the order statistics for $n-1$ independent random variables having the uniform distribution on the interval $[0, n]$. Almost surely, conditional on $\mathcal{B}$, the random variable $(1/n)E_n(0, \tau_1, \ldots, \tau_{n-1}, n)$ satisfies a large deviation principle with rate function $(f-1)^*$.*

Proof. Choose $\beta \in \mathbf{R}$; then by (1.1)

$$\mathsf{E}\left[\exp\{\beta E_n(0, \tau_1, \ldots, \tau_{n-1}, n)\} \mid \mathcal{B}\right] = \frac{(n-1)!}{n^{n-1}} Z_n(\beta)$$

hence by Stirling's formula and Theorem 2.5,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathsf{E}\left[\exp\{\beta E_n(0, \tau_1, \ldots, \tau_{n-1}, n)\} \mid \mathcal{B}\right] = f(\beta) - 1 \tag{4.3}$$

almost surely. Now if $\alpha < \nu$ then $Z_n(\alpha) \le \exp\{-(\nu - \alpha)V_n\} Z_n(\nu)$, where V_n was defined in (3.6); hence for $\alpha, \beta \in \mathbf{Q}$ with $\beta > \alpha$

$$\begin{aligned} f(\alpha) &\le \liminf_{n\to\infty} \frac{1}{n} \log \inf_{\nu\in(\alpha,\beta)} Z_n(\nu) + (\beta - \alpha)c \\ &\le \limsup_{n\to\infty} \frac{1}{n} \log \sup_{\nu\in(\alpha,\beta)} Z_n(\nu) + (\beta - \alpha)c \\ &\le f(\beta) + 2(\beta - \alpha)c \end{aligned}$$

almost surely, where c was defined in Lemma 3.1. Therefore by the continuity of f, there exists a null set $\mathcal{N}$ such that on $\mathcal{N}^c$, the convergence in (4.3) holds for all $\beta \in \mathbf{R}$. □

5. Connection with random matrices and a Brownian directed percolation problem

It was shown in [11] that with L_n defined as in (3.1), for each $t \ge 0$,

$$\lim_{n\to\infty} \frac{1}{n} L_n(nt) = 2\sqrt{t}$$

almost surely. In the notation of Lemma 3.1 this says that $c = 2$. This result can also be deduced from random matrix theory using the fact [3,10] that $L_n(1)$ has the same law as the largest eigenvalue of a $n \times n$ GUE random matrix. We note here that the present work allows us to deduce the inequality $c \geq 2$. Since $L_n(n) \geq E_n(0, t_1, \ldots, t_{n-1}, n)$ we have that given β,

$$\frac{1}{n} L_n(n) \geq \frac{1}{\beta} \frac{1}{n} \log \mathsf{E} \left[\exp\{\beta E_n(0, \tau_1, \ldots, \tau_{n-1}, n)\} \mid \mathcal{B} \right]$$

almost surely. Letting $n \to \infty$ and then $\beta \to \infty$, using (4.3) and Theorem 4.1 gives the required inequality.

Acknowledgement

Research supported in part by Science Foundation Ireland, Grant Number SFI 04/RP1/I512. The authors would also like to thank the anonymous referee for helpful suggestions which have led to a much improved presentation.

References

[1] M. Abramowitz and I.A. Stegun (eds.) (1992) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover Publications Inc., New York. Reprint of the 1972 edition.

[2] H. Alzer (2004) Sharp inequalities for the digamma and polygamma functions. *Forum Math.* **16** (2), 181–221.

[3] Yu. Baryshnikov (2001) GUEs and queues. *Probab. Theory and Relat. Fields* **119** (2), 256–274.

[4] Ph. Carmona and Y. Hu (2002) On the partition function of a directed polymer in a Gaussian random environment. *Probab. Theory and Relat. Fields* **124** (3), 431–457.

[5] F. Comets and N. Yoshida (2005) Brownian directed polymers in random environment. *Commun. Math. Phys.* **254** (2), 257–287.

[6] B. Derrida (1990) Directed polymers in a random medium. *Physica A* **163** (1), 71–84. Proceedings of the Statphys. 17 Conference, Rio de Janeiro, July, 1989.

[7] D. Dufresne (2001) An affine property of the reciprocal Asian option process. *Osaka J. Math.* **38** (2), 379–381.

[8] D. Dufresne (2001) The integral of geometric Brownian motion. *Adv. Appl. Probab.* **33** (1), 223–241.

[9] R. Durrett (1996) *Probability: Theory and Examples.* Duxbury Press, Belmont, CA.

[10] J. Gravner, C.A. Tracy and H. Widom (2001) Limit theorems for height fluctuations in a class of discrete space and time growth models. *J. Stat. Phys.* **102** (5-6), 1085–1132.

[11] B.M. Hambly, J.B. Martin and N. O'Connell (2002) Concentration results for a Brownian directed percolation problem. *Stoch. Process. Appl.* **102** (2), 207–220.

[12] H. Matsumoto and M. Yor (1999) A version of Pitman's $2M - X$ theorem for geometric Brownian motions. *C. R. Acad. Sci. Paris Sér.* I *Math.* **328** (11), 1067–1074.

[13] H. Matsumoto and M. Yor (2001) A relationship between Brownian motions with opposite drifts via certain enlargements of the Brownian filtration. *Osaka J. Math.* **38** (2), 383–398.

[14] N. O'Connell and M. Yor (2001) Brownian analogues of Burke's theorem. *Stoch. Process. Appl.* **96** (2), 285–304.