

## PINNING AND WETTING TRANSITION FOR (1+1)-DIMENSIONAL FIELDS WITH LAPLACIAN INTERACTION

BY FRANCESCO CARAVENNA<sup>1</sup> AND JEAN-DOMINIQUE DEUSCHEL

*Università degli Studi di Padova and Technische Universität Berlin*

We consider a random field  $\varphi : \{1, \dots, N\} \rightarrow \mathbb{R}$  as a model for a linear chain attracted to the defect line  $\varphi = 0$ , that is, the  $x$ -axis. The free law of the field is specified by the density  $\exp(-\sum_i V(\Delta\varphi_i))$  with respect to the Lebesgue measure on  $\mathbb{R}^N$ , where  $\Delta$  is the discrete Laplacian and we allow for a very large class of potentials  $V(\cdot)$ . The interaction with the defect line is introduced by giving the field a reward  $\varepsilon \geq 0$  each time it touches the  $x$ -axis. We call this model the *pinning model*. We consider a second model, the *wetting model*, in which, in addition to the pinning reward, the field is also constrained to stay nonnegative.

We show that both models undergo a phase transition as the intensity  $\varepsilon$  of the pinning reward varies: both in the pinning ( $a = p$ ) and in the wetting ( $a = w$ ) case, there exists a critical value  $\varepsilon_c^a$  such that when  $\varepsilon > \varepsilon_c^a$  the field touches the defect line a positive fraction of times (localization), while this does not happen for  $\varepsilon < \varepsilon_c^a$  (delocalization). The two critical values are nontrivial and distinct:  $0 < \varepsilon_c^p < \varepsilon_c^w < \infty$ , and they are the only nonanalyticity points of the respective free energies. For the pinning model the transition is of second order, hence the field at  $\varepsilon = \varepsilon_c^p$  is delocalized. On the other hand, the transition in the wetting model is of *first order* and for  $\varepsilon = \varepsilon_c^w$  the field is localized. The core of our approach is a Markov renewal theory description of the field.

### 1. Introduction and main results.

1.1. *Definition of the models.* We are going to define two distinct but related models for a (1+1)-dimensional random field. These models depend on a measurable function  $V(\cdot) : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ , the *potential*. We require that  $x \mapsto \exp(-V(x))$  is bounded and continuous and that  $\int_{\mathbb{R}} \exp(-V(x)) dx < \infty$ . Since a global shift on  $V(\cdot)$  is irrelevant for our purposes, we will actually impose the stronger condition

$$(1.1) \quad \int_{\mathbb{R}} e^{-V(x)} dx = 1.$$

---

Received March 2007; revised October 2007.

<sup>1</sup>Supported in part by the German Research Foundation–Research Group 718 during his stay at TU Berlin in January 2007.

*AMS 2000 subject classifications.* 60K35, 60F05, 82B41.

*Key words and phrases.* Pinning model, wetting model, phase transition, entropic repulsion, Markov renewal theory, local limit theorem, Perron–Frobenius theorem.

The last assumptions we make on  $V(\cdot)$  are that  $V(0) < \infty$ , that is,  $\exp(-V(0)) > 0$ , and that

$$(1.2) \quad \int_{\mathbb{R}} x^2 e^{-V(x)} dx =: \sigma^2 < \infty \quad \text{and} \quad \int_{\mathbb{R}} x e^{-V(x)} dx = 0.$$

A typical example is of course the Gaussian case  $V(x) = x^2/(2\sigma^2) + \log \sqrt{2\pi\sigma^2}$ , but we stress that we do not make any convexity assumption on  $V(\cdot)$ . Next we introduce the Hamiltonian  $\mathcal{H}_{[a,b]}(\varphi)$ , defined for  $a, b \in \mathbb{Z}$ , with  $b - a \geq 2$ , and for  $\varphi : \{a, \dots, b\} \rightarrow \mathbb{R}$  by

$$(1.3) \quad \mathcal{H}_{[a,b]}(\varphi) := \sum_{n=a+1}^{b-1} V(\Delta\varphi_n),$$

where  $\Delta$  denotes the discrete Laplacian:

$$(1.4) \quad \Delta\varphi_n := (\varphi_{n+1} - \varphi_n) - (\varphi_n - \varphi_{n-1}) = \varphi_{n+1} + \varphi_{n-1} - 2\varphi_n.$$

We are ready to introduce our first model, the *pinning model* (p-model for short)  $\mathbb{P}_{\varepsilon,N}^p$ , that is, the probability measure on  $\mathbb{R}^{N-1}$  defined by

$$(1.5) \quad \mathbb{P}_{\varepsilon,N}^p(d\varphi_1 \cdots d\varphi_{N-1}) := \frac{\exp(-\mathcal{H}_{[-1,N+1]}(\varphi))}{Z_{\varepsilon,N}^p} \prod_{i=1}^{N-1} (\varepsilon \delta_0(d\varphi_i) + d\varphi_i)$$

where  $N \in \mathbb{N}$ ,  $\varepsilon \geq 0$ ,  $d\varphi_i$  is the Lebesgue measure on  $\mathbb{R}$ ,  $\delta_0(\cdot)$  is the Dirac mass at zero and  $Z_{\varepsilon,N}^p$  is the normalization constant, usually called partition function. To complete the definition, in order to make sense of  $\mathcal{H}_{[-1,N+1]}(\varphi)$ , we have to specify:

$$(1.6) \quad \text{the boundary conditions } \varphi_{-1} = \varphi_0 = \varphi_N = \varphi_{N+1} := 0.$$

We fix zero boundary conditions for simplicity, but our approach works for arbitrary choices (as long as they are bounded in  $N$ ).

The second model we consider, the *wetting model* (w-model for short)  $\mathbb{P}_{\varepsilon,N}^w$ , is a variant of the pinning model defined by

$$(1.7) \quad \begin{aligned} \mathbb{P}_{\varepsilon,N}^w(d\varphi_1 \cdots d\varphi_{N-1}) &:= \mathbb{P}_{\varepsilon,N}^p(d\varphi_1 \cdots d\varphi_{N-1} | \varphi_1 \geq 0, \dots, \varphi_{N-1} \geq 0) \\ &= \frac{\exp(-\mathcal{H}_{[-1,N+1]}(\varphi))}{Z_{\varepsilon,N}^w} \prod_{i=1}^{N-1} (\varepsilon \delta_0(d\varphi_i) + d\varphi_i \mathbf{1}_{(\varphi_i \geq 0)}), \end{aligned}$$

that is, we replace the measure  $d\varphi_i$  by  $d\varphi_i \mathbf{1}_{(\varphi_i \geq 0)}$  and  $Z_{\varepsilon,N}^p$  by a new normalization  $Z_{\varepsilon,N}^w$ .

Both  $\mathbb{P}_{\varepsilon,N}^p$  and  $\mathbb{P}_{\varepsilon,N}^w$  are (1+1)-dimensional models for a linear chain of length  $N$  which is attracted to a defect line, the  $x$ -axis, and the parameter  $\varepsilon \geq 0$  tunes the strength of the attraction. By “(1+1)-dimensional” we mean that the configurations of the linear chain are described by the trajectories  $\{(i, \varphi_i)\}_{0 \leq i \leq N}$  of the

field, so that we are dealing with directed models (see Figure 1 for a graphical representation). We point out that linear chain models with Laplacian interaction appear naturally in the physical literature in the context of *semiflexible polymers*; cf. [6, 17, 21] (however, the scaling they consider is different from the one we look at in this paper). One note about the terminology: while “pinning” refers of course to the attraction terms  $\varepsilon\delta_0(d\varphi_i)$ , the use of the term “wetting” is somewhat customary in the presence of a positivity constraint and refers to the interpretation of the field as an effective model for the interface of separation between a liquid above a wall and a gas; see [10] for more details.

The purpose of this paper is to investigate the behavior of  $\mathbb{P}_{\varepsilon,N}^{\text{p}}$  and  $\mathbb{P}_{\varepsilon,N}^{\text{w}}$  in the large  $N$  limit; in particular we wish to understand whether and when the reward  $\varepsilon \geq 0$  is strong enough to pin the chain at the defect line, a phenomenon that we will call *localization*. We point out that these kinds of questions have been answered in depth in the case of gradient interaction, that is, when the Laplacian  $\Delta$  appearing in (1.3) is replaced by the discrete gradient  $\nabla\varphi_n := \varphi_n - \varphi_{n-1}$  (cf. [2, 9–11, 13,

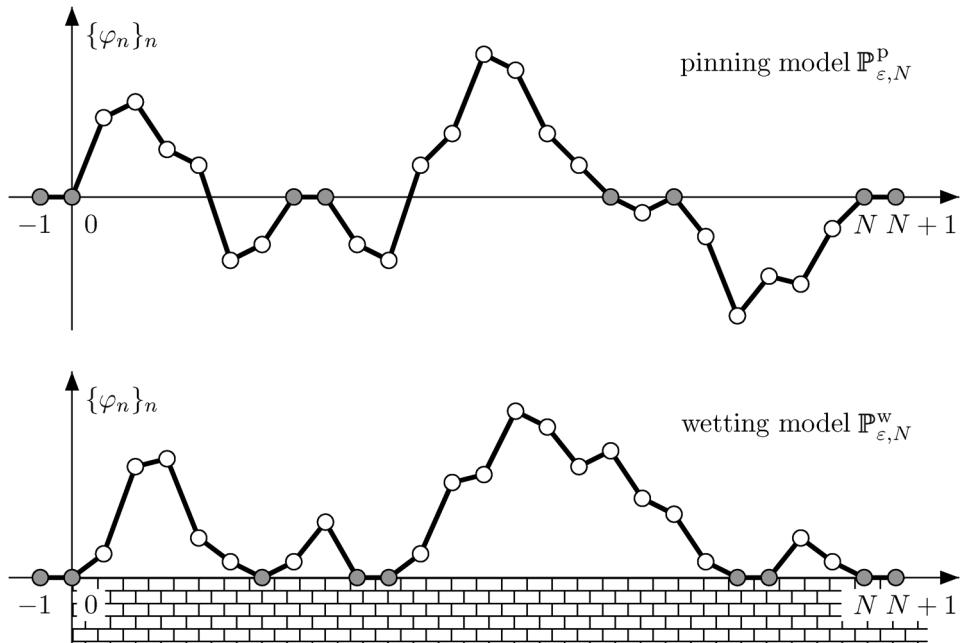


FIG. 1. A graphical representation of the pinning model  $\mathbb{P}_{\varepsilon,N}^{\text{p}}$  (top) and of the wetting model  $\mathbb{P}_{\varepsilon,N}^{\text{w}}$  (bottom), for  $N = 25$  and  $\varepsilon > 0$ . The trajectories  $\{(n, \varphi_n)\}_{0 \leq n \leq N}$  of the field describe the configurations of a linear chain attracted to a defect line, the  $x$ -axis. The gray circles represent the pinned sites, that is, the points in which the chain touches the defect line, which are energetically favored. Note that in the pinning case the chain can cross the defect line without touching it, while this does not happen in the wetting case due to the presence of a wall, that is, of a constraint for the chain to stay nonnegative; the repulsion effect of entropic nature that arises is responsible for the different critical behavior of the models.

[18]); we will refer to this as the *gradient case*. As we are going to see, the behavior in the Laplacian case turns out to be sensibly different.

**1.2. The free energy and the main results.** A convenient way to define localization for our models is by looking at the Laplace asymptotic behavior of the partition function  $Z_{\varepsilon,N}^a$  as  $N \rightarrow \infty$ . More precisely, for  $a \in \{p, w\}$  we define the *free energy*  $F^a(\varepsilon)$  by

$$(1.8) \quad F^a(\varepsilon) := \lim_{N \rightarrow \infty} F_N^a(\varepsilon), \quad F_N^a(\varepsilon) := \frac{1}{N} \log Z_{\varepsilon,N}^a,$$

where the existence of this limit (that will follow as a by-product of our approach) can be proven with a standard superadditivity argument. The basic observation is that the free energy is nonnegative. In fact, setting  $\Omega^w := [0, \infty)$  and  $\Omega^p := \mathbb{R}$ , we have  $\forall N \in \mathbb{N}$

$$(1.9) \quad \begin{aligned} Z_{\varepsilon,N}^a &= \int \exp(-\mathcal{H}_{[-1,N+1]}(\varphi)) \prod_{i=1}^{N-1} (\varepsilon \delta_0(d\varphi_i) + d\varphi_i \mathbf{1}_{(\varphi_i \in \Omega^a)}) \\ &\geq \int \exp(-\mathcal{H}_{[-1,N+1]}(\varphi)) \prod_{i=1}^{N-1} d\varphi_i \mathbf{1}_{(\varphi_i \in \Omega^a)} = Z_{0,N}^a \geq \frac{c_1}{N^{c_2}}, \end{aligned}$$

where  $c_1, c_2$  are positive constants and the polynomial bound for  $Z_{0,N}^a$  (analogous to what happens in the gradient case; cf. [10]) is proven in (2.14). Therefore  $F^a(\varepsilon) \geq F^a(0) = 0$  for every  $\varepsilon \geq 0$ . Since this lower bound has been obtained by ignoring the contribution of the paths that touch the defect line, one is led to the following

**DEFINITION 1.1.** For  $a \in \{p, w\}$ , the  $a$ -model  $\{\mathbb{P}_{\varepsilon,N}^a\}_N$  is said to be localized if  $F^a(\varepsilon) > 0$ .

The first problem is to understand for which values of  $\varepsilon$  (if any) there is localization. Some considerations can be drawn easily. We introduce for convenience for  $t \in \mathbb{R}$

$$(1.10) \quad \tilde{F}_N^a(t) := F_N^a(e^t), \quad \tilde{F}^a(t) := F^a(e^t).$$

It is easy to show (see Appendix A) that  $\tilde{F}_N^a(\cdot)$  is convex, therefore also  $\tilde{F}^a(\cdot)$  is convex. In particular, the free energy  $F^a(\varepsilon) = \tilde{F}^a(\log \varepsilon)$  is a *continuous function*, as long as it is finite.  $F^a(\cdot)$  is also *nondecreasing*, because  $Z_{\varepsilon,N}^a$  is increasing in  $\varepsilon$  [cf. the first line of (1.9)]. This observation implies that, for both  $a \in \{p, w\}$ , there is a critical value  $\varepsilon_c^a \in [0, \infty]$  such that the  $a$ -model is localized if and only if  $\varepsilon > \varepsilon_c^a$ . Moreover  $\varepsilon_c^p \leq \varepsilon_c^w$ , since  $Z_{\varepsilon,N}^p \geq Z_{\varepsilon,N}^w$ .

However, it is still not clear that a phase transition really exists, that is, that  $\varepsilon_c^a \in (0, \infty)$ . Indeed, in the gradient case the transition is nontrivial only for the

wetting model, that is,  $0 < \varepsilon_c^{w,\nabla} < \infty$  while  $\varepsilon_c^{p,\nabla} = 0$ ; cf. [10, 13]. Our first theorem shows that in the Laplacian case both the pinning and the wetting models undergo a nontrivial transition, and gives further properties of the free energy  $F^a(\cdot)$ .

**THEOREM 1.2** (Localization transition). *The following relations hold:*

$$\varepsilon_c^p \in (0, \infty), \quad \varepsilon_c^w \in (0, \infty), \quad \varepsilon_c^p < \varepsilon_c^w.$$

We have  $F^a(\varepsilon) = 0$  for  $\varepsilon \in [0, \varepsilon_c^a]$ , while  $0 < F^a(\varepsilon) < \infty$  for  $\varepsilon \in (\varepsilon_c^a, \infty)$ , and as  $\varepsilon \rightarrow \infty$

$$(1.11) \quad F^a(\varepsilon) = \log \varepsilon (1 + o(1)), \quad a \in \{p, w\}.$$

Moreover the function  $F^a(\varepsilon)$  is real analytic on  $(\varepsilon_c^a, \infty)$ .

One may ask why in the Laplacian case we have  $\varepsilon_c^p > 0$ , unlike in the gradient case. Heuristically, we could say that the Laplacian interaction (1.3) describes a stiffer chain, more rigid to bending with respect to the gradient interaction, and therefore Laplacian models require a stronger reward in order to localize. Note in fact that in the Gaussian case  $V(x) = x^2/(2\sigma^2) + \log \sqrt{2\pi\sigma^2}$  the ground state of the gradient interaction is just the horizontally flat line, whereas the Laplacian interaction favors rather *affine configurations*, penalizing curvature and bendings.

It is worth stressing that the free energy has a direct translation in terms of some path properties of the field. Defining the contact number  $\ell_N$  by

$$(1.12) \quad \ell_N := \#\{i \in \{1, \dots, N\} : \varphi_i = 0\},$$

a simple computation (see Appendix A) shows that for every  $\varepsilon > 0$  and  $N \in \mathbb{N}$

$$(1.13) \quad D_N^a(\varepsilon) := \mathbb{E}_{\varepsilon, N}^a \left( \frac{\ell_N}{N} \right) = (\tilde{F}_N^a)'(\log \varepsilon) = \varepsilon \cdot (F_N^a)'(\varepsilon).$$

Then, introducing the nonrandom quantity  $D^a(\varepsilon) := \varepsilon \cdot (F^a)'(\varepsilon)$  (which is well defined for  $\varepsilon \neq \varepsilon_c^a$  by Theorem 1.2), a simple convexity argument shows that  $D_N^a(\varepsilon) \rightarrow D^a(\varepsilon)$  as  $N \rightarrow \infty$ , for every  $\varepsilon \neq \varepsilon_c^a$ . Indeed much more can be said (see Appendix A):

- When  $\varepsilon > \varepsilon_c^a$  we have that  $D^a(\varepsilon) > 0$ , and for every  $\delta > 0$  and  $N \in \mathbb{N}$

$$(1.14) \quad \mathbb{P}_{\varepsilon, N}^a \left( \left| \frac{\ell_N}{N} - D^a(\varepsilon) \right| > \delta \right) \leq \exp(-c_3 N),$$

where  $c_3$  is a positive constant. This shows that, when the  $a$ -model is localized according to Definition 1.1, its typical paths touch the defect line a positive fraction of times, equal to  $D^a(\varepsilon)$ . Notice that, by (1.11) and convexity arguments,  $D^a(\varepsilon)$  converges to 1 as  $\varepsilon \rightarrow \infty$ , that is, a strong reward pins the field at the defect line in a very effective way (observe that  $\ell_N/N \leq 1$ ).

- On the other hand, when  $\varepsilon < \varepsilon_c^a$  we have  $D^a(\varepsilon) = 0$  and for every  $\delta > 0$  and  $N \in \mathbb{N}$

$$(1.15) \quad \mathbb{P}_{\varepsilon, N}^a \left( \frac{\ell_N}{N} > \delta \right) \leq \exp(-c_4 N),$$

where  $c_4$  is a positive constant. Thus for  $\varepsilon < \varepsilon_c^a$  the typical paths of the  $a$ -model touch the defect line only  $o(N)$  times; when this happens it is customary to say that the model is *delocalized*.

What is left out from this analysis is the critical regime  $\varepsilon = \varepsilon_c^a$ . The behavior of the model in this case is sharply linked to the way in which the free energy  $F^a(\varepsilon)$  vanishes as  $\varepsilon \downarrow \varepsilon_c^a$ . If  $F^a(\cdot)$  is differentiable also at  $\varepsilon = \varepsilon_c^a$  (transition of second or higher order), then  $(F^a)'(\varepsilon_c^a) = 0$  and relation 1.15 holds, that is, the  $a$ -model for  $\varepsilon = \varepsilon_c^a$  is delocalized. The other possibility is that  $F^a(\cdot)$  is not differentiable at  $\varepsilon = \varepsilon_c^a$  (transition of first order), which happens when the right-derivative is positive:  $(F^a)'_+(\varepsilon_c^a) > 0$ . In this case the behavior of  $\mathbb{P}_{\varepsilon, N}^a$  for large  $N$  may depend on the choice of the boundary conditions.

We first consider the critical regime for the wetting model, where the transition turns out to be of *first order*. Recall the definition 1.13 of  $D_N^a(\varepsilon)$ .

**THEOREM 1.3** (Critical wetting model). *For the wetting model we have*

$$(1.16) \quad \liminf_{N \rightarrow \infty} D_N^w(\varepsilon_c^w) > 0.$$

*Therefore  $(F^w)'_+(\varepsilon_c^w) > 0$  and the phase transition is of first order.*

Notice that (1.13) and (1.16) yield

$$\liminf_{N \rightarrow \infty} \mathbb{E}_{\varepsilon_c^w, N}^w \left( \frac{\ell_N}{N} \right) > 0,$$

and in this sense the wetting model at the critical point exhibits a *localized* behavior. This is in sharp contrast with the gradient case, where it is well known that the wetting model at criticality is delocalized and in fact the transition is of *second order*; cf. [9–11, 18]. The emergence of a first-order transition in the case of Laplacian interaction is interesting in view of the possible applications of  $\mathbb{P}_{\varepsilon, N}^w$  as a model for the *DNA denaturation transition*, where the nonnegative field  $\{\varphi_i\}_i$  describes the distance between the two DNA strands. In fact for the DNA denaturation something close to a first-order phase transition is experimentally observed; we refer to [13], Section 1.4, for a detailed discussion (cf. also [19, 27]).

Finally we consider the critical pinning model, where the transition is of second order.

**THEOREM 1.4** (Critical pinning model). *For the pinning model we have*

$$(1.17) \quad \limsup_{\varepsilon \downarrow \varepsilon_c^p} \limsup_{N \rightarrow \infty} D_N^p(\varepsilon) = 0.$$

Therefore  $F^p(\varepsilon)$  is differentiable at  $\varepsilon = \varepsilon_c^p$  and  $(F^p)'(\varepsilon_c^p) = 0$ . Moreover there exists  $c_5 > 0$  such that for  $\delta$  sufficiently small we have

$$(1.18) \quad F^p(\varepsilon_c^p + \delta, 0) \geq c_5 \frac{\delta}{\log 1/\delta},$$

that is, the transition is exactly of second order.

Although the relation  $(F^p)'(\varepsilon_c^p) = 0$  yields  $\ell_N = o(N)$ , in a delocalized fashion, the pinning model at  $\varepsilon = \varepsilon_c^p$  is actually somewhat borderline between localization and delocalization, as (1.18) suggests and as we point out in the next paragraph.

**1.3. Further path results.** A direct application of the techniques that we develop in this paper yields further path properties of the field. Let us introduce the *maximal gap*

$$\Delta_N := \max\{n \leq N : \varphi_{k+1} \neq 0, \varphi_{k+2} \neq 0, \dots, \varphi_{k+n} \neq 0 \text{ for some } k \leq N-n\}.$$

One can show that, for both  $a \in \{p, w\}$  and for  $\varepsilon > \varepsilon_c^a$ , the following relations hold:

$$(1.19) \quad \begin{aligned} \forall \delta > 0: \lim_{N \rightarrow \infty} \mathbb{P}_{\varepsilon, N}^a \left( \frac{\Delta_N}{N} \geq \delta \right) &= 0, \\ \lim_{L \rightarrow \infty} \limsup_{N \rightarrow \infty} \max_{i=1, \dots, N-1} \mathbb{P}_{\varepsilon, N}^a (|\varphi_i| \geq L) &= 0. \end{aligned}$$

In particular for  $\varepsilon > \varepsilon_c^a$  each component  $\varphi_i$  of the field is at finite distance from the defect line and this is a clear localization path statement. On the other hand, in the pinning case  $a = p$  we can strengthen (1.15) to the following relation: for every  $\varepsilon < \varepsilon_c^p$

$$(1.20) \quad \lim_{L \rightarrow \infty} \limsup_{N \rightarrow \infty} \mathbb{P}_{\varepsilon, N}^p (\Delta_N \leq N - L) = 0,$$

that is, for  $\varepsilon < \varepsilon_c^p$  the field touches the defect line at a finite number of sites, all at *finite distance* from the boundary points  $\{0, N\}$ . We expect that the same relation holds true also in the wetting case  $a = w$ , but at present we cannot prove it: what is missing are more precise estimates on the entropic repulsion problem; see Section 1.5 for a detailed discussion. It is interesting to note that we can prove that the first relation in (1.19) holds true also in the pinning case  $a = p$  at the critical point  $\varepsilon = \varepsilon_c^p$ , and this shows that the pinning model at criticality has also features of localized behavior.

We do not give an explicit proof of the above relations in this paper, both for conciseness and because in a second paper [7] we focus on the scaling limits of the pinning model, obtaining (de)localization path statements that are much more precise than (1.19) and (1.20) [under stronger assumptions on the potential  $V(\cdot)$ ]. We show in particular that for all  $\varepsilon \in (0, \varepsilon_c^p)$  the natural rescaling of  $\mathbb{P}_{\varepsilon, N}^p$  converges in distribution in  $C([0, 1])$  to the same limit that one obtains in the free case

$\varepsilon = 0$ , that is, the integral process of a Brownian bridge. On the other hand, for every  $\varepsilon \geq \varepsilon_c^p$  the natural rescaling of  $\mathbb{P}_{\varepsilon,N}^p$  yields the trivial process which is identically zero. We stress that  $\varepsilon = \varepsilon_c^p$  is included in the last statement; this is in sharp contrast with the gradient case, where both the pinning and the wetting models at criticality have a nontrivial scaling limit, respectively the Brownian bridge and the reflected Brownian bridge. This shows again the peculiarity of the critical pinning model in the Laplacian case. Indeed, by lowering the scaling constants with suitable logarithmic corrections, we are able to extract a nontrivial scaling limit (in a distributional sense) for the law  $\mathbb{P}_{\varepsilon_c^p,N}^p$  in terms of a symmetric stable Lévy process. We stress that the techniques and results of the present paper play a crucial role for [7].

**1.4. Outline of the paper: approach and techniques.** Although our main results are about the free energy, the core of our approach is a precise pathwise description of the field based on Markov renewal theory. In analogy to [10, 9] and especially to [8], we would like to stress the power of (Markov) renewal theory techniques for the study of  $(1 + 1)$ -dimensional linear chain models. The other basic techniques that we use are local limit theorems and an infinite-dimensional version of the Perron–Frobenius theorem. Let us describe more in detail the structure of the paper.

In Section 2 we study the pinning and wetting models in the free case  $\varepsilon = 0$ , showing that these models are sharply linked to the integral of a random walk. More precisely, let  $\{Y_n\}_{n \geq 0}$  denote a random walk starting at zero and with step law  $\mathbf{P}(Y_1 \in dx) = \exp(-V(x))dx$  [the walk has zero mean and finite variance by (1.2)] and let us denote by  $Z_n := Y_1 + \dots + Y_n$  the corresponding *integrated random walk process*. In Proposition 2.2 we show that the law  $\mathbb{P}_{0,N}^a$  is nothing but a bridge of length  $N$  of the process  $\{Z_n\}_n$ , with the further conditioning to stay nonnegative in the wetting case  $a = w$ . Therefore we focus on the asymptotic properties of the process  $\{Z_n\}_n$ , obtaining a basic local limit theorem (cf. Proposition 2.3) and some bounds for the probability that  $\{Z_n\}_n$  stays positive (connected to the problem of entropic repulsion that we discuss below; cf. Section 1.5).

In Section 3, which is the core of the paper, we show that for  $\varepsilon > 0$  the law  $\mathbb{P}_{\varepsilon,N}^a$  admits a crucial description in terms of Markov renewal theory. More precisely, we show that the zeros of the field  $\{i \leq N : \varphi_i = 0\}$  under  $\mathbb{P}_{\varepsilon,N}^a$  are distributed according to the law of a (hidden) Markov renewal process conditioned to hit  $\{N, N + 1\}$ ; cf. Proposition 3.1. We thus obtain an explicit expression for the partition function  $Z_{\varepsilon,N}^a$  in terms of this Markov renewal process, which is the key to our main results.

Section 4 is devoted to proving some analytical results that underlie the construction of the Markov renewal process appearing in Section 3. The main tool is an infinite-dimensional version of the classical Perron–Frobenius theorem (cf. [33]), and a basic role is played by the asymptotic estimates obtained in Section 2. A by-product of this analysis is an explicit formula [cf. 4.10], that links  $F^a(\cdot)$  and  $\varepsilon_c^a$  to the spectral radius of a suitable integral operator and that will be exploited later.

Sections 5, 6 and 7 contain the proofs of Theorems 1.2, 1.3 and 1.4, respectively. In view of the description given in Section 3, all the results to prove can be rephrased in the language of Markov renewal theory. The proofs are then carried out exploiting the asymptotic estimates derived in Sections 2 and 4 together with some algebraic manipulation of the kernel that gives the law of the hidden Markov renewal process. Finally, the Appendixes contain the proof of some technical results.

We conclude the introduction by discussing briefly two interesting problems that are linked to our models, namely the *entropic repulsion* in Section 1.5 and the *smoothing effect of disorder* in Section 1.6. Finally, Section 1.7 contains some recurrent notations, especially about kernels, used throughout the paper.

**1.5. Entropic repulsion.** We recall that  $(\{Y_n\}_{n \geq 0}, \mathbf{P})$  is the random walk with step  $\mathbf{P}(Y_1 \in dx) = e^{-V(x)} dx$  and that  $Z_n = Y_1 + \dots + Y_n$ . The analysis of the wetting model requires estimating the decay as  $N \rightarrow \infty$  of the probabilities  $\mathbf{P}(\Omega_N^+)$  and  $\mathbf{P}(\Omega_N^+ | Z_{N+1} = 0, Z_{N+2} = 0)$ , where we set  $\Omega_N^+ := \{Z_1 \geq 0, \dots, Z_N \geq 0\}$ . This type of problem is known in the literature as *entropic repulsion* and it has received a lot of attention; see [32] for a recent overview. In the Laplacian case that we consider here, this problem has been solved in the Gaussian setting (i.e., when  $V(x) = x^2/(2\sigma^2) + \log \sqrt{2\pi\sigma^2}$ ) in  $(d+1)$ -dimension with  $d \geq 5$ ; cf. [28, 22]. Little is known in the  $(1+1)$ -dimensional setting, apart from the following result of Sinai's [29] in the very special case when  $\{Y_n\}_n$  is the simple random walk on  $\mathbb{Z}$ :

$$(1.21) \quad \frac{c}{N^{1/4}} \leq \mathbf{P}(\Omega_N^+) \leq \frac{C}{N^{1/4}},$$

where  $c, C$  are positive constants. The proof of this bound relies on the exact combinatorial results available in the simple random walk case and it appears difficult to extend it to our situation. We point out that the same exponent  $1/4$  appears in related continuous models dealing with the integral of Brownian motion; cf. [23, 25]. Based on Sinai's result, which we believe to hold for general random walks with zero mean and finite variance, we expect that for the bridge case one should have the bound

$$(1.22) \quad \frac{c}{N^{1/2}} \leq \mathbf{P}(\Omega_N^+ | Z_{N+1} = 0, Z_{N+2} = 0) \leq \frac{C}{N^{1/2}}.$$

We cannot derive precise bounds as (1.21) and (1.22); however, for the purpose of this paper the following weaker result suffices:

**PROPOSITION 1.5.** *There exist constants  $c, C, c_- > 0$  and  $c_+ > 1$  such that  $\forall N \in \mathbb{N}$*

$$(1.23) \quad \frac{c}{N^{c_-}} \leq \mathbf{P}(\Omega_N^+) \leq \frac{C}{(\log N)^{c_+}},$$

$$(1.24) \quad \frac{c}{N^{c_-}} \leq \mathbf{P}(\Omega_N^+ | Z_{N+1} = 0, Z_{N+2} = 0) \leq \frac{C}{(\log N)^{c_+}}.$$

We prove this proposition in Appendix C. We stress that the most delicate point is the upper bound in (1.23): the idea of the proof is to dilute the system on super-exponentially spaced times. We also point out that in the Gaussian case, that is, when  $V(x) = x^2/(2\sigma^2) + \log \sqrt{2\pi\sigma^2}$ , the upper bound can be easily strengthened to  $\mathbf{P}(\Omega_N^+) \leq (\text{const.})/N^a$ , for some  $a > 0$ , by diluting the system on exponentially spaced times and using the results in the paper [24].

**1.6. Smoothing effect of quenched disorder.** The models we consider in this work are *homogeneous*, in that the reward  $\varepsilon$  is deterministic and constant. However, one can define in a natural way a disordered version of our model, where the reward is itself random and may vary from site to site. We stress that disordered pinning models have attracted a lot of attention recently; cf. [1, 2, 16, 30, 31] (see also [13] for an overview).

Let us be more precise: we take a sequence  $\omega = \{\omega_n\}_{n \in \mathbb{N}}$  of i.i.d. standard Gaussian random variables, defined on some probability space  $(\mathcal{S}, P)$ , and we denote by  $M(\beta) := E(\exp(\beta\omega_1)) = \exp(\beta^2/2)$  the corresponding moment generating function (we focus on Gaussian variables only for the sake of simplicity; we could more generally take variables with zero mean, unit variance and finite exponential moments). Then, given a realization  $\omega$ , we introduce for  $\beta \geq 0$  a *quenched* version  $\mathbb{P}_{\varepsilon, \beta, \omega, N}^a$  of our polymer measure, just replacing the reward  $\varepsilon$  by  $\varepsilon \exp(\beta\omega_i)$ :

$$\begin{aligned} \mathbb{P}_{\varepsilon, \beta, \omega, N}^a(d\varphi_1 \cdots d\varphi_{N-1}) \\ := \frac{\exp(-\mathcal{H}_{[-1, N+1]}(\varphi))}{Z_{\varepsilon, \beta, \omega, N}^a} \prod_{i=1}^{N-1} (\varepsilon e^{\beta\omega_i} \delta_0(d\varphi_i) + d\varphi_i \mathbf{1}_{\{\varphi_i \in \Omega^a\}}), \end{aligned}$$

where of course  $a \in \{p, w\}$  and  $\Omega^p = \mathbb{R}$  while  $\Omega^w = \mathbb{R}^+$ .

The analysis of the model goes along the same line as in the homogeneous case. Namely, the free energy  $F^a(\varepsilon, \beta)$  is defined by

$$F^a(\varepsilon, \beta) := \lim_{N \rightarrow \infty} F_N^a(\varepsilon, \beta, \omega), \quad F_N^a(\varepsilon, \beta, \omega) := \frac{1}{N} \log Z_{\varepsilon, \beta, \omega, N}^a,$$

where the existence of this limit, both  $P(d\omega)$ -a.s. and in  $L_1(dP)$ , follows by Kingman's Superadditive Ergodic Theorem (cf. [13], Section 4.4.2), and moreover  $F^a(\varepsilon, \beta)$  does not depend on  $\omega$ , that is, it is  $P(d\omega)$ -a.s. a constant (*self-averaging* of the free energy).

Restricting to the paths that do not touch the defect line, exactly like in (1.9), yields the lower bound  $F^a(\varepsilon, \beta) \geq 0$  for all  $\varepsilon, \beta$ . Therefore, as in the homogeneous case, we say that the  $a$ -model with parameters  $(\varepsilon, \beta)$  is *localized* if  $F^a(\varepsilon, \beta) > 0$ . By the monotonicity in  $\varepsilon$  it follows that, for every value of  $\beta \geq 0$ , there is a critical value  $\varepsilon_c^a(\beta) \in [0, \infty]$  such that the  $a$ -model is localized if and

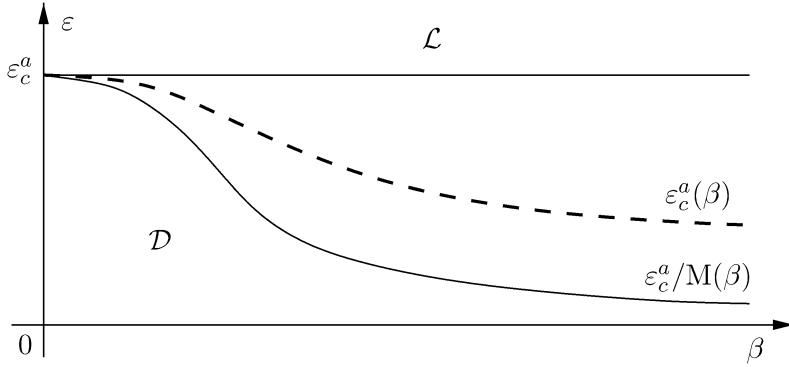


FIG. 2. The critical value  $\varepsilon_c^a(\beta)$  as a function of  $\beta$ , with the upper and lower bound given in (1.25).  $\mathcal{L} = \{(\beta, \varepsilon) : \varepsilon > \varepsilon_c^a(\beta)\}$  and  $\mathcal{D} = \{(\beta, \varepsilon) : \varepsilon < \varepsilon_c^a(\beta)\}$  represent the localized and delocalized region, respectively.

only if  $\varepsilon > \varepsilon_c^a(\beta)$ . Of course  $\varepsilon_c^a(0)$  equals the value  $\varepsilon_c^a$ , defined before Theorem 1.2. Using convexity arguments and Jensen's inequality (*annealed bound*), in close analogy with [13], Proposition 5.1, one proves easily that

$$F^a(\varepsilon, 0) \leq F^a(\varepsilon, \beta) \leq F^a(\varepsilon M(\beta), 0) \quad \forall \beta \geq 0.$$

This yields immediately that for every  $\beta \geq 0$

$$(1.25) \quad \frac{\varepsilon_c^a}{M(\beta)} \leq \varepsilon_c^a(\beta) \leq \varepsilon_c^a;$$

see Figure 2. In particular,  $0 < \varepsilon_c^a(\beta) < \infty$  for every  $\beta \geq 0$ , by Theorem 1.2. By convexity arguments one also shows that the curve  $\beta \mapsto \varepsilon_c^a(\beta)$  is continuous and nonincreasing.

The picture emerging from these simple considerations is that, also in presence of disorder, there is still a nontrivial phase transition: by increasing the parameter  $\varepsilon$  one passes from the delocalized region  $\varepsilon < \varepsilon_c^a(\beta)$  to the localized region  $\varepsilon > \varepsilon_c^a(\beta)$ . When disorder is absent, that is, for  $\beta = 0$ , the behavior of the free energy near the phase transition critical point  $\varepsilon_c^a = \varepsilon_c^a(0)$  is described by Theorems 1.3 and 1.4, for the wetting and the pinning models, respectively, and it is natural to ask what happens when disorder is switched on, that is, for  $\beta > 0$ . It turns out that the presence of disorder has a strong smoothing effect on the phase transition. More precisely, one can prove the following result.

**PROPOSITION 1.6.** *For every  $\beta > 0$  there is a constant  $C_\beta \in (0, \infty)$  such that for both  $a = p, w$  the following inequality holds true for small  $\delta > 0$ :*

$$(1.26) \quad F^a(\varepsilon_c^a(\beta) + \delta, \beta) \leq C_\beta \delta^2.$$

This means that, as soon as disorder is present, the phase transition is at least of second order, with *bounded* second derivative of the free energy. Let us stress the sharp difference with the results we have found for the homogeneous case  $\beta = 0$ :

- in the wetting case, by Theorem 1.3 the transition is of first order and in fact

$$F^W(\varepsilon_c^W + \delta, 0) \geq c_6 \delta,$$

for small  $\delta > 0$  and for some positive constant  $c_6$ , in contrast with (1.26);

- even in the pinning case, where the transition is of second order, (1.18) shows that the second derivative of the free energy from the right at  $\varepsilon = \varepsilon_c^P$  is *infinite*, in contrast with (1.26).

We omit a detailed proof of Proposition 1.6, because it follows exactly the same line as that of [15], Theorem 2.1, and we just sketch it. The idea is to search for *long, atypical stretches* in the disorder sequence  $\omega$  and to give an explicit *lower bound* on the partition function, by making the field come back to the defect line only in correspondence to these stretches. Then, in a clever way, one extracts from this lower bound an *upper bound* on the free energy near the critical point (for a beautiful heuristic description see [14]).

The original proof of [15], Theorem 2.1, holds for a general class of disordered pinning models, for which the free models (i.e., without pinning) enjoy the following properties:

1. the epochs of the returns to zero have a *genuine renewal structure*;
2. the corresponding interarrival time probability has a power-law lower bound.

This class of models includes in particular  $(1 + 1)$ -dimensional pinning and wetting models with *gradient interaction*. As we already mentioned, in our Laplacian-interaction case the returns to zero have only a *Markov* renewal structure. Nevertheless, the set  $\{i \in \mathbb{N} : \varphi_i = \varphi_{i-1} = 0\}$  of *double zeros* has a genuine renewal structure, as we show in Section 5.1, and this is sufficient for the proof, because making the field come back to the defect line with double zeros gives a lower bound. Finally, the power-law lower bound for the interarrival time probability in the free model (i.e., when  $\varepsilon = 0$ ) is given by Proposition 2.3 in the pinning case and by Proposition 1.5 in the wetting case. This shows again the usefulness of bounds like (1.23) and (1.24), despite the fact that they are not sharp.

**1.7. Some recurrent notation.** Throughout the paper, generic positive and finite constants will be denoted by *(const.)*, *(const.)'*. For us  $\mathbb{N} = \{1, 2, \dots\}$ ,  $\mathbb{Z}^+ = \mathbb{N} \cup \{0\}$  and  $\mathbb{R}^+ := [0, \infty)$ . Given two positive sequences  $(a_n)$ ,  $(b_n)$ , by  $a_n \sim b_n$  we mean that  $a_n/b_n \rightarrow 1$  as  $n \rightarrow \infty$ . For  $x \in \mathbb{R}$  we denote as usual by  $\lfloor x \rfloor := \max\{n \in \mathbb{Z} : n \leq x\}$  its integer part.

In this paper we deal with kernels of two kinds. Kernels of the first kind are just  $\sigma$ -finite kernels on  $\mathbb{R}$ , that is, functions  $A_{\cdot, \cdot} : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}^+$ , where  $\mathcal{B}(\mathbb{R})$

denotes the Borel  $\sigma$ -field of  $\mathbb{R}$ , such that  $A_{x,\cdot}$  is a  $\sigma$ -finite Borel measure on  $\mathbb{R}$  for every  $x \in \mathbb{R}$  and  $A_{\cdot,F}$  is a Borel function for every  $F \in \mathcal{B}(\mathbb{R})$ . Given two such kernels  $A_{x,dy}$ ,  $B_{x,dy}$ , their *composition* is denoted as usual by  $(A \circ B)_{x,dy} := \int_{z \in \mathbb{R}} A_{x,dz} B_{z,dy}$  and  $A_{x,dy}^{\circ k}$  denotes the  $k$ -fold composition of  $A$  with itself, where  $A_{x,dy}^{\circ 0} := \delta_x(dy)$ . We also use the standard notation

$$(1 - A)_{x,dy}^{-1} := \sum_{k=0}^{\infty} A_{x,dy}^{\circ k},$$

which of course in general may be infinite.

The second kind of kernel is obtained by letting a kernel of the first kind depend on the further parameter  $n \in \mathbb{Z}^+$ , that is, we consider objects of the form  $\mathbf{A}_{x,dy}(n)$  with  $x, y \in \mathbb{R}$  and  $n \in \mathbb{Z}^+$ . Given two such kernels  $\mathbf{A}_{x,dy}(n)$ ,  $\mathbf{B}_{x,dy}(n)$ , we define their *convolution* by

$$\begin{aligned} (\mathbf{A} * \mathbf{B})_{x,dy}(n) &:= \sum_{m=0}^n (\mathbf{A}(m) \circ \mathbf{B}(n-m))_{x,dy} \\ &= \sum_{m=0}^n \int_{z \in \mathbb{R}} \mathbf{A}_{x,dz}(m) \mathbf{B}_{z,dy}(n-m), \end{aligned}$$

and the  $k$ -fold convolution of the kernel  $\mathbf{A}_{x,dy}(n)$  with itself will be denoted by  $\mathbf{A}_{x,dy}^{*k}(n)$ , where by definition  $\mathbf{A}_{x,dy}^{*0}(n) := \delta_0(dy) \mathbf{1}_{(n=0)}$ . Finally, given two kernels  $\mathbf{A}_{x,dy}(n)$  and  $B_{x,dy}$  and a positive sequence  $(a_n)$ , we will write

$$\mathbf{A}_{x,dy}(n) \sim \frac{B_{x,dy}}{a_n} \quad (n \rightarrow \infty)$$

to mean  $\mathbf{A}_{x,F}(n) \sim B_{x,F}/a_n$  as  $n \rightarrow \infty$ ,  $\forall x \in \mathbb{R}$  and for every *bounded* Borel set  $F \subset \mathbb{R}$ .

**2. The free case  $\epsilon = 0$ : a random walk viewpoint.** In this section we study in detail the free laws  $\mathbb{P}_{0,N}^{\text{P}}$  and  $\mathbb{P}_{0,N}^{\text{W}}$  and their link with the integral of a random walk. The main results are a basic local limit theorem and some asymptotic estimates.

**2.1. Integrated random walk.** Given  $a, b \in \mathbb{R}$ , let  $(\Omega, \mathcal{F}, \mathbf{P} = \mathbf{P}^{(a,b)})$  be a probability space on which are defined the processes  $\{X_i\}_{i \in \mathbb{N}}$ ,  $\{Y_i\}_{i \in \mathbb{Z}^+}$  and  $\{Z_i\}_{i \in \mathbb{Z}^+}$  with the following properties:

- $\{X_i\}_{i \in \mathbb{N}}$  is a sequence of independent and identically distributed random variables, with marginal laws  $X_1 \sim \exp(-V(x)) dx$ . We recall that by our assumptions on  $V(\cdot)$  it follows that  $\mathbf{E}(X_1) = 0$  and  $\mathbf{E}(X_1^2) = \sigma^2 \in (0, \infty)$ ; cf. (1.2).
- $\{Y_i\}_{i \in \mathbb{Z}^+}$  is the random walk associated to  $\{X_i\}$ , with starting point  $a$ , that is,

$$(2.1) \quad Y_0 = a, \quad Y_n = a + X_1 + \cdots + X_n.$$

- $\{Z_i\}_{i \in \mathbb{Z}^+}$  is the *integrated random walk process* with initial value  $b$ ; that is,  $Z_0 = b$  and for  $n \in \mathbb{N}$

$$(2.2) \quad Z_n = b + Y_1 + \cdots + Y_n = b + na + nX_1 + (n-1)X_2 + \cdots + X_n.$$

From (2.1) and (2.2) it follows that

$$(2.3) \quad \{(Y_n, Z_n)\}_n \text{ under } \mathbf{P}^{(a,b)} \stackrel{d}{=} \{(Y_n + a, Z_n + b + na)\}_n \text{ under } \mathbf{P}^{(0,0)}.$$

The marginal distributions of the process  $\{Z_n\}_n$  are specified in the following lemma.

LEMMA 2.1. *For every  $n \in \mathbb{N}$ , the law of the vector  $(Z_1, \dots, Z_n)$  under  $\mathbf{P}^{(a,b)}$  is given by*

$$(2.4) \quad \begin{aligned} & \mathbf{P}^{(a,b)}((Z_1, \dots, Z_n) \in (dz_1, \dots, dz_n)) \\ &= \exp(-\mathcal{H}_{[-1,n]}(z_{-1}, z_0, z_1, \dots, z_n)) \prod_{i=1}^n dz_i, \end{aligned}$$

where we set  $z_{-1} := b - a$  and  $z_0 := b$ .

PROOF. By (2.2) we have  $Y_n = Z_n - Z_{n-1}$  for  $n \geq 1$ . Then, setting  $y_i := z_i - z_{i-1}$  for  $i \geq 2$  and  $y_1 := z_1 - b$ , it suffices to show that, under the measure given by the r.h.s. of (2.4), the variables  $(y_i)_{i=1, \dots, n}$  are distributed like the first  $n$  steps of a random walk starting at  $a$  and with step law  $\exp(-V(x)) dx$ . But for this it suffices to rewrite the Hamiltonian as

$$\begin{aligned} \mathcal{H}_{[-1,n]}(z) &= V((z_1 - b) - (b - (b - a))) + \sum_{i=1}^{n-1} V((z_{i+1} - z_i) - (z_i - z_{i-1})) \\ &= V(y_1 - a) + \sum_{i=1}^{n-1} V(y_{i+1} - y_i), \end{aligned}$$

and the proof is completed.  $\square$

By construction  $\{(Y_n, Z_n)\}_{n \in \mathbb{Z}^+}$  under  $\mathbf{P}^{(a,b)}$  is a Markov process with starting values  $Y_0 = a$ ,  $Z_0 = b$ . On the other hand, the process  $\{Z_n\}_n$  alone is not a Markov process; it is rather a process with *finite memory*  $m = 2$ , that is, for every  $n \in \mathbb{N}$

$$(2.5) \quad \begin{aligned} \mathbf{P}^{(a,b)}(\{Z_{n+k}\}_{k \geq 0} \in \cdot \mid Z_i, i \leq n) &= \mathbf{P}^{(a,b)}(\{Z_{n+k}\}_{k \geq 0} \in \cdot \mid Z_{n-1}, Z_n) \\ &= \mathbf{P}^{(Z_n - Z_{n-1}, Z_n)}(\{Z_k\}_{k \geq 0} \in \cdot), \end{aligned}$$

as follows from Lemma 2.1. For this reason the law  $\mathbf{P}^{(a,b)}$  may be viewed as

$$(2.6) \quad \mathbf{P}^{(a,b)} = \mathbf{P}(\cdot \mid Z_{-1} = b - a, Z_0 = b).$$

**2.2. The link with  $\mathbb{P}_{0,N}^a$ .** In the r.h.s. of (2.4) we see exactly the same density appearing in the definitions (1.5) and (1.7) of our models  $\mathbb{P}_{0,N}^a$ . As an immediate consequence we have the following proposition, which states that  $\mathbb{P}_{0,N}^a$  is nothing but a bridge of the process  $\{Z_n\}_n$  for  $a = p$ , with the further constraint to stay nonnegative for  $a = w$ .

**PROPOSITION 2.2.** *The following statements hold:*

- (1) *The pinning model  $\mathbb{P}_{0,N}^p$  is the law of the vector  $(Z_1, \dots, Z_{N-1})$  under the measure  $\mathbf{P}^{(0,0)}(\cdot | Z_N = 0, Z_{N+1} = 0)$ . The partition function  $Z_{0,N}^p$  is the value at  $(0, 0)$  of the density of the vector  $(Y_{N+1}, Z_{N+1})$  under the law  $\mathbf{P}^{(0,0)}$ .*
- (2) *Setting  $\Omega_n^+ := \{S_1 \geq 0, \dots, S_n \geq 0\}$ , the wetting model  $\mathbb{P}_{0,N}^w$  is the law of the vector  $(Z_1, \dots, Z_{N-1})$  under the measure  $\mathbf{P}^{(0,0)}(\cdot | \Omega_{N-1}^+, Z_N = 0, Z_{N+1} = 0)$ . The partition function  $Z_{0,N}^w$  is the value at  $(0, 0)$  of the density of the vector  $(Y_{N+1}, Z_{N+1})$  under the law  $\mathbf{P}^{(0,0)}(\cdot | \Omega_{N-1}^+)$ , multiplied by  $\mathbf{P}^{(0,0)}(\Omega_{N-1}^+)$ .*

For the statement on the partition function, observe that the density at  $(0, 0)$  of the vector  $(Y_{N+1}, Z_{N+1})$  coincides with the one of the vector  $(Z_N, Z_{N+1})$ , since  $Y_{N+1} = Z_{N+1} - Z_N$ .

**2.3. A local limit theorem.** In view of Proposition 2.2, we study the asymptotic behavior as  $n \rightarrow \infty$  of the vector  $(Y_n, Z_n)$  under the law  $\mathbf{P}^{(a,b)}$ .

Let us denote by  $\{B_t\}_{t \in [0,1]}$  a standard Brownian motion and by  $\{I_t\}_{t \in [0,1]}$  its integral process  $I_t := \int_0^t B_s ds$ . A simple application of Donsker's invariance principle shows that the vector  $(Y_n/(\sigma\sqrt{n}), Z_n/(\sigma n^{3/2}))$  under  $\mathbf{P}^{(0,0)}$  converges in distribution as  $n \rightarrow \infty$  toward the law of the centered Gaussian vector  $(B_1, I_1)$ , whose density  $g(y, z)$  is

$$(2.7) \quad g(y, z) = \frac{6}{\pi} \exp(-2y^2 - 6z^2 + 6yz).$$

We want to reinforce this convergence in the form of a local limit theorem. To this purpose, we introduce the density of  $(Y_n, Z_n)$  under  $\mathbf{P}^{(a,b)}$ , setting for  $n \geq 2$

$$(2.8) \quad \varphi_n^{(a,b)}(y, z) = \frac{\mathbf{P}^{(a,b)}((Y_n, Z_n) \in (dy, dz))}{dy dz}.$$

From 2.3 it follows that

$$(2.9) \quad \varphi_n^{(a,b)}(y, z) = \varphi_n^{(0,0)}(y - a, z - b - na),$$

hence it suffices to focus on  $\varphi_n^{(0,0)}(\cdot, \cdot)$ . We set for short  $\varphi_n^{(0,0)}(\mathbb{R}, z) := \int_{\mathbb{R}} \varphi_n^{(0,0)}(y, z) dy$ , that is, the density of  $Z_n$  under  $\mathbf{P}^{(0,0)}$ , and  $g(\mathbb{R}, z) := \int_{\mathbb{R}} g(y, z) dy$ . We are ready to state the main result of this section.

**PROPOSITION 2.3** (Local limit theorem). *The following relations hold as  $n \rightarrow \infty$ :*

$$(2.10) \quad \begin{aligned} & \sup_{(y,z) \in \mathbb{R}^2} |\sigma^2 n^2 \varphi_n^{(0,0)}(\sigma\sqrt{n}y, \sigma n^{3/2}z) - g(y, z)| \rightarrow 0, \\ & \sup_{z \in \mathbb{R}} |\sigma n^{3/2} \varphi_n^{(0,0)}(\mathbb{R}, \sigma n^{3/2}z) - g(\mathbb{R}, z)| \rightarrow 0. \end{aligned}$$

The proof, based on Fourier analysis, is deferred to Appendix B. We stress that this result retains a crucial importance in the rest of the paper. Notice that an analogous local limit theorem holds also for  $\varphi_n^{(0,0)}(y, \mathbb{R})$ , that is, the density of  $Y_n$ , but we do not state it explicitly because we will not need it.

**2.4. The positivity constraint.** To deal with the wetting model we need to study the law of the random vector  $(Y_n, Z_n)$ , or equivalently of  $(Z_{n-1}, Z_n)$ , conditionally on the event  $\Omega_{n-2}^+ = \{Z_1 \geq 0, \dots, Z_{n-2} \geq 0\}$ . To this purpose we set for  $x, y \in \mathbb{R}$  and  $n \geq 3$

$$(2.11) \quad w_{x,y}(n) := \mathbf{1}_{(x \geq 0, y \geq 0)} \cdot \mathbf{P}^{(-x,0)}(\Omega_{n-2}^+ | Z_{n-1} = y, Z_n = 0),$$

while for  $n = 1, 2$  we simply set  $w_{x,y}(n) := \mathbf{1}_{(x \geq 0, y \geq 0)}$ . We are interested in the rate of decay of  $w_{x,y}(n)$  as  $n \rightarrow \infty$ . To this purpose we claim that there exists a positive constant  $c_+ > 1$  such that the following upper bound holds: for all  $n \in \mathbb{N}$

$$(2.12) \quad \sup_{0 \leq x, y \leq \sqrt{n}} w_{x,y}(n) \leq \frac{(const.)}{(\log n)^{c_+}}.$$

Moreover we have the following lower bound for  $x, y = 0$  and for all  $n \in \mathbb{N}$ :

$$(2.13) \quad w_{0,0}(n) \geq \frac{(const.)}{n^{c_-}},$$

for some positive constant  $c_-$ . Notice that by Proposition 2.2 we have  $Z_{0,N}^p = \varphi_{N+1}^{(0,0)}(0, 0)$  and  $Z_{0,N}^w = Z_{0,N}^p \cdot w_{0,0}(N+1)$ , hence by (2.10) and (2.13) we have for every  $N \in \mathbb{N}$

$$(2.14) \quad Z_{0,N}^p \geq Z_{0,N}^w \geq \frac{(const.)}{N^{2+c_-}},$$

so that the last inequality in (1.9) is proven.

We prove the lower bound (2.13) in Appendix C.1; the idea is to restrict the expectation that defines  $w_{0,0}(n)$  on a suitable subset of paths, whose probability can be estimated. On the other hand, the upper bound (2.12) follows directly combining the following lemma with the upper bound in (1.23) (which is proven in Appendix C.2).

LEMMA 2.4. *There exists a positive constant  $C$  such that for every  $n \in \mathbb{N}$*

$$(2.15) \quad \sup_{0 \leq x, y \leq \sqrt{n}} w_{x,y}(n) \leq C \cdot \mathbf{P}^{(0,0)}(\Omega_{\lfloor n/2 \rfloor}^+).$$

PROOF. The inclusion bound gives (for  $x, y \geq 0$ )

$$\begin{aligned} w_{x,y}(n) &= \mathbf{P}^{(-x,0)}(\Omega_{n-2}^+ | Z_{n-1} = y, Z_n = 0) \\ &\leq \mathbf{P}^{(-x,0)}(\Omega_{\lfloor n/2 \rfloor}^+ | Z_{n-1} = y, Z_n = 0), \end{aligned}$$

and disintegrating over the variables  $Z_{\lfloor n/2 \rfloor - 1}, Z_{\lfloor n/2 \rfloor}$  we get [recall (2.6) and (2.9)]

$$\begin{aligned} w_{x,y}(n) &= \int_{a,b \in \mathbb{R}^+} \frac{\mathbf{P}^{(-x,0)}(\Omega_{\lfloor n/2 \rfloor}^+, Z_{\lfloor n/2 \rfloor - 1} \in da, Z_{\lfloor n/2 \rfloor} \in db) \cdot \varphi_{n-\lfloor n/2 \rfloor}^{(b-a,b)}(-y, 0)}{\varphi_n^{(-x,0)}(-y, 0)}. \end{aligned}$$

Proposition 2.3 together with (2.9) yields

$$\sup_{a,b,y \in \mathbb{R}} \varphi_{n-\lfloor n/2 \rfloor}^{(b-a,b)}(-y, 0) \leq \frac{(const.)}{n^2}, \quad \inf_{0 \leq x, y \leq \sqrt{n}} \varphi_n^{(-x,0)}(-y, 0) \geq \frac{(const.')}{n^2},$$

hence setting  $C := (const.)/(const.') > 0$  we have

$$\begin{aligned} w_{x,y}(n) &\leq C \int_{a,b \in \mathbb{R}^+} \mathbf{P}^{(-x,0)}(\Omega_{\lfloor n/2 \rfloor}^+, Z_{\lfloor n/2 \rfloor - 1} \in da, Z_{\lfloor n/2 \rfloor} \in db) \\ &= C \mathbf{P}^{(-x,0)}(\Omega_{\lfloor n/2 \rfloor}^+), \end{aligned}$$

and the proof is completed noting that  $\mathbf{P}^{(-x,0)}(\Omega_{\lfloor n/2 \rfloor}^+) \leq \mathbf{P}^{(0,0)}(\Omega_{\lfloor n/2 \rfloor}^+)$  by (2.3).  $\square$

**3. The interacting case  $\varepsilon > 0$ : a renewal theory description.** In this section we study in detail the laws  $\mathbb{P}_{\varepsilon,N}^p$  and  $\mathbb{P}_{\varepsilon,N}^w$  in the case  $\varepsilon > 0$ . The crucial result is that the contact set  $\{i \in \mathbb{Z}^+ : \varphi_i = 0\}$  can be described in terms of a Markov renewal process. Throughout the section we assume that  $\varepsilon > 0$ .

3.1. *The law of the contact set.* We introduce the *contact set*  $\tau$  by:

$$(3.1) \quad \tau := \{i \in \mathbb{Z}^+ : \varphi_i = 0\} \subset \mathbb{Z}^+,$$

where we set by definition  $\varphi_0 = 0$ , so that  $0 \in \tau$ . It is practical to identify the set  $\tau$  with the increasing sequence of random variables  $\{\tau_k\}_{k \geq 0}$  defined by

$$(3.2) \quad \tau_0 := 0, \quad \tau_{k+1} := \inf\{i > \tau_k : \varphi_i = 0\}.$$

Observe that the random variable  $\ell_N$ , introduced in (1.12), may be expressed as  $\ell_N := \max\{k : \tau_k \leq N\}$ . Next we introduce the process  $\{J_k\}_{k \geq 0}$  that gives the height of the field before the contact points:

$$(3.3) \quad J_0 := 0, \quad J_k := \varphi_{\tau_k - 1}.$$

The basic observation is that the joint law of the process  $\{\ell_N, (\tau_k)_{k \leq \ell_N}, (J_k)_{k \leq \ell_N}\}$  under  $\mathbb{P}_{\varepsilon, N}^a$  can be written in the following “product form”: for  $k \in \mathbb{N}$ , for  $(t_i)_{i=1, \dots, k} \in \mathbb{N}$  with  $0 < t_1 < \dots < t_{k-1} < t_k := N$  and for  $(y_i)_{i=1, \dots, k} \in \mathbb{R}$  we have

$$(3.4) \quad \begin{aligned} & \mathbb{P}_{\varepsilon, N}^a(\ell_N = k, \tau_i = t_i, J_i \in dy_i, i = 1, \dots, k) \\ &= \frac{1}{Z_{\varepsilon, N}^a} \varepsilon^{k-1} \mathsf{F}_{0, dy_1}^a(t_1) \\ & \quad \times \mathsf{F}_{y_1, dy_2}^a(t_2 - t_1) \cdots \cdots \mathsf{F}_{y_{k-1}, dy_k}^a(N - t_{k-1}) \cdot \mathsf{F}_{y_k, \{0\}}^a(1), \end{aligned}$$

for a suitable kernel  $\mathsf{F}_{x, dy}^a(n)$  that we now define. For  $a = p$  we have

$$(3.5) \quad \begin{aligned} & \mathsf{F}_{x, dy}^p(n) \\ &:= \begin{cases} e^{-\mathcal{H}_{[-1,1]}(x, 0, 0)} \delta_0(dy) = e^{-V(x)} \delta_0(dy), & \text{if } n = 1, \\ e^{-\mathcal{H}_{[-1,2]}(x, 0, y, 0)} dy = e^{-V(x+y) - V(-2y)} dy, & \text{if } n = 2, \\ \left( \int_{\mathbb{R}^{n-2}} e^{-\mathcal{H}_{[-1,n]}(\varphi_{-1}, \dots, \varphi_n)} d\varphi_1 \cdots d\varphi_{n-2} \right) dy & \text{if } n \geq 3, \\ \text{where } \varphi_{-1} = x, \varphi_0 = 0, \varphi_{n-1} = y, \varphi_n = 0 \end{cases} \end{aligned}$$

and the definition of  $\mathsf{F}_{x, dy}^w(n)$  is analogous: we just have to impose that  $x, y \geq 0$  and for  $n \geq 3$  we also have to restrict the integral in (3.5) on  $(\mathbb{R}^+)^{n-2}$ . Although these formulas may appear quite involved, they follow easily from the definition of  $\mathbb{P}_{\varepsilon, N}^a$ . In fact it suffice, to expand the product of measures in the r.h.s. of (1.5) and (1.7) as a sum of “monomials,” according to the elementary formula (where we set  $\Omega^p := \mathbb{R}$  and  $\Omega^w := \mathbb{R}^+$ )

$$\prod_{i=1}^{N-1} (\varepsilon \delta_0(d\varphi_i) + d\varphi_i \mathbf{1}_{(\varphi_i \in \Omega^a)}) = \sum_{A \subset \{1, \dots, N-1\}} \varepsilon^{|A|} \prod_{m \in A} \delta_0(d\varphi_m) \prod_{n \in A^c} d\varphi_n \mathbf{1}_{(\varphi_n \in \Omega^a)}.$$

It is then clear that  $A = \{\tau_1, \dots, \tau_{\ell_N-1}\}$  and integrating over the variables  $\varphi_i$  with index  $i \notin A \cup (A - 1)$  one gets to (3.4). We stress that the algebraic structure of (3.4) retains a crucial importance, that we are going to exploit in the next paragraph.

From (3.5) it follows that the kernel  $\mathsf{F}_{x, dy}^a(n)$  is a Dirac mass in  $y$  for  $n = 1$  while it is absolutely continuous for  $n \geq 2$ . Then it is convenient to introduce the  $\sigma$ -finite measure  $\mu(dx) := \delta_0(dx) + dx$ , so that we can write

$$(3.6) \quad \mathsf{F}_{x, dy}^a(n) = \mathsf{f}_{x, y}^a(n) \mu(dy).$$

The interesting fact is that the density  $\mathsf{f}_{x, y}^a(n)$  can be rephrased explicitly in terms of the process  $\{Z_k\}_k$  introduced in Section 2.1. Let us start with the pinning case  $a = p$ : From Lemma 2.1 and from (3.5) it follows that, for  $n \geq 2$ ,

$f_{x,y}^p(n)$  is nothing but the density of  $(Z_{n-1}, Z_n)$  at  $(y, 0)$ , under  $\mathbf{P}^{(-x,0)}$ . Recalling the definition (2.8) of  $\varphi_n^{(\cdot,\cdot)}(\cdot, \cdot)$  and the fact that  $Z_{n-1} = Z_n - Y_n$ , we can write

$$(3.7) \quad f_{x,y}^p(n) = \begin{cases} e^{-V(x)} \mathbf{1}_{(y=0)}, & \text{if } n = 1, \\ \varphi_n^{(-x,0)}(-y, 0) \mathbf{1}_{(y \neq 0)} \\ \quad = \varphi_n^{(0,0)}(-y + x, nx) \mathbf{1}_{(y \neq 0)}, & \text{if } n \geq 2, \end{cases}$$

where we have used relation (2.9). Analogously, recalling the definition (2.11) of  $w_{x,y}(n)$ , in the wetting case we have

$$(3.8) \quad f_{x,y}^w(n) = f_{x,y}^p(n) \cdot w_{x,y}(n).$$

Equations (3.6), (3.7) and (3.8) provide a description of the kernel  $F_{x,dy}^a(n)$  which is both simpler and more useful than the original definition (3.5).

**3.2. A Markov renewal theory interpretation.** Equation (3.4) expresses the law of  $\{(\tau_k, J_k)\}_k$  under  $\mathbb{P}_{\varepsilon,N}^a$  in terms of an explicit kernel  $F_{x,dy}^a(n)$ . The crucial point is that the algebraic structure of equation (3.4) allows to modify the kernel, in order to give this formula a direct renewal theory interpretation. In fact we set

$$(3.9) \quad K_{x,dy}^{a,\varepsilon}(n) := \varepsilon F_{x,dy}^a(n) e^{-F^a(\varepsilon)n} \frac{v_\varepsilon^a(y)}{v_\varepsilon^a(x)},$$

where the number  $F^a(\varepsilon) \in [0, \infty)$  and the positive real function  $v_\varepsilon^a(\cdot)$  will be defined explicitly in Section 4. Of course this is an abuse of notation, because the symbol  $F^a(\varepsilon)$  was already introduced to denote the free energy [cf. (1.8)], but we will show in Section 5.2 that the two quantities indeed coincide. We denote by  $k_{x,y}^{a,\varepsilon}(n)$  the density of  $K_{x,dy}^{a,\varepsilon}(n)$  with respect to  $\mu(dy)$ , that is,

$$(3.10) \quad k_{x,y}^{a,\varepsilon}(n) := \varepsilon f_{x,y}^a(n) e^{-F^a(\varepsilon)n} \frac{v_\varepsilon^a(y)}{v_\varepsilon^a(x)}.$$

The reason for introducing the kernel  $K_{x,dy}^{a,\varepsilon}(n)$  lies in the following fundamental fact: the number  $F^a(\varepsilon)$  and the function  $v_\varepsilon^a(\cdot)$  appearing in (3.9) can be chosen such that,

$$(3.11) \quad \forall x \in \mathbb{R}: \quad \int_{y \in \mathbb{R}} \sum_{n \in \mathbb{N}} K_{x,dy}^{a,\varepsilon}(n) = \frac{\varepsilon}{\varepsilon_c^a} \wedge 1 \leq 1,$$

where  $\varepsilon_c^a \in (0, \infty)$  is a fixed number. A detailed proof and discussion of this fact, with an explicit definition of  $\varepsilon_c^a$ ,  $F^a(\varepsilon)$  and  $v_\varepsilon^a(\cdot)$ , is deferred to Section 4; for the moment we focus on its consequences.

Thanks to (3.11), we can define the law  $\mathcal{P}_\varepsilon^a$  under which the joint process  $\{(\tau_k, J_k)\}_{k \geq 0}$  is a (possibly defective) Markov chain on  $\mathbb{Z}^+ \times \mathbb{R}$ , with starting value  $(\tau_0, J_0) = (0, 0)$  and with transition kernel given by

$$(3.12) \quad \mathcal{P}_\varepsilon^a((\tau_{k+1}, J_{k+1}) \in (\{n\}, dy) | (\tau_k, J_k) = (m, x)) = K_{x,dy}^{a,\varepsilon}(n - m).$$

An alternative (and perhaps more intuitive) definition is as follows:

- First sample the process  $\{J_k\}_{k \geq 0}$  as a (defective if  $\varepsilon < \varepsilon_c^a$ ) Markov chain on  $\mathbb{R}$ , with  $J_0 = 0$  and with transition kernel

$$(3.13) \quad \mathcal{P}_\varepsilon^a(J_{k+1} \in dy \mid J_k = x) = \sum_{n \in \mathbb{N}} K_{x,dy}^{a,\varepsilon}(n) =: D_{x,dy}^{a,\varepsilon}.$$

In the defective case we take  $\infty$  as cemetery.

- Then sample the increments  $\{T_k := \tau_k - \tau_{k-1}\}_{k \in \mathbb{N}}$  as a sequence of independent, but not identically distributed, random variables, according to the conditional law:

$$\mathcal{P}_\varepsilon^a(T_k = n \mid \{J_i\}_{i \geq 0}) = \begin{cases} \mathbf{1}_{(n=1)}, & \text{if } J_k = 0, \\ \frac{K_{J_{k-1},J_k}^{a,\varepsilon}(n)\mathbf{1}_{(n \geq 2)}}{\sum_{m \geq 2} K_{J_{k-1},J_k}^{a,\varepsilon}(m)}, & \text{if } J_k \neq 0, J_k \neq \infty, \\ \mathbf{1}_{(n=\infty)} & \text{if } J_k = \infty. \end{cases}$$

We stress that the process  $\{(\tau_k, J_k)\}_{k \geq 0}$  is defective if  $\varepsilon < \varepsilon_c^a$  and proper if  $\varepsilon \geq \varepsilon_c^a$ ; cf. (3.11). The process  $\{\tau_k\}_{k \geq 0}$  under  $\mathcal{P}_\varepsilon^a$  is what is called a *Markov renewal process* and  $\{J_k\}_{k \geq 0}$  is its *modulating chain*. This is a generalization of classical renewal processes, because  $\tau_n = T_1 + \dots + T_n$  where the variables  $\{T_k\}_{k \in \mathbb{N}}$  are allowed to have a special kind of dependence, namely they are independent conditionally on the modulating chain  $\{J_k\}_{k \geq 0}$ . For a detailed account on Markov renewal processes we refer to [4].

Now let us come back to (3.4). We perform the substitution  $F_{x,dy}^a(n) \rightarrow K_{x,dy}^{a,\varepsilon}(n)$ , defined in (3.9): the boundary terms  $v_\varepsilon^a(y)/v_\varepsilon^a(x)$  get simplified and the exponential term  $e^{-F^a(\varepsilon)n}$  factorizes, so that we get

$$(3.14) \quad \begin{aligned} \mathbb{P}_{\varepsilon,N}^a(\ell_N = k, \tau_i = t_i, J_i \in dy_i, i = 1, \dots, k) \\ = \frac{e^{F^a(\varepsilon)N}}{\varepsilon^2 Z_{\varepsilon,N}^a} K_{0,dy_1}^{a,\varepsilon}(t_1) \cdot K_{y_1,dy_2}^{a,\varepsilon}(t_2 - t_1) \cdots \\ \cdot K_{y_{k-1},dy_k}^{a,\varepsilon}(N - t_{k-1}) \cdot K_{y_k,\{0\}}^{a,\varepsilon}(1). \end{aligned}$$

Moreover, since the partition function  $Z_{\varepsilon,N}^a$  is the normalizing constant that makes  $\mathbb{P}_{\varepsilon,N}^a$  a probability, it can be expressed as

$$(3.15) \quad \begin{aligned} Z_{\varepsilon,N}^a = \frac{e^{F^a(\varepsilon)N}}{\varepsilon^2} \sum_{k=1}^N \sum_{\substack{t_i \in \mathbb{N}, i=1, \dots, k \\ 0 < t_1 < \dots < t_k := N}} \int_{\mathbb{R}^k} \left( \prod_{i=1}^k K_{y_{i-1},dy_i}^{a,\varepsilon}(t_i - t_{i-1}) \right) \\ \cdot K_{y_k,\{0\}}^{a,\varepsilon}(1). \end{aligned}$$

We are finally ready to make explicit the link between the law  $\mathcal{P}_\varepsilon^a$  and our model  $\mathbb{P}_{\varepsilon,N}^a$ . Let us introduce the event

$$(3.16) \quad \mathcal{A}_N := \{ \{N, N+1\} \subset \tau \} = \{ \exists k \geq 0 : \tau_k = N, \tau_{k+1} = N+1 \}.$$

The following proposition is an immediate consequence of (3.12), (3.14) and (3.15).

**PROPOSITION 3.1.** *For any  $N \in \mathbb{N}$  and  $\varepsilon > 0$ , the vector  $\{\ell_N, (\tau_i)_{i \leq \ell_N}, (J_i)_{i \leq \ell_N}\}$  has the same law under  $\mathbb{P}_{\varepsilon, N}^a$  and under the conditional law  $\mathcal{P}_\varepsilon^a(\cdot | \mathcal{A}_N)$ : for all  $k$ ,  $\{t_i\}_i$  and  $\{y_i\}_i$*

$$\begin{aligned} & \mathbb{P}_{\varepsilon, N}^a(\ell_N = k, \tau_i = t_i, J_i \in dy_i, i \leq k) \\ &= \mathcal{P}_\varepsilon^a(\ell_N = k, \tau_i = t_i, J_i \in dy_i, i \leq k | \mathcal{A}_N). \end{aligned}$$

Moreover the partition function can be expressed as  $Z_{\varepsilon, N}^a = (e^{F^a(\varepsilon)N}/\varepsilon^2) \cdot \mathcal{P}_\varepsilon^a(\mathcal{A}_N)$ .

Thus we have shown that the contact set  $\tau \cap [0, N]$  under the pinning law  $\mathbb{P}_{\varepsilon, N}^a$  is distributed like a Markov renewal process (of law  $\mathcal{P}_\varepsilon^a$ ) conditioned to visit  $\{N, N+1\}$ . The crucial point is that  $\mathcal{P}_\varepsilon^a$  does not have any dependence on  $N$ , therefore all the dependence on  $N$  of  $\mathbb{P}_{\varepsilon, N}^a$  is contained in the conditioning on the event  $\mathcal{A}_N$ . As will be clear in the next sections, this fact is the key to all our results.

**4. An infinite-dimensional Perron–Frobenius problem.** In this section we prove that for every  $\varepsilon > 0$  the nonnegative number  $F^a(\varepsilon)$  and the positive real function  $v_\varepsilon^a(\cdot) : \mathbb{R} \rightarrow (0, \infty)$  appearing in the definition of  $K_{x, dy}^{a, \varepsilon}(n)$ , [cf. (3.9)], can be chosen in such a way that (3.11) holds true.

**4.1. Some analytical preliminaries.** We recall that the kernel  $F_{x, dy}^a(n)$  and its density  $f_{x, y}^a(n)$  are defined in (3.6), (3.7) and (3.8), and that  $\mu(dx) = \delta_0(dx) + dx$ . In particular we have  $0 \leq f_{x, y}^w(n) \leq f_{x, y}^p(n)$ . We first list some important properties of  $f_{x, y}^p(n)$ :

- Uniformly for  $x, y$  in compact sets we have

$$(4.1) \quad f_{x, y}^p(n) \sim \frac{c}{n^2} \quad (n \rightarrow \infty),$$

where  $c := 6/(\pi\sigma^2)$ ; moreover there exists  $C > 0$  such that  $\forall x, y \in \mathbb{R}$  and  $n \in \mathbb{N}$

$$(4.2) \quad f_{x, y}^p(n) \leq \frac{C}{n^2}, \quad \int_{\mathbb{R}} dz f_{x, z}^p(n) \leq \frac{C}{n^{3/2}}, \quad \int_{\mathbb{R}} dz f_{z, y}^p(n) \leq \frac{C}{n^{3/2}}.$$

Both the above relations follow comparing (3.7) with Proposition (2.3).

- For  $n \geq 2$  we have

$$(4.3) \quad \int_{\mathbb{R}^2} dx dy f_{x, y}^p(n) = \frac{1}{n},$$

as follows from (3.7) recalling that  $\varphi_n^{(0, 0)}(\cdot, \cdot)$  is a probability density.

For  $\lambda \geq 0$  we introduce the kernel

$$(4.4) \quad B_{x,dy}^{a,\lambda} := \sum_{n \in \mathbb{N}} e^{-\lambda n} F_{x,dy}^a(n),$$

which induces the integral operator:  $(B^{a,\lambda} h)(x) := \int_{\mathbb{R}} B_{x,dy}^{a,\lambda} h(y)$ . Note that for every  $x \in \mathbb{R}$  the kernel  $B_{x,dy}^{a,\lambda}$  is absolutely continuous with respect to the measure  $\mu$ , so that we can write  $B_{x,dy}^{a,\lambda} = b^{a,\lambda}(x, y) \mu(dy)$ , where the density  $b^{a,\lambda}(x, y)$  is given by [cf. (3.7)]

$$(4.5) \quad b^{a,\lambda}(x, y) = e^{-\lambda} f_{x,0}^a(1) \mathbf{1}_{(y=0)} + \sum_{n \geq 2} e^{-\lambda n} f_{x,y}^a(n) \mathbf{1}_{(y \neq 0)}.$$

The following result is of basic importance.

**LEMMA 4.1.** *For every  $\lambda \geq 0$ ,  $B^{a,\lambda}$  is a compact operator on the Hilbert space  $L^2(\mathbb{R}, d\mu)$ .*

**PROOF.** We are going to check the stronger condition that  $B^{a,\lambda}$  is Hilbert–Schmidt, that is,

$$(4.6) \quad \int_{\mathbb{R}^2} b^{a,\lambda}(x, y)^2 \mu(dx) \mu(dy) < \infty.$$

Since  $0 \leq f_{x,y}^w(n) \leq f_{x,y}^p(n)$  it suffices to focus on the case  $a = p$ . Setting  $\lambda = 0$  in the r.h.s. of (4.5) we obtain

$$b^{p,\lambda}(x, y)^2 \leq f_{x,0}^p(1)^2 \mathbf{1}_{(y=0)} + \sum_{n,m \geq 2} f_{x,y}^p(n) f_{x,y}^p(m) \mathbf{1}_{(y \neq 0)},$$

hence

$$(4.7) \quad \begin{aligned} & \int_{\mathbb{R}^2} b^{p,\lambda}(x, y)^2 \mu(dx) \mu(dy) \\ & \leq \int_{\mathbb{R}} f_{x,0}^p(1)^2 \mu(dx) + \sum_{n,m \geq 2} \int_{\mathbb{R}} f_{0,y}^p(n) f_{0,y}^p(m) dy \\ & \quad + \sum_{n,m \geq 2} \int_{\mathbb{R}^2} f_{x,y}^p(n) f_{x,y}^p(m) dx dy. \end{aligned}$$

From the definition (3.7) of  $f_{x,0}^p(1)$  it is immediate to check that the first term in the r.h.s. is finite. For the second term, we apply the first and the second relation in (4.2), getting

$$\sum_{n,m \geq 2} \int_{\mathbb{R}} f_{0,y}^p(n) f_{0,y}^p(m) dy \leq \sum_{n,m \geq 2} \frac{C}{n^2} \int_{\mathbb{R}} f_{0,y}^p(m) dy \leq \sum_{n,m \geq 2} \frac{C}{n^2} \frac{C}{m^{3/2}} < \infty.$$

For the third term, it is convenient first to exploit the symmetry between  $n$  and  $m$ , restricting the sum on the set  $n \geq m$ :

$$\sum_{n,m \geq 2} \int_{\mathbb{R}^2} f_{x,y}^p(n) f_{x,y}^p(m) dx dy \leq 2 \sum_{m \geq 2} \sum_{n \geq m} \int_{\mathbb{R}^2} f_{x,y}^p(n) f_{x,y}^p(m) dx dy.$$

Then applying relations (4.2) and (4.3) we get

$$\sum_{m \geq 2} \sum_{n \geq m} \int_{\mathbb{R}^2} f_{x,y}^p(n) f_{x,y}^p(m) dx dy \leq \sum_{m \geq 2} \sum_{n \geq m} \frac{1}{m} \frac{C}{n^2} = C \sum_{m \geq 2} \frac{1}{m} \sum_{n \geq m} \frac{1}{n^2}.$$

However, the last sum is bounded by  $(const.)/m$ , hence the r.h.s. is finite.  $\square$

**4.2. A formula for the free energy.** Lemma 4.1 allows us to apply an infinite-dimensional analogue of the classical Perron–Frobenius theorem. We first introduce the function  $\delta^a(\lambda) \in [0, \infty)$  defined for  $\lambda \geq 0$  by

$$(4.8) \quad \delta^a(\lambda) := \text{spectral radius of the operator } B^{a,\lambda}.$$

Notice that  $\{0\}$  is a *proper atom*, hence a *small set*, of the kernel  $B^{a,\lambda}$ ; cf. [26], Section 4.2. Therefore by [26], Section 3.2 one can define  $\delta^a(\lambda)$  equivalently as

$$(4.9) \quad \delta^a(\lambda) := \inf \left\{ R > 0 : \sum_{n=0}^{\infty} \frac{(B^{a,\lambda})_{0,\{0\}}^{\circ n}}{R^n} < \infty \right\},$$

where the convolution  $\circ$  between kernels is defined in Section 1.7. One checks easily that indeed  $\delta^a(\lambda) \in (0, \infty)$  for every  $\lambda \geq 0$ . By Theorem 1 in [33],  $\delta^a(\lambda)$  is an isolated and *simple* eigenvalue of  $B^{a,\lambda}$ . The function  $\delta^a(\cdot)$  is nonincreasing, continuous on  $[0, \infty)$  and analytic on  $(0, \infty)$ , because the operator  $B^{a,\lambda}$  has these properties and  $\delta^a(\lambda)$  is simple and isolated; cf. [20], Chapter VII, Section 1.3. The analyticity and the fact that  $\delta^a(\cdot)$  is not constant (because  $\delta^a(\lambda) \rightarrow 0$  as  $\lambda \rightarrow +\infty$ ) force  $\delta^a(\cdot)$  to be strictly decreasing. Denoting by  $(\delta^a)^{-1}(\cdot)$  the inverse function, defined on the domain  $(0, \delta^a(0)]$ , we can now define  $\varepsilon_c^a$  and  $F^a(\varepsilon)$  by

$$(4.10) \quad \varepsilon_c^a := \frac{1}{\delta^a(0)}, \quad F^a(\varepsilon) := \begin{cases} (\delta^a)^{-1}(1/\varepsilon), & \text{if } \varepsilon \geq \varepsilon_c^a, \\ 0, & \text{if } \varepsilon \leq \varepsilon_c^a. \end{cases}$$

From now on we focus on the operator  $B^{a,F^a(\varepsilon)}$ , that is, on  $B^{a,\lambda}$  for  $\lambda = F^a(\varepsilon)$ , whose spectral radius equals  $1/\varepsilon \wedge 1/\varepsilon_c^a$  by construction. Notice that  $b^{a,F^a(\varepsilon)}(x, y) > 0$  for every  $x, y \in \Omega^a$ , where  $\Omega^p = \mathbb{R}$  and  $\Omega^w = \mathbb{R}^+$ . Then Theorem 1 in [33] gives the existence of the so-called right and left *Perron–Frobenius eigenfunctions* of  $B^{a,F^a(\varepsilon)}$ , that is, of two functions  $v_\varepsilon^a(\cdot), w_\varepsilon^a(\cdot) \in L^2(\mathbb{R}, d\mu)$  such that  $v_\varepsilon^a(x) > 0$  and  $w_\varepsilon^a(x) > 0$ , for  $\mu$ -a.e.  $x \in \Omega^a$ , and such that

$$(4.11) \quad \begin{aligned} \int_{y \in \mathbb{R}} B_{x,dy}^{a,F^a(\varepsilon)} v_\varepsilon^a(y) &= \left( \frac{1}{\varepsilon} \wedge \frac{1}{\varepsilon_c^a} \right) v_\varepsilon^a(x), \\ \int_{x \in \mathbb{R}} w_\varepsilon^a(x) B_{x,dy}^{a,F^a(\varepsilon)} \mu(dx) &= \left( \frac{1}{\varepsilon} \wedge \frac{1}{\varepsilon_c^a} \right) w_\varepsilon^a(y) \mu(dy). \end{aligned}$$

Notice that in the wetting case  $v_\varepsilon^w(x) = w_\varepsilon^w(x) = 0$  for all  $x < 0$ . Spelling out the first equation in (4.11) we have

$$(4.12) \quad v_\varepsilon^a(x) = \frac{1}{1/\varepsilon \wedge 1/\varepsilon_c^a} \sum_{n \in \mathbb{N}} e^{-F^a(\varepsilon)n} \int_{y \in \mathbb{R}} f_{x,y}^a(n) v_\varepsilon^a(y) \mu(dy).$$

This yields easily that  $v_\varepsilon^a(x) > 0$  for every  $x \in \Omega^a$  (and not only  $\mu$ -a.e.). One shows analogously that  $w_\varepsilon^a(x) > 0$  for every  $x \in \Omega^a$ .

Having defined the quantities  $\varepsilon_c^a$ ,  $F^a(\varepsilon)$  and  $v_\varepsilon^a(\cdot)$ , it remains to check that (3.11) indeed holds true. But this is a straightforward consequence of (4.11); in fact by the definition (3.9) of  $K_{x,dy}^{a,\varepsilon}(n)$  we have

$$\begin{aligned} \int_{y \in \mathbb{R}} \sum_{n \in \mathbb{N}} K_{x,dy}^{a,\varepsilon}(n) &= \frac{\varepsilon}{v_\varepsilon^a(x)} \int_{y \in \mathbb{R}} \left( \sum_{n \in \mathbb{N}} F_{x,dy}^a(n) e^{-F^a(\varepsilon)n} \right) v_\varepsilon^a(y) \\ &= \frac{\varepsilon}{v_\varepsilon^a(x)} \int_{y \in \mathbb{R}} B_{x,dy}^{a,F^a(\varepsilon)} v_\varepsilon^a(y) = \varepsilon \left( \frac{1}{\varepsilon} \wedge \frac{1}{\varepsilon_c^a} \right) = 1 \wedge \frac{\varepsilon}{\varepsilon_c^a}. \end{aligned}$$

We note that the functions  $v_\varepsilon^a(\cdot)$  and  $w_\varepsilon^a(\cdot)$  are uniquely defined up to a multiplicative constant and we use this degree of freedom to fix  $\langle v_\varepsilon^a, w_\varepsilon^a \rangle = \int_{\mathbb{R}} v_\varepsilon^a w_\varepsilon^a d\mu = 1$ . In other words, the measure  $v_\varepsilon^a$  defined by

$$(4.13) \quad v_\varepsilon^a(dx) := w_\varepsilon^a(x) v_\varepsilon^a(x) \mu(dx)$$

is a probability measure:  $v_\varepsilon^a(\mathbb{R}) = 1$ . An important observation is that for  $\varepsilon \geq \varepsilon_c^a$  the probability measure  $v_\varepsilon^a$  is invariant for the kernel  $D_{x,dy}^{a,\varepsilon}$ , as follows from (3.13) and (3.9). This means that the Markov chain  $\{J_k\}_k$  is *positive recurrent*; cf. [26].

We conclude this section with a simple perturbation result that will be useful later. Let  $A_{x,dy}$  and  $C_{x,dy}$  be two nonnegative kernels that induce two compact operators on  $L^2(\mathbb{R}, d\mu)$ . Assume that the spectral radius of  $C_{x,dy}$  is strictly positive. For  $t \in [0, \infty)$  we set  $\gamma(t) :=$  spectral radius of  $(A_{x,dy} + t \cdot C_{x,dy})$ . Then we have the following:

**LEMMA 4.2.** *The function  $t \mapsto \gamma(t)$  is strictly increasing; in particular  $\gamma(0) < \gamma(1)$ .*

**PROOF.** The function  $\gamma(\cdot)$  is clearly nondecreasing. It follows by Theorem 1 in [33] that  $\gamma(t)$  is a simple and isolated eigenvalue of  $A_{x,dy} + t \cdot C_{x,dy}$  for every  $t \geq 0$ ; therefore by perturbation theory [20], Chapter VII, Section 1.3 the function  $\gamma(\cdot)$  is analytic. Since  $\gamma(t) \geq (\text{const.}) \cdot t$  (here we use the hypothesis that the spectral radius of  $C_{x,dy}$  is strictly positive), the function  $\gamma(\cdot)$  is nonconstant and therefore it must be strictly increasing.  $\square$

**5. Proof of Theorem 1.2.** We apply the content of the preceding sections to prove Theorem 1.2. Before that, we show that from the Markov renewal structure described in Section 3 one can extract a genuine renewal process, which will be a basic technical tool.

**5.1. From Markov renewals to true renewals.** Recalling the definition (3.1) of the contact set  $\tau$ , we introduce the subset  $\chi$  of the *adjacent contact points* defined by

$$(5.1) \quad \chi := \{i \in \mathbb{Z}^+ : \varphi_{i-1} = \varphi_i = 0\} \subset \tau \subset \mathbb{Z}^+,$$

where we set by definition  $\varphi_{-1} = \varphi_0 = 0$ , so that  $0 \in \chi$ . Note that  $\chi = \tau \cap (\tau + 1)$  and that the points of  $\chi$  are the random variables  $\{\chi_k\}_{k \geq 0}$  defined by

$$(5.2) \quad \chi_0 := 0, \quad \chi_{k+1} := \inf\{i > \chi_k : \varphi_{i-1} = \varphi_i = 0\}.$$

By (3.16) the event  $\mathcal{A}_N$  can be written as  $\mathcal{A}_N = \{N + 1 \in \chi\}$ ; therefore by Proposition 3.1 the partition function can be written as

$$(5.3) \quad Z_{\varepsilon, N}^a = \frac{e^{F^a(\varepsilon)N}}{\varepsilon^2} \cdot \mathcal{P}_\varepsilon^a(N + 1 \in \chi).$$

The reason for focusing on the process  $\{\chi_k\}_k$  is explained by the following proposition.

**PROPOSITION 5.1.** *The process  $\{\chi_k\}_{k \geq 0}$  under  $\mathcal{P}_\varepsilon^a$  is a classical (i.e., not Markov) renewal process, which is nonterminating for  $\varepsilon \geq \varepsilon_c^a$ .*

**PROOF.** We introduce the epochs  $\{\zeta_k\}_k$  of return to zero of the process  $\{J_k\}_k$ :

$$(5.4) \quad \zeta_0 := 0, \quad \zeta_{n+1} := \inf\{k > \zeta_n : J_k = 0\}.$$

It is clear that the variables  $\{\zeta_k - \zeta_{k-1}\}_{k \in \mathbb{N}}$  are i.i.d. under  $\mathcal{P}_\varepsilon^a$ , because  $\{J_k\}_k$  is a Markov chain. Observe that  $\chi_k = \tau_{\zeta_k}$ , for every  $k \geq 0$ , as it follows by (3.2), (3.3) and (5.2). This fact implies that also the variables  $\{\chi_k - \chi_{k-1}\}_{k \in \mathbb{N}} = \{\tau_{\zeta_k} - \tau_{\zeta_{k-1}}\}_{k \in \mathbb{N}}$  are i.i.d. under  $\mathcal{P}_\varepsilon^a$ , because the transition kernel of the process  $\{(\tau_k, J_k)\}_{k \geq 0}$  is a function of  $(n - m)$ ; cf. the r.h.s. of (3.12). Therefore  $\{\chi_k\}_{k \geq 0}$  under  $\mathcal{P}_\varepsilon^a$  is a genuine renewal process.

We have already observed that for  $\varepsilon \geq \varepsilon_c^a$  the Markov chain  $\{J_k\}_{k \geq 0}$  is *positive recurrent*, because for  $\varepsilon \geq \varepsilon_c^a$  the probability measure  $v_\varepsilon^a$  defined in (4.13) is by construction an invariant measure for its transition kernel  $D_{x, dy}^{a, \varepsilon}$ ; cf. (3.13) and (3.9). Since  $v_\varepsilon^a(\{0\}) > 0$ , the state 0 is an *atom* for  $\{J_k\}_{k \geq 0}$  and then it is a classical result that the returns of  $\{J_k\}_{k \geq 0}$  to 0 are not only  $\mathcal{P}_\varepsilon^a$ -a.s. finite, but also integrable:

$$(5.5) \quad \mathcal{P}_\varepsilon^a(\zeta_1 < \infty) = 1, \quad \mathbb{E}_\varepsilon^a(\zeta_1) = \frac{1}{v_\varepsilon^a(\{0\})} < \infty, \quad (\varepsilon \geq \varepsilon_c^a);$$

cf. [26], Chapter 5. Therefore also  $\chi_1 = \tau_{\zeta_1}$  is  $\mathcal{P}_\varepsilon^a$ -a.s. a finite random variable for  $\varepsilon \geq \varepsilon_c^a$ .  $\square$

**5.2. Proof of Theorem 1.2.** We start showing that the quantities  $F^a(\varepsilon)$  and  $\varepsilon_c^a$ , that were defined in (4.10) and appear in the definition (3.9) of the kernel  $K_{x,dy}^{a,\varepsilon}(n)$ , indeed coincide with the corresponding quantities defined in the Introduction. By (5.3) we have

$$\frac{1}{N} \log Z_{\varepsilon,N}^a = F^a(\varepsilon) - \frac{2}{N} \log \varepsilon + \frac{1}{N} \log \mathcal{P}_\varepsilon^a(N+1 \in \chi).$$

Since by (1.9) we have  $\liminf_{N \rightarrow \infty} N^{-1} \log Z_{\varepsilon,N}^a \geq 0$ , when  $F^a(\varepsilon) = 0$  the trivial bound  $\mathcal{P}_\varepsilon^a(N+1 \in \chi) \leq 1$  shows that (1.8) holds true. Therefore to complete the identification of  $F^a$  it suffices to show that when  $F^a(\varepsilon) > 0$ , that is, when  $\varepsilon > \varepsilon_c^a$ , we have

$$(5.6) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{P}_\varepsilon^a(N+1 \in \chi) = 0.$$

However, it is well known (and easy to prove) that this relation is true in complete generality for any nonterminating aperiodic renewal process, and Proposition 5.1 shows that for  $\varepsilon > \varepsilon_c^a$  the process  $\{\chi_k\}_k$  under  $\mathcal{P}_\varepsilon^a$  is indeed a genuine nonterminating renewal process, which is aperiodic because  $\mathcal{P}_\varepsilon^a(\chi_1 = 1) > 0$ . This completes the identification of  $F^a(\varepsilon)$ , as defined in (4.10), with the free energy defined by (1.8).

Completing the proof of Theorem 1.2 is now easy. By definition we have  $B_{x,dy}^{p,0} \geq B_{x,dy}^{w,0}$  and one checks easily that the nonnegative kernel  $B_{x,dy}^{p,0} - B_{x,dy}^{w,0}$  has strictly positive spectral radius. Then Lemma 4.2 with  $A_{x,dy} = B_{x,dy}^{w,0}$  and  $C_{x,dy} = B_{x,dy}^{p,0} - B_{x,dy}^{w,0}$  yields  $\delta^w(0) < \delta^p(0)$ , that is,  $\varepsilon_c^w > \varepsilon_c^p$  by (4.10). The analyticity of  $F^a(\cdot)$  on  $(\varepsilon_c^a, \infty)$  has been already discussed in Section 4.2. It remains to prove (1.11). Note that  $e^{-\lambda} F_{x,dy}^a(1) \leq B_{x,dy}^{a,\lambda} \leq e^{-\lambda} B_{x,dy}^{a,0}$  by (4.4), hence  $c e^{-\lambda} \leq \delta^a(\lambda) \leq c' e^{-\lambda}$  by (4.9), with  $c, c' > 0$ . Taking  $\lambda = (\delta^a)^{-1}(1/\varepsilon)$  and using (4.10) we finally obtain  $F^a(\varepsilon) \sim \log(\varepsilon)$  as  $\varepsilon \rightarrow \infty$ .

**6. Proof of Theorem 1.3.** In this section we prove Theorem 1.3, that is, we show that (1.16) holds true:

$$\liminf_{N \rightarrow \infty} \mathbb{E}_{\varepsilon_c^w, N}^w \left( \frac{\ell_N}{N} \right) > 0,$$

where we have used (1.13). We introduce the quantity

$$(6.1) \quad \iota_N := \#\{\chi \cap [0, N]\} = \max\{k \geq 0 : \chi_k \leq N\},$$

and since  $\ell_N \geq \iota_N$ , by Proposition 3.1 it is sufficient to show that

$$(6.2) \quad \liminf_{N \rightarrow \infty} \mathbb{E}_{\varepsilon_c^w, N}^w \left( \frac{\iota_N}{N} \right) = \liminf_{N \rightarrow \infty} \mathbb{E}_{\varepsilon_c^w}^w \left( \frac{\iota_N}{N} \mid \mathcal{A}_N \right) > 0.$$

We recall that the process  $\{\chi_k\}_k$  under  $\mathcal{P}_{\varepsilon_c^w}^w$  is a classical aperiodic renewal process by Proposition 5.1. Moreover we claim that  $b := \mathbb{E}_{\varepsilon_c^w}^w(\chi_1) < \infty$ . Then by

the strong law of large numbers we have  $\iota_N/N \rightarrow 1/b$ ,  $\mathcal{P}_{\varepsilon_c^W}$ -a.s., and by the renewal theorem  $\mathcal{P}_{\varepsilon_c^W}^W(\mathcal{A}_N) = \mathcal{P}_{\varepsilon_c^W}^W(N+1 \in \chi) \rightarrow 1/b > 0$  as  $N \rightarrow \infty$ . It follows that

$$\mathcal{E}_{\varepsilon_c^W}^W\left(\frac{\iota_N}{N} \mid \mathcal{A}_N\right) \rightarrow \frac{1}{b} > 0 \quad (N \rightarrow \infty),$$

and (6.2) follows. It only remains to check that  $\mathcal{E}_{\varepsilon_c^W}^W(\chi_1) < \infty$ .

**6.1. A formula for  $\mathcal{E}_{\varepsilon_c^W}^W(\chi_1)$ .** The dependency on  $\varepsilon_c^W$  will be omitted from now on for notational convenience. We recall that  $\chi_1 := \tau_1 + \dots + \tau_{\zeta_1}$ , where  $\zeta_1 := \inf\{n \geq 0 : J_n = 0\}$ . We introduce the kernel  $\widehat{K}_{x, dy}^W(n) := K_{x, dy}^W(n) \mathbf{1}_{(y \neq 0)} (= K_{x, dy}^W(n) \mathbf{1}_{(n \geq 2)})$  that gives the transition probabilities of the process  $\{(\tau_k, J_k)\}_k$  before the chain  $\{J_k\}_k$  comes back to zero. Summing over the possible values of the variable  $\zeta_1$ , we obtain the expression

$$(6.3) \quad \mathcal{P}^W(\chi_1 = n) = \int_{y \in \mathbb{R}} \sum_{k=0}^{\infty} (\widehat{K}^W)_{0, dy}^{*k}(n-1) \cdot K_{y, \{0\}}^W(1),$$

where the convolution  $*$  of kernels is defined in Section 1.7. Now observe that

$$\begin{aligned} & \sum_{m \in \mathbb{N}} m \cdot \sum_{k=0}^{\infty} (\widehat{K}^W)_{0, dy}^{*k}(m) \\ &= \sum_{k=1}^{\infty} \sum_{m=1}^{\infty} \sum_{\substack{t_1, \dots, t_k \in \mathbb{N} \\ t_1 + \dots + t_k = m}} (t_1 + \dots + t_k) (\widehat{K}^W(t_1) \circ \dots \circ \widehat{K}^W(t_k))_{0, dy} \\ &= \sum_{k=1}^{\infty} \sum_{i=1}^k ((\widehat{D}^W)^{\circ(i-1)} \circ \widehat{M}^W \circ (\widehat{D}^W)^{\circ(k-i)})_{0, dy} \\ &= ((1 - \widehat{D}^W)^{-1} \circ \widehat{M}^W \circ (1 - \widehat{D}^W)^{-1})_{0, dy}, \end{aligned}$$

where  $\widehat{D}_{x, dy}^W := \sum_{n \in \mathbb{N}} \widehat{K}_{x, dy}^W(n) = D_{x, dy}^W \mathbf{1}_{(y \neq 0)}$  and  $\widehat{M}_{x, dy}^W := \sum_{n \in \mathbb{N}} n \cdot \widehat{K}_{x, dy}^W(n)$ . Notice that  $K_{y, \{0\}}^W(1) = D_{y, \{0\}}^W$  by (3.13), therefore by (6.3) we have

$$\begin{aligned} \mathcal{E}^W(\chi_1) &= \sum_{n \in \mathbb{N}} n \cdot \mathcal{P}^W(\chi_1 = n) = 1 + \sum_{n \in \mathbb{N}} (n-1) \cdot \mathcal{P}^W(\chi_1 = n) \\ &= 1 + ((1 - \widehat{D}^W)^{-1} \circ \widehat{M}^W \circ (1 - \widehat{D}^W)^{-1} \circ D^W)_{0, \{0\}}. \end{aligned}$$

We recall that  $D_{x, dy}^W$  is the transition kernel of the Markov chain  $\{J_k\}_k$  under  $\mathcal{P}^W$ , which is positive recurrent with invariant probability measure  $\nu^W(\cdot) = \nu_{\varepsilon_c^W}^W(\cdot)$  defined in (4.13). Since  $\nu^W(\{0\}) > 0$ , the state 0 is an atom for  $\{J_k\}_k$  and therefore by

[26] the following formulas hold: for all  $x, y \in \mathbb{R}$

$$(1 - \widehat{D}^w)^{-1}_{0,dx} = \frac{v^w(dx)}{v^w(\{0\})} = \frac{v^w(x)w^w(x)}{v^w(0)w^w(0)}\mu(dx),$$

$$((1 - \widehat{D}^w)^{-1} \circ D^w)_{y,\{0\}} = 1.$$

Then we finally come to the expression

$$(6.4) \quad \mathcal{E}^w(\chi_1) = 1 + \varepsilon_c^w \int_{x,y \in \mathbb{R}^+} \mu(dx) \mu(dy) \frac{w^w(x)}{w^w(0)} \left( \sum_{n \in \mathbb{N}} n f_{x,y}^w(n) \right) \frac{v^w(y)}{v^w(0)}.$$

6.2. *Finiteness of  $\mathcal{E}_{\varepsilon_c^w}^w(\chi_1)$ .* First we state two relations that we prove below: for both  $a \in \{p, w\}$  and for all  $\varepsilon > 0$ :

$$(6.5) \quad \begin{aligned} \int_{\mathbb{R}} \mu(dx) v_{\varepsilon}^a(x) \log(x^2 \vee 1) &< \infty, \\ \int_{\mathbb{R}} \mu(dx) w_{\varepsilon}^a(x) \log(x^2 \vee 1) &< \infty. \end{aligned}$$

We aim at showing that the r.h.s. of (6.4) is finite. We start considering the terms in the sum with  $n \geq (x^2 \vee y^2)$ : applying (3.8), (2.12) and the first relation in (4.2) we obtain

$$\sum_{n \geq (x^2 \vee y^2)} n f_{x,y}^w(n) \leq \sum_{n \geq (x^2 \vee y^2)} \frac{(const.)}{n (\log n)^{c_+}} \leq (const.'),$$

because  $c_+ > 1$  and hence the sum converges. Therefore the contribution to the r.h.s. of (6.4) is bounded by

$$\begin{aligned} \varepsilon_c^w (const.') \int_{x,y \in \mathbb{R}^+} \mu(dx) \mu(dy) \frac{w^w(x)}{w^w(0)} \frac{v^w(y)}{v^w(0)} \\ \leq \varepsilon_c^w (const.') \frac{\|v^w\|_1 \|w^w\|_1}{v^w(0) w^w(0)}, \end{aligned}$$

where  $\|\cdot\|_1$  denotes the norm in  $L^1(\mathbb{R}, d\mu)$ . Notice that  $\|v^w\|_1 < \infty$ ,  $\|w^w\|_1 < \infty$  by (6.5).

Next we deal with the terms in the r.h.s. of (6.4) with  $n < (x^2 \vee y^2)$ . From the bound  $f_{x,y}^w(n) \leq f_{x,y}^p(n) \leq C/n^2$  [cf. (4.2)], we have

$$\sum_{n < (x^2 \vee y^2)} n f_{x,y}^w(n) \leq C \sum_{n < (x^2 \vee y^2)} \frac{1}{n} \leq C(\log(x^2 \vee 1) + \log(y^2 \vee 1)),$$

and using again (6.5) we see that the r.h.s. of (6.4) is indeed finite.

6.3. *Proof of (6.5).* We focus on the first relation, the second one being analogous. By (4.12) we have for both  $a \in \{p, w\}$

$$v_\varepsilon^a(x) \leq (\text{const.}) \sum_{n \in \mathbb{N}} \int_{y \in \mathbb{R}} f_{x,y}^p(n) v_\varepsilon^a(y) \mu(dy),$$

because  $f_{x,y}^w(n) \leq f_{x,y}^p(n)$ . Setting  $u_n(y) := \int_{x \in \mathbb{R}} f_{x,y}^p(n) \log(x^2 \vee 1) \mu(dx)$  and applying the Cauchy–Schwarz inequality we get

$$\begin{aligned} \int_{x \in \mathbb{R}} \mu(dx) v_\varepsilon^a(x) \log(x^2 \vee 1) &\leq (\text{const.}) \sum_{n \in \mathbb{N}} \int_{y \in \mathbb{R}} \mu(dy) u_n(y) v_\varepsilon^a(y) \\ (6.6) \quad &\leq (\text{const.}) \|v_\varepsilon^a\|_2 \cdot \left( \sum_{n \in \mathbb{N}} \|u_n\|_2 \right), \end{aligned}$$

where  $\|v_\varepsilon^a\|_2 := (\int_{\mathbb{R}} (v_\varepsilon^a(x))^2 \mu(dx))^{1/2}$  and likewise for  $\|u_n\|_2$ . Setting for short  $f_{\mathbb{R},y}^p(n) := \int_{\mathbb{R}} f_{x,y}^p(n) \mu(dx)$ , by Jensen’s inequality we have

$$\begin{aligned} \|u_n\|_2^2 &= \int_{y \in \mathbb{R}} (u_n(y))^2 \mu(dy) = \int_{y \in \mathbb{R}} \mu(dy) \left( \int_{x \in \mathbb{R}} f_{x,y}^p(n) \log(x^2 \vee 1) \mu(dx) \right)^2 \\ &\leq \int_{y \in \mathbb{R}} \mu(dy) (f_{\mathbb{R},y}^p(n))^2 \int_{x \in \mathbb{R}} \frac{f_{x,y}^p(n)}{f_{\mathbb{R},y}^p(n)} \log^2(x^2 \vee 1) \mu(dx), \end{aligned}$$

and since  $f_{\mathbb{R},y}^p(n) \leq C/n^{3/2}$  by (4.2), Fubini’s theorem yields

$$\int_{y \in \mathbb{R}} (u_n(y))^2 \mu(dy) \leq \frac{C}{n^{3/2}} \frac{1}{n} \int_{x \in \mathbb{R}} n \cdot f_{x,\mathbb{R}}^p(n) \log^2(x^2 \vee 1) \mu(dx).$$

Observe that  $f_{\mathbb{R},\mathbb{R}}^p(n) := \int_{x \in \mathbb{R}} f_{x,\mathbb{R}}^p(n) \mu(dx) = 1/n$  by (4.3), therefore  $x \mapsto n \cdot f_{x,\mathbb{R}}^p(n)$  is a probability density; in fact it is the density of the random variable  $Z_n/n$  under  $\mathbf{P}^{(0,0)}$ , as follows from (3.7). Since  $\log^2(x^2 \vee 1) \leq \log^2(x^2 \vee e)$  and the function  $z \mapsto \log^2(z)$  is concave on the half-line  $[e, \infty)$ , by Jensen’s inequality we get

$$\begin{aligned} \int_{x \in \mathbb{R}} n \cdot f_{x,\mathbb{R}}^p(n) \log^2(x^2 \vee 1) \mu(dx) &\leq \log^2 \left( \int_{x \in \mathbb{R}} n \cdot f_{x,\mathbb{R}}^p(n) (x^2 \vee e) \mu(dx) \right) \\ &\leq \log^2 \left( \mathbf{E}^{(0,0)} \left( \frac{Z_n^2}{n^2} \vee e \right) \right) \\ &\leq \log^2(e + (\text{const.}) \cdot n), \end{aligned}$$

because by (2.2) we have  $\mathbf{E}^{(0,0)}(Z_n^2) \sim \sigma^2 \cdot n^3/3$  as  $n \rightarrow \infty$ . It follows that

$$\int_{y \in \mathbb{R}} (u_n(y))^2 \mu(dy) \leq \frac{C}{n^{5/2}} \log^2(e + (\text{const.}) \cdot n)$$

and therefore

$$\sum_{n \in \mathbb{N}} \|u_n\|_2 = \sum_{n \in \mathbb{N}} \sqrt{\int_{y \in \mathbb{R}} (u_n(y))^2 \mu(dy)} < \infty.$$

Looking back to (6.6), equation (6.5) is proven.

**7. Proof of Theorem 1.4.** In this section we prove Theorem 1.4. We start proving (1.18), that we rewrite for convenience: there exists a positive constant  $c_5$  such that for small  $\delta > 0$

$$(7.1) \quad F^p(\varepsilon_c^p + \delta) \geq c_5 \frac{\delta}{\log 1/\delta}.$$

To this purpose, we observe that the following relation holds true for all  $\varepsilon, \varepsilon' > 0$ :

$$\frac{Z_{\varepsilon', N}^p}{Z_{\varepsilon, N}^p} = \mathbb{E}_{\varepsilon, N}^p \left( \left( \frac{\varepsilon'}{\varepsilon} \right)^{\ell_N - 1} \right),$$

where we recall that  $\ell_N = \#\{i \in \{1, \dots, N\} : \varphi_i = 0\}$ . This relation follows easily from the definition (1.5) of our model, because the differences between  $Z_{\varepsilon, N}^p$  and  $Z_{\varepsilon', N}^p$  are only given by the different size of the pinning reward for each time the path touches the  $x$ -axis (see also Appendix A). Then we have the easy lower bound

$$\frac{Z_{\varepsilon', N}^p}{Z_{\varepsilon, N}^p} \geq \mathbb{E}_{\varepsilon, N}^p \left( \left( \frac{\varepsilon'}{\varepsilon} \right)^{\iota_N - 1} \right) = \mathbb{E}_{\varepsilon}^p \left( \left( \frac{\varepsilon'}{\varepsilon} \right)^{\iota_N - 1} \mid N + 1 \in \chi \right),$$

where we recall that  $\iota_N = \#\{i \in \{1, \dots, N\} : \chi_i = 0\}$  and we have applied Proposition 3.1. Setting  $\varepsilon = \varepsilon_c^p$ ,  $\varepsilon' = \varepsilon_c^p + \delta$ , taking logarithms and letting  $N \rightarrow \infty$  yields

$$\begin{aligned} F^p(\varepsilon_c^p + \delta) &\geq F^p(\varepsilon_c^p) + \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}_{\varepsilon_c^p}^p \left( \left( \frac{\varepsilon_c^p + \delta}{\varepsilon_c^p} \right)^{\iota_N - 1} \mid N + 1 \in \chi \right) \\ &=: F^p(\varepsilon_c^p) + G(\delta), \end{aligned}$$

and since  $F^p(\varepsilon_c^p) = 0$  we get

$$F^p(\varepsilon_c^p + \delta) \geq G(\delta).$$

Now notice that under  $\mathcal{P}_{\varepsilon_c^p}^p$  the process  $\{\chi_k\}_k$  is a classical, nonterminating renewal process, by Proposition 5.1. Therefore  $G(\delta)$  is just the free energy of a *classical pinning model*, that is, a renewal process rewarded  $\frac{\varepsilon_c^p + \delta}{\varepsilon_c^p}$  at each renewal epoch. The asymptotic behavior of  $G(\delta)$  is then given by Theorem 2.1 of [13]: since  $\mathcal{P}_{\varepsilon_c^p}^p(\chi_1 = n) \sim c/n^2$  as  $n \rightarrow \infty$ , as we prove in Proposition 7.1 below, we have  $G(\delta) \sim c'\delta/\log(1/\delta)$  as  $\delta \downarrow 0$ , and (7.1) is proven.

It only remains to show that (1.17) holds true, that is,

$$(7.2) \quad \limsup_{\varepsilon \downarrow \varepsilon_c^P} \limsup_{N \rightarrow \infty} \mathcal{E}_\varepsilon^P \left( \frac{\ell_N}{N} \mid \mathcal{A}_N \right) = 0,$$

where we have used (1.13) and Proposition 3.1. The idea is to focus first on the variable  $\iota_N$ , which is easier to handle, and then to make the comparison with  $\ell_N$ .

7.1. *From  $\iota_N$  to  $\ell_N$ .* We recall that the process  $\{\chi_k\}_k$  under the law  $\mathcal{P}_\varepsilon^P$  is a classical aperiodic renewal process; see Proposition 5.1. We introduce the step law

$$(7.3) \quad q_\varepsilon(n) := \mathcal{P}_\varepsilon^P(\chi_1 = n),$$

whose asymptotic behavior as  $n \rightarrow \infty$ , when  $\varepsilon$  is close to  $\varepsilon_c^P$ , is given by the following:

**PROPOSITION 7.1.** *There exists  $\alpha > 0$  such that for every  $\varepsilon \in [\varepsilon_c^P, \varepsilon_c^P + \alpha]$  we have*

$$(7.4) \quad q_\varepsilon(n) \sim \frac{C_\varepsilon}{n^2} \exp(-F^P(\varepsilon) \cdot n) \quad (n \rightarrow \infty),$$

where  $C_\varepsilon \in (0, \infty)$  is a continuous function of  $\varepsilon$ .

We postpone the proof to the next paragraphs; for the moment we focus on the consequences. We assume in the following that  $\varepsilon_c^P < \varepsilon \leq \varepsilon_c^P + \alpha$ .

Let us set  $G_\varepsilon := \mathcal{E}_\varepsilon^P(\chi_1) < \infty$  by Proposition 7.1. A standard Tauberian theorem (cf. [5], Theorem 1.7.1) gives the asymptotic behavior of  $G_\varepsilon$  as  $\varepsilon \downarrow \varepsilon_c^P$ :

$$(7.5) \quad G_\varepsilon \sim (\text{const.}) \log \frac{1}{F^P(\varepsilon)} \quad (\varepsilon \downarrow \varepsilon_c^P).$$

Notice that the classical Renewal theorem yields

$$(7.6) \quad \mathcal{P}_\varepsilon^P(\mathcal{A}_N) = \mathcal{P}_\varepsilon^P(N+1 \in \chi) \rightarrow \frac{1}{G_\varepsilon} > 0 \quad (N \rightarrow \infty).$$

Therefore by the weak law of large numbers for the process  $\{\chi_k\}_k$  we have

$$(7.7) \quad \begin{aligned} \mathcal{P}_\varepsilon^P \left( \frac{\iota_N}{N} \geq \frac{2}{G_\varepsilon} \mid \mathcal{A}_N \right) &= \frac{\mathcal{P}_\varepsilon^P(\chi_{\lfloor 2N/G_\varepsilon \rfloor} \leq N)}{\mathcal{P}_\varepsilon^P(\mathcal{A}_N)} \\ &= \frac{\mathcal{P}_\varepsilon^P(\chi_{\lfloor 2N/G_\varepsilon \rfloor}/\lfloor 2N/G_\varepsilon \rfloor \leq G_\varepsilon/2)}{\mathcal{P}_\varepsilon^P(\mathcal{A}_N)} \xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

We recall that  $\zeta_k$  denotes the epoch of the  $k$ th return of the process  $\{J_k\}_k$  to the state zero [cf. 5.4], and that  $\{\zeta_k\}_{k \geq 0}$  under  $\mathcal{P}_\varepsilon^P$  is a nonterminating renewal process

with finite mean  $m_\varepsilon := \mathbb{E}_\varepsilon^P(\zeta_1) = 1/v_\varepsilon^P(\{0\}) < \infty$ ; cf. (5.5). Then by the weak law of large numbers

$$(7.8) \quad \mathcal{P}_\varepsilon^P\left(\frac{\zeta k}{k} > 2m_\varepsilon\right) \rightarrow 0 \quad (k \rightarrow \infty).$$

We are finally ready to estimate  $\ell_N$ . A trivial bound yields

$$\begin{aligned} \mathcal{P}_\varepsilon^P\left(\frac{\ell_N}{N} \geq \frac{4m_\varepsilon}{G_\varepsilon} \mid \mathcal{A}_N\right) &\leq \mathcal{P}_\varepsilon^P\left(\frac{\ell_N}{N} \geq \frac{2}{G_\varepsilon} \mid \mathcal{A}_N\right) \\ &\quad + \mathcal{P}_\varepsilon^P\left(\frac{\ell_N}{N} \geq \frac{4m_\varepsilon}{G_\varepsilon}, \frac{\ell_N}{N} < \frac{2}{G_\varepsilon} \mid \mathcal{A}_N\right). \end{aligned}$$

The first term in the r.h.s. vanishes as  $N \rightarrow \infty$  by (7.7). For the second term we observe that by definition  $\ell_N = \zeta_{\ell_N}$  on the event  $\mathcal{A}_N$ , hence by an inclusion argument we have

$$\mathcal{P}_\varepsilon^P\left(\frac{\ell_N}{N} \geq \frac{4m_\varepsilon}{G_\varepsilon}, \frac{\ell_N}{N} < \frac{2}{G_\varepsilon} \mid \mathcal{A}_N\right) \leq \mathcal{P}_\varepsilon^P\left(\frac{\zeta_{\lfloor 2N/G_\varepsilon \rfloor}}{\lfloor 2N/G_\varepsilon \rfloor} > 2m_\varepsilon \mid \mathcal{A}_N\right) \xrightarrow{N \rightarrow \infty} 0,$$

having used (7.8) and (7.6). Since  $\ell_N/N \leq 1$ , we can finally write

$$\limsup_{N \rightarrow \infty} \mathbb{E}_\varepsilon^P\left(\frac{\ell_N}{N} \mid \mathcal{A}_N\right) \leq \frac{4m_\varepsilon}{G_\varepsilon} + \limsup_{N \rightarrow \infty} \mathcal{P}_\varepsilon^P\left(\frac{\ell_N}{N} \geq \frac{4m_\varepsilon}{G_\varepsilon} \mid \mathcal{A}_N\right) = \frac{4m_\varepsilon}{G_\varepsilon}.$$

Now observe that as  $\varepsilon \downarrow \varepsilon_c^P$  we have  $m_\varepsilon \rightarrow 1/v_{\varepsilon_c^P}^P(\{0\}) < \infty$  and moreover  $G_\varepsilon \rightarrow +\infty$  by (7.5). Then we let  $\varepsilon \downarrow \varepsilon_c^P$  in the last equation and (7.2) is proven.

**7.2. A Markov renewal theorem with infinite mean.** Before proving Proposition 7.1, we derive a generalized renewal theorem in our Markovian setting. Since the steps are more transparent if carried out in a general setting, we assume that we are given a kernel  $\mathsf{A}_{x,dy}(n)$  satisfying the following assumptions:

1. the spectral radius of  $G_{x,dy} := \sum_{n \in \mathbb{N}} \mathsf{A}_{x,dy}(n)$  is strictly less than 1;
2. we have  $\mathsf{A}_{x,dy}(n) \sim L_{x,dy}/n^2$  as  $n \rightarrow \infty$ , for some kernel  $L_{x,dy}$  (for the precise meaning of this relation we refer to Section 1.7), and moreover  $\mathsf{A}_{x,dy}(n) \leq cL_{x,dy}/n^2$  for every  $n \in \mathbb{N}$ , where  $c$  is a positive constant;
3. there exists  $\beta > 1$  such that  $((1 - \beta G)^{-1} \circ L \circ (1 - \beta G)^{-1})_{x,F} < \infty$  for every  $x \in \mathbb{R}$  and for every bounded Borel set  $F \subset \mathbb{R}$ .

The result we are going to prove is the following asymptotic relation:

$$(7.9) \quad \sum_{k=0}^{\infty} \mathsf{A}_{x,dy}^{*k}(n) \sim \frac{((1 - G)^{-1} \circ L \circ (1 - G)^{-1})_{x,dy}}{n^2} \quad (n \rightarrow \infty).$$

The path we follow is close to [8], Section 3.4. We start proving by induction the following bound: for all  $k, n \in \mathbb{N}$  and  $x, y \in \mathbb{R}$

$$(7.10) \quad \mathsf{A}_{x,dy}^{*k}(n) \leq c k^2 \frac{\sum_{i=0}^{k-1} (G^{\circ i} \circ L \circ G^{\circ[(k-1)-i]})_{x,dy}}{n^2}.$$

The  $k = 1$  case holds by assumption (2). Then we consider the even- $k$  case: by the definition of the convolution  $*$  we have

$$\mathsf{A}_{x,dy}^{*(2k)}(n) \leq \sum_{h=1}^{\lceil n/2 \rceil} \int_{z \in \mathbb{R}} (\mathsf{A}_{x,dz}^{*k}(h) \cdot \mathsf{A}_{z,dy}^{*k}(n-h) + \mathsf{A}_{x,dz}^{*k}(n-h) \cdot \mathsf{A}_{z,dy}^{*k}(h)).$$

Observing that  $\sum_{h \in \mathbb{N}} \mathsf{A}_{x,dy}^{*k}(h) = G_{x,dy}^{\circ k}$  and applying the inductive step we get

$$\begin{aligned} \mathsf{A}_{x,dy}^{*(2k)}(n) &\leq \frac{c k^2}{(n/2)^2} \left[ \left( G^{\circ k} \circ \left( \sum_{i=0}^{k-1} G^{\circ i} \circ L \circ G^{\circ[(k-1)-i]} \right) \right)_{x,dy} \right. \\ &\quad \left. + \left( \left( \sum_{i=0}^{k-1} G^{\circ i} \circ L \circ G^{\circ[(k-1)-i]} \right) \circ G^{\circ k} \right)_{x,dy} \right] \\ &= \frac{c(2k)^2}{n^2} \sum_{i=0}^{2k-1} (G^{\circ i} \circ L \circ G^{\circ[(2k-1)-i]})_{x,dy}, \end{aligned}$$

so that (7.10) is proven, the odd- $k$  case being analogous. Note that, choosing a constant  $c' > 0$  such that  $k^2 \leq c' \beta^k$  for every  $k$ , assumption (3) yields that for every  $x \in \mathbb{R}$  and for every bounded Borel set  $F \subset \mathbb{R}$

$$\begin{aligned} (7.11) \quad &\sum_{k=1}^{\infty} \sum_{i=0}^{k-1} k^2 (G^{\circ i} \circ L \circ G^{\circ[(k-1)-i]})_{x,F} \\ &\leq c' \sum_{k=1}^{\infty} \sum_{i=0}^{k-1} ((\beta G)^{\circ i} \circ L \circ (\beta G)^{\circ[(k-1)-i]})_{x,F} \\ &= c' ((1 - \beta G)^{-1} \circ L \circ (1 - \beta G)^{-1})_{x,F} < \infty. \end{aligned}$$

Next we claim that

$$(7.12) \quad \mathsf{A}_{x,dy}^{*k}(n) \sim \frac{\sum_{i=0}^{k-1} (G^{\circ i} \circ L \circ G^{\circ[(k-1)-i]})_{x,dy}}{n^2} \quad (n \rightarrow \infty).$$

We proceed by induction: the  $k = 1$  case holds by assumption (1), while for general  $k$

$$\begin{aligned} \mathsf{A}_{x,dy}^{*k}(n) &= \sum_{h=1}^{n/2} \int_{z \in \mathbb{R}} (\mathsf{A}_{x,dz}^{*(k-1)}(h) \cdot \mathsf{A}_{z,dy}(n-h) \\ &\quad + \mathsf{A}_{x,dz}^{*(k-1)}(n-h) \cdot \mathsf{A}_{z,dy}(h)). \end{aligned}$$

Applying the induction step and using dominated convergence, thanks to (7.10)

and (7.11), we have [observe that  $\sum_{h \in \mathbb{N}} A_{x,dy}^{*m}(h) = G_{x,dy}^{\circ m}$ ]

$$\begin{aligned} n^2 A_{x,dy}^{*k}(n) &\xrightarrow{n \rightarrow \infty} \left( G^{\circ(k-1)} \circ L + \left( \sum_{i=0}^{k-2} G^{\circ i} \circ L \circ G^{\circ[(k-2)-i]} \right) \circ G \right)_{x,dy} \\ &= \sum_{i=0}^{k-1} (G^{\circ i} \circ L \circ G^{\circ[(k-1)-i]})_{x,dy}, \end{aligned}$$

and (7.12) is proven. Finally we can write as  $n \rightarrow \infty$

$$\begin{aligned} n^2 \sum_{k \geq 0} A_{x,dy}^{*k}(n) &\rightarrow \sum_{k \geq 0} \sum_{i=0}^{k-1} (G^{\circ i} \circ L \circ G^{\circ[(k-1)-i]})_{x,dy} \\ &= ((1-G)^{-1} \circ L \circ (1-G)^{-1})_{x,dy}, \end{aligned}$$

where we have applied (7.12) and again dominated convergence, using (7.10) and (7.11). This completes the proof of (7.9).

**7.3. Proof of Proposition 7.1.** We start from a close analog of (6.3), namely

$$\begin{aligned} (7.13) \quad q_\varepsilon(n) &= \int_{y \in \mathbb{R}} \left( \sum_{k=0}^{\infty} (\widehat{K}_{x,dy}^{p,\varepsilon})_{0,dy}^{*k}(n-1) \right) \cdot K_{y,\{0\}}^{p,\varepsilon}(1) \\ &= \varepsilon \cdot e^{-F(\varepsilon)n} \cdot \int_{y \in \mathbb{R}} \left( \sum_{k=0}^{\infty} \varepsilon^k (\widehat{F}_{x,dy}^p)_{0,dy}^{*k}(n-1) \right) \cdot F_{y,\{0\}}^p(1), \end{aligned}$$

where we have set  $\widehat{K}_{x,dy}^{p,\varepsilon}(n) := K_{x,dy}^{p,\varepsilon}(n) \mathbf{1}_{(y \neq 0)}$  and  $\widehat{F}_{x,dy}^p(n) := F_{x,dy}^p(n) \mathbf{1}_{(y \neq 0)}$ . What we need is the asymptotic behavior as  $n \rightarrow \infty$  of the r.h.s. of (7.13) and to this purpose we are going to apply the results of Section 7.2 to the kernel  $A_{x,dy}(n) = \varepsilon \cdot \widehat{F}_{x,dy}^p(n)$ .

We need to check that the assumptions (1)–(3) are satisfied. The asymptotic behavior of  $\widehat{F}_{x,dy}^p(n)$  is obtained by (4.1):

$$(7.14) \quad \widehat{F}_{x,dy}^p(n) \sim \frac{c}{n^2} dy \quad (n \rightarrow \infty),$$

and from (4.2) we see that assumption (2) is checked. We set for simplicity

$$\widehat{B}_{x,dy}^p := \sum_{n \in \mathbb{N}} \widehat{F}_{x,dy}^p(n) = B_{x,dy}^{p,0} \mathbf{1}_{(y \neq 0)},$$

where the kernel  $B_{x,dy}^{p,0}$  was defined in (4.4). The spectral radius of  $\varepsilon_c^p \cdot B_{x,dy}^{p,0}$  equals 1 by the very definition of  $\varepsilon_c^p$ . Then applying Lemma 4.2 with  $A_{x,dy} = \varepsilon_c^p \cdot \widehat{B}_{x,dy}^p$  and  $C_{x,dy} = \varepsilon_c^p \cdot B_{x,dy}^{p,0} - \varepsilon_c^p \cdot \widehat{B}_{x,dy}^p$  [it is easily seen that the spectral radius of  $C_{x,dy} = \varepsilon_c^p \cdot B_{x,dy}^{p,0} \mathbf{1}_{(y=0)} = \varepsilon_c^p \cdot e^{-V(x)} \delta_0(dy)$  is strictly positive] we have that the

spectral radius of  $\varepsilon_c^p \cdot \widehat{B}_{x,dy}^p$  is strictly smaller than 1. By continuity there exists  $\alpha > 0$  such that the spectral radius of  $\varepsilon \cdot \widehat{B}_{x,dy}^p$  is strictly smaller than 1 for every  $\varepsilon \in [\varepsilon_c^p, \varepsilon_c^p + \alpha]$ . Then assumption (1) is verified and it only remains to check assumption (3), that is,

$$(7.15) \quad \left( \int_{y \in \mathbb{R}} (1 - \beta \varepsilon \widehat{B}_{x,dy}^p)^{-1} \right) \left( \int_{z \in \mathbb{R}} \mu(dz) \int_{w \in \mathcal{A}} (1 - \beta \varepsilon \widehat{B}_{z,dw}^p)^{-1} v(w) \right) < \infty,$$

for some  $\beta > 1$ . Let us focus on the first integral: we can write

$$\int_{y \in \mathbb{R}} (1 - \beta \varepsilon \widehat{B}_{x,dy}^p)^{-1} = 1 + \sum_{n \geq 0} (\beta \varepsilon) \int_{z \in \mathbb{R}} (\beta \varepsilon \widehat{B}_{x,dz}^p)^{\circ n} g(z),$$

where  $g(z) := \int_{y \in \mathbb{R}} \widehat{B}_{z,dy}^p$ . We choose  $\beta$  sufficiently close to 1 so that the spectral radius of  $\beta \varepsilon \widehat{B}^p$ , let us call it  $\rho$ , is strictly smaller than 1. Denoting by  $\|\cdot\|$  the operator norm in  $L^2(\mathbb{R}, d\mu)$ , a classical result gives  $\|(\beta \varepsilon \widehat{B}^p)^{\circ n}\|^{1/n} \rightarrow \rho < 1$  as  $n \rightarrow \infty$ ; cf. [20], Chapter III, Section 6.2. Therefore, if we show that  $g(\cdot) \in L^2(\mathbb{R}, d\mu)$ , we obtain

$$\int_{y \in \mathbb{R}} (1 - \beta \varepsilon \widehat{B}_{x,dy}^p)^{-1} \leq 1 + (\beta \varepsilon) \sum_{n \geq 0} \|(\beta \varepsilon \widehat{B}^p)^{\circ n}\| \|g\| < \infty.$$

To prove that  $g(\cdot) \in L^2(\mathbb{R}, d\mu)$ , we observe that  $g(z) = \sum_{n \in \mathbb{N}} \int_{y \in \mathbb{R}} \widehat{F}_{z,dy}^p(n)$  so that

$$\int_{x \in \mathbb{R}} g(x)^2 \mu(dx) = \sum_{n \in \mathbb{N}} \sum_{m \in \mathbb{N}} \int_{x \in \mathbb{R}} \left( \int_{y \in \mathbb{R}} \widehat{F}_{x,dy}^p(n) \right) \left( \int_{z \in \mathbb{R}} \widehat{F}_{x,dz}^p(m) \right) \mu(dx).$$

Using the symmetry in  $n, m$  and applying (4.2) and (4.3), we finally obtain

$$\int_{x \in \mathbb{R}} g(x)^2 \mu(dx) \leq 2 \sum_{m \in \mathbb{N}} \sum_{n \geq m} \frac{C}{n^{3/2} m} < \infty.$$

With similar arguments one shows that also the second integral term in (7.15) is finite.

We can finally apply (7.9) in our setting, getting

$$\sum_{k=0}^{\infty} \varepsilon^k (\widehat{F}^p)_{0,dy}^{*k}(n) \sim \frac{c\varepsilon}{n^2} \left( \int_{z \in \mathbb{R}} (1 - \varepsilon \widehat{B}_{0,dz}^p)^{-1} \right) \cdot \left( \int_{x \in \mathbb{R}} dx (1 - \varepsilon \widehat{B}_{x,dy}^p)^{-1} \right) \quad (n \rightarrow \infty).$$

Coming back to (7.13), we can apply dominated convergence thanks to (7.10) and (7.11); we thus obtain (7.4), with  $C_\varepsilon$  given by

$$C_\varepsilon := c\varepsilon^2 \left( \int_{z \in \mathbb{R}} (1 - \varepsilon \widehat{B}_{0,dz}^p)^{-1} \right) \cdot \left( \int_{x,y \in \mathbb{R}} dx (1 - \varepsilon \widehat{B}_{x,dy}^p)^{-1} e^{-V(y)} \right),$$

and the proof is complete.

## APPENDIX A: CONVEXITY OF THE FREE ENERGY

Recall the definition (1.12) of the contact number  $\ell_N$  and observe that in any case  $\ell_N \geq 1$  under  $\mathbb{P}_{\varepsilon, N}^a$ , because  $\varphi_N = 0$ . Setting  $\Omega^p := \mathbb{R}$  and  $\Omega^w := \mathbb{R}^+$ , from the definitions (1.5) and (1.7) of our models we can write for  $k \in \mathbb{Z}^+$

$$\begin{aligned} \mathbb{P}_{\varepsilon, N}^a(\ell_N = k + 1) \\ = \frac{\varepsilon^k}{Z_{\varepsilon, N}^a} \left\{ \sum_{\substack{A \subset \{1, \dots, N-1\} \\ |A|=k}} \int e^{-\mathcal{H}_{[-1, N+1]}(\varphi)} \prod_{m \in A} \delta_0(d\varphi_m) \prod_{n \in A^c} d\varphi_n \mathbf{1}_{(\varphi_n \in \Omega^a)} \right\}. \end{aligned}$$

The term in braces in the r.h.s. is a positive number depending on  $a, k, N$  but not on  $\varepsilon$ : let us call it  $C^a(k, N)$ . Summing over  $k = 0, \dots, N-1$  we obtain

$$Z_{\varepsilon, N}^a = \sum_{k=0}^{N-1} \varepsilon^k C^a(k, N), \quad \tilde{F}_N^a(t) = \frac{1}{N} \log \left( \sum_{k=0}^{N-1} e^{tk} C^a(k, N) \right),$$

where  $\tilde{F}_N^a(t) := F_N^a(e^t)$ ; cf. (1.10). Differentiating it the variable  $t$  we have

$$(\tilde{F}_N^a)'(t) = \frac{1}{N} \mathbb{E}_{e^t, N}^a(\ell_N), \quad (\tilde{F}_N^a)''(t) = \frac{1}{N} \text{var}_{\mathbb{P}_{e^t, N}^a}(\ell_N) \geq 0,$$

which proves (1.13) and the convexity of  $\tilde{F}_N^a(t)$ .

Now fix  $x \in [0, 1]$ . For every  $\alpha \geq 0$  we have

$$(A.1) \quad \mathbb{P}_{\varepsilon, N}^a(\ell_N/N > x) = \mathbb{P}_{\varepsilon, N}^a(e^{\alpha \ell_N} > e^{\alpha x N}) \leq e^{-\alpha x N} \mathbb{E}_{\varepsilon, N}^a(e^{\alpha \ell_N}).$$

Using the above relations we can write

$$\begin{aligned} \mathbb{E}_{\varepsilon, N}^a(e^{\alpha \ell_N}) &= \sum_{k=0}^{N-1} e^{\alpha(k+1)} \mathbb{P}_{\varepsilon, N}^a(\ell_N = k + 1) \\ &= e^\alpha \sum_{k=0}^{N-1} (e^\alpha \varepsilon)^k \frac{C^a(k, N)}{Z_{\varepsilon, N}^a} = e^\alpha \frac{Z_{e^\alpha \varepsilon, N}^a}{Z_{\varepsilon, N}^a}, \end{aligned}$$

therefore by (1.8) we have  $N^{-1} \log \mathbb{E}_{\varepsilon, N}^a(e^{\alpha \ell_N}) \rightarrow F^a(e^\alpha \varepsilon) - F^a(\varepsilon)$  as  $N \rightarrow \infty$ . If  $\varepsilon \neq \varepsilon_c^a$ , the free energy  $F^a$  is differentiable at  $\varepsilon$  by Theorem 1.2, therefore as  $\alpha \downarrow 0$  we have  $F^a(e^\alpha \varepsilon) - F^a(\varepsilon) = D^a(\varepsilon) \cdot \alpha + o(\alpha)$ , where  $D^a(\varepsilon) = \varepsilon \cdot (F^a)'(\varepsilon)$ . Plugging  $x = D^a(\varepsilon) + \delta$  (with  $\delta > 0$ ) and  $\alpha$  small into (A.1) we obtain

$$\mathbb{P}_{\varepsilon, N}^a(\ell_N/N > D^a(\varepsilon) + \delta) \leq e^{-(\text{const.})N}.$$

With almost identical arguments one shows that  $\mathbb{P}_{\varepsilon, N}^a(\ell_N/N < D^a(\varepsilon) - \delta) \leq e^{-(\text{const.}')N}$ , therefore (1.14) and (1.15) are proven.

## APPENDIX B: LLT FOR THE INTEGRATED RANDOM WALK

We are going to prove Proposition 2.3. We will stick for conciseness to the first relation in (2.10), the second one being analogous. We recall that the density of the random vector  $(B_1, I_1)$  is given by (2.7). Then its characteristic function  $\Psi(s, t)$  is given by

$$\Psi(s, t) = \exp(-s^2/2 - t^2/6 - st/2).$$

We denote by  $\psi_n(u, v) := \mathbf{E}^{(0,0)}[\exp(i(uY_n + vZ_n))]$  the characteristic function of  $(Y_n, Z_n)$ . An application of the Fourier-transform inversion formula gives

$$\begin{aligned} & |\sigma^2 n^2 \varphi_n^{(0,0)}(\sigma\sqrt{n}y, \sigma n^{3/2}z) - g(y, z)| \\ & \leq \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left| \psi_n\left(\frac{s}{\sigma\sqrt{n}}, \frac{t}{\sigma n^{3/2}}\right) - \Psi(s, t) \right| ds dt. \end{aligned}$$

The proof consists in showing that the r.h.s. vanishes as  $n \rightarrow \infty$ . More precisely, following the proof of Theorem 2 in [12], Section XV.5, we consider separately the three domains

$$\begin{aligned} \mathcal{D}_1 &= \{(s^2 + t^2) \leq A\}, \\ \mathcal{D}_2 &= \{A < (s^2 + t^2) \leq B^2 n\}, \\ \mathcal{D}_3 &= \{(s^2 + t^2) > B^2 n\}, \end{aligned}$$

and we show that, for a suitable choice of the positive constants  $A$  and  $B$  and for large  $n$ , the integral in the r.h.s. above is less than  $\varepsilon$  on each domain, for every fixed  $\varepsilon > 0$ .

*The domain  $\mathcal{D}_1$ .* Denoting by  $\xi(u) := \mathbf{E}[\exp(iuX_1)]$  the characteristic function of  $X_1$ , from (2.1) and (2.2) we have

$$(B.1) \quad \psi_n(u, v) = \mathbf{E}^{(0,0)}\left[\prod_{m=1}^n e^{i(u+mv)X_{n+1-m}}\right] = \prod_{m=1}^n \xi(u+mv).$$

Since by hypothesis  $\mathbf{E}(X_1) = 0$  and  $\mathbf{E}(X_1^2) = \sigma^2 \in (0, \infty)$ , it follows that

$$(B.2) \quad \xi(u) = 1 - \frac{\sigma^2}{2}u^2 + o(u^2) \quad (u \rightarrow 0),$$

hence, uniformly for  $(s, t)$  such that  $(s^2 + t^2) \leq A$ , we get from (B.1)

$$\psi_n\left(\frac{s}{\sigma\sqrt{n}}, \frac{t}{\sigma n^{3/2}}\right) = \exp\left[-\frac{1}{2} \sum_{m=1}^n \left(\frac{s}{\sqrt{n}} + m\frac{t}{n^{3/2}}\right)^2 + o(1)\right] \xrightarrow{n \rightarrow \infty} \Psi(s, t).$$

Therefore, for every choice of the parameter  $A$ , we can find  $n_0 = n_0(A)$  such that the integral  $\int_{\mathcal{D}_1} |\psi_n(\frac{s}{\sigma\sqrt{n}}, \frac{t}{\sigma n^{3/2}}) - \Psi(s, t)| ds dt$  is smaller than  $\varepsilon$  for  $n \geq n_0$ .

*The domain  $\mathcal{D}_2$ .* From B.2 it follows that  $|\xi(u)| = 1 - \sigma^2 u^2/2 + o(u^2)$  and therefore we can fix  $B > 0$  such that

$$|\xi(u)| \leq \exp\left(-\frac{\sigma^2}{4}u^2\right) \quad \text{for } |u| \leq \frac{2B}{\sigma}.$$

Using (B.1) and some rough bounds, we get for  $(s, t) \in \mathcal{D}_2$  and for all  $n \in \mathbb{N}$

$$\begin{aligned} \left| \psi_n\left(\frac{s}{\sigma\sqrt{n}}, \frac{t}{\sigma n^{3/2}}\right) \right| &\leq \exp\left[-\frac{1}{4} \sum_{m=1}^n \left(\frac{s}{\sqrt{n}} + m \frac{t}{n^{3/2}}\right)^2\right] \\ &\leq \exp\left[-\frac{1}{4}\left(s^2 + \frac{t^2}{3} + st\right)\right]. \end{aligned}$$

Then by the triangle inequality

$$\begin{aligned} \int_{\mathcal{D}_2} \left| \psi_n\left(\frac{s}{\sigma\sqrt{n}}, \frac{t}{\sigma n^{3/2}}\right) - \Psi(s, t) \right| ds dt \\ \leq \int_{\{s^2+t^2>A\}} (e^{-1/4(s^2+t^2/3+st)} + \Psi(s, t)) ds dt, \end{aligned}$$

and note that the r.h.s. can be made smaller than  $\varepsilon$  by choosing  $A$  large (this fixes  $A$ ).

*The domain  $\mathcal{D}_3$ .* By the triangle inequality we have

$$\begin{aligned} \int_{\mathcal{D}_3} \left| \psi_n\left(\frac{s}{\sigma\sqrt{n}}, \frac{t}{\sigma n^{3/2}}\right) - \Psi(s, t) \right| ds dt \\ \leq \int_{\{s^2+t^2>B^2n\}} \left| \psi_n\left(\frac{s}{\sigma\sqrt{n}}, \frac{t}{\sigma n^{3/2}}\right) \right| ds dt + \int_{\{s^2+t^2>B^2n\}} \Psi(s, t) ds dt. \end{aligned}$$

It is clear that the second integral in the r.h.s. vanishes as  $n \rightarrow \infty$  and it remains to show that the same is true for the first integral  $I_1$ . With the change of variables  $s/(\sigma\sqrt{n}) = r \cos(\theta)$ ,  $t/(\sigma\sqrt{n}) = r \sin(\theta)$  and using (B.1), we can rewrite  $I_1$  as

$$\begin{aligned} I_1 &= \sigma^2 n \int_{\{\theta \in [0, 2\pi), r > B\}} \left| \psi_n\left(r \cos(\theta), \frac{r \sin(\theta)}{n}\right) \right| r dr d\theta \\ (B.3) \quad &= \sigma^2 n \int_{\{\theta \in [0, 2\pi), r > B\}} \left\{ \prod_{m=1}^n |\xi|\left(r \left(\cos(\theta) + \frac{m}{n} \sin(\theta)\right)\right) \right\} r dr d\theta, \end{aligned}$$

where by  $|\xi|(\cdot)$  we mean the function  $u \mapsto |\xi(u)|$ . It is convenient to divide the domain of integration over  $\theta$  in the two subsets

$$\Theta_a := \{\theta \in [0, 2\pi) : |\cos(\theta)| > 1/2\}, \quad \Theta_b := \{\theta \in [0, 2\pi) : |\cos(\theta)| \leq 1/2\},$$

and to split accordingly the integral  $I_1 = I_{1,a} + I_{1,b}$ , with obvious notation. We are going to show that both  $I_{1,a}$  and  $I_{1,b}$  vanish as  $n \rightarrow \infty$ .

Since  $\xi(\cdot)$  is the characteristic function of the absolutely continuous random variables  $X_1$ , we have  $|\xi(u)| < 1$  for all  $u \neq 0$ ; cf. Lemma 4 in [12], Section XV.1, and moreover  $|\xi(u)| \rightarrow 0$  as  $u \rightarrow \infty$  by the Riemann–Lebesgue lemma; cf. Lemma 3 in [12], Section XV.4. Therefore

$$\Delta := \sup_{\{u \in \mathbb{R}: |u| \geq B/10\}} |\xi(u)| < 1.$$

We are ready to bound  $I_{1,a}$  and  $I_{1,b}$ . For  $r > B$ ,  $\theta \in \Theta_a$  and for  $m \leq \lfloor n/4 \rfloor$  we have

$$r \left| \cos(\theta) + \frac{m}{n} \sin(\theta) \right| \geq \frac{B}{4} \geq \frac{B}{10},$$

and therefore  $|\xi|(r(\cos(\theta) + \frac{m}{n} \sin(\theta))) \leq \Delta$ . Since  $|\xi| \leq 1$ , coming back to (B.3) we can bound  $I_{1,a}$  from above (for  $n \geq 4$ ) by

$$\begin{aligned} I_{1,a} &\leq \sigma^2 n \Delta^{\lfloor n/4 \rfloor} \int_{\{\theta \in \Theta_a, r > B\}} \left\{ \prod_{m=n-3}^n |\xi| \left( r \left( \cos(\theta) + \frac{m}{n} \sin(\theta) \right) \right) \right\} r dr d\theta \\ (B.4) \quad &\leq \sigma^2 n \Delta^{\lfloor n/4 \rfloor} \int_{\mathbb{R}^2} \prod_{m=n-3}^n |\xi| \left( x + \frac{m}{n} y \right) dx dy. \end{aligned}$$

The bound for  $I_{1,b}$  is analogous: for  $r > B$ ,  $\theta \in \Theta_b$  and for  $m \geq \lfloor (3n)/4 \rfloor$  we have

$$r \left| \cos(\theta) + \frac{m}{n} \sin(\theta) \right| \geq B \left( \frac{3\sqrt{3}}{4} - \frac{1}{2} \right) \geq \frac{B}{10},$$

and therefore  $|\xi|(r(\cos(\theta) + \frac{m}{n} \sin(\theta))) \leq \Delta$ . Since  $|\xi| \leq 1$ , in analogy to (B.4) we can bound  $I_{1,b}$  from above by

$$(B.5) \quad \leq \sigma^2 n \Delta^{\lfloor n/4 \rfloor - 4} \int_{\mathbb{R}^2} \prod_{m=n-3}^n |\xi| \left( x + \frac{m}{n} y \right) dx dy.$$

Combining (B.4) and (B.5), we can finally bound  $I_1 = I_{1,a} + I_{1,b}$  from above by

$$(B.6) \quad I_1 \leq 2\sigma^2 n \Delta^{\lfloor n/4 \rfloor - 4} \int_{\mathbb{R}^2} \prod_{m=n-3}^n |\xi| \left( x + \frac{m}{n} y \right) dx dy.$$

Since  $\Delta < 1$ , if we prove that the integral in the r.h.s. is bounded by  $C \cdot n$  for some positive constant  $C$ , then it follows that  $I_1 \rightarrow 0$  as  $n \rightarrow \infty$  and the proof is completed.

Notice that  $|\xi|^2(\cdot) = \xi(\cdot) \xi^*(\cdot)$  is the characteristic function of the random variable  $X_1 - X_2$ , which has an absolutely continuous law with *bounded density* [this is because the density of  $X_1$ , that is  $\exp(-V(\cdot))$ , is bounded by hypothesis]. Since  $|\xi|^2(\cdot) \geq 0$ , it follows from the corollary to Theorem 3 in [12], Section XV.3, that

$|\xi|^2(\cdot)$  is integrable over the whole real line, that is,  $\|\xi\|_2^2 := \int_{\mathbb{R}} |\xi|^2(x) dx < \infty$ . By Young's inequality, we can bound the integral in the r.h.s. of (B.6) by

$$\begin{aligned} & \int_{\mathbb{R}^2} \prod_{m=n-3}^n |\xi| \left( x + \frac{m}{n} y \right) dx dy \\ & \leq \frac{1}{2} \int_{\mathbb{R}^2} |\xi|^2(x+y) \cdot |\xi|^2 \left( x + \frac{n-1}{n} y \right) dx dy \\ & \quad + \frac{1}{2} \int_{\mathbb{R}^2} |\xi|^2 \left( x + \frac{n-2}{n} y \right) \cdot |\xi|^2 \left( x + \frac{n-3}{n} y \right) dx dy. \end{aligned}$$

However, by a simple change of variables it is easy to see that both the integrals in the r.h.s. equal  $n \cdot (\|\xi\|_2^2)^2$ , and the proof is completed.

## APPENDIX C: ENTROPIC REPULSION

We are going to prove Proposition 1.5. Notice that the lower bound in (1.23) is easy:  $\mathbf{P}(\Omega_N^+) \geq \mathbf{P}(Y_1 \geq 0, \dots, Y_N \geq 0) \sim (\text{const.})/\sqrt{N}$ , where the last asymptotic behavior is a classical result of fluctuation theory for random walks with zero mean and finite variance; cf. [12]. Moreover the upper bound in (1.24) follows immediately from the upper bound in (1.23) and Lemma 2.4 [recall the definition (2.11)]. Therefore it remains to prove the lower bound in (1.24), or equivalently (2.13), and the upper bound in (1.23).

**C.1. Proof of (2.13).** We want to get a polynomial lower bound for  $w_{0,0}(N)$  as  $N \rightarrow \infty$ . The difficulty comes from the fact that the process  $\{Z_n\}_n$  is conditioned to come back to zero and therefore the comparison with the process  $\{Y_n\}_n$  is not straightforward.

For simplicity we limit ourselves to the odd case  $N = 2n + 1$  with  $n \in \mathbb{N}$ . Recalling the definition  $\Omega_k^+ := \{Z_1 \geq 0, \dots, Z_k \geq 0\}$ , by Lemma 2.1 we can write

$$\begin{aligned} w_{0,0}(2n+1) &= \mathbf{P}^{(0,0)}(\Omega_{2n-1}^+ | Z_{2n} = 0, Z_{2n+1} = 0) \\ &= \int_{(\mathbb{R}^+)^{2n-1}} e^{-\sum_{k=0}^{2n} V(\Delta\varphi_k)} \prod_{k=1}^{2n-1} d\varphi_k, \end{aligned}$$

where we fix  $\varphi_{-1} = \varphi_0 = \varphi_{2n} = \varphi_{2n+1} = 0$ . The first step is to restrict the integration on the set  $\mathcal{C}_n(\varepsilon) := (\mathbb{R}^+)^{2n-1} \cap \{|\varphi_n - \varphi_{n-1}| \leq \varepsilon, |\varphi_n - \varphi_{n+1}| \leq \varepsilon\}$ , on which  $|\Delta\varphi_n| \leq 2\varepsilon$ . Since  $V(\cdot)$  is continuous and  $V(0) < \infty$ , we can choose  $\varepsilon$  sufficiently small such that  $V(x) \leq V(0) + 1$  for all  $|x| \leq 2\varepsilon$ , so that in particular  $V(\Delta\varphi_n) \leq V(0) + 1$  on the event  $\mathcal{C}_n(\varepsilon)$ . This observation yields the lower bound

$$w_{0,0}(2n+1) \geq e^{-(V(0)+1)} \int_{\mathcal{C}_n(\varepsilon)} e^{-\sum_{k=0}^{n-1} V(\Delta\varphi_k)} \cdot e^{-\sum_{n+1}^{2n} V(\Delta\varphi_k)} \prod_{k=1}^{2n-1} d\varphi_k.$$

Setting  $\mathcal{C}'_n(\varepsilon) := (\mathbb{R}^+)^{n-1} \cap \{|\varphi_n - \varphi_{n-1}| \leq \varepsilon\}$ , the symmetry  $k \rightarrow 2n - k$  gives

$$w_{0,0}(2n+1) \geq e^{-(V(0)+1)} \int_0^\infty d\varphi_n \left( \int_{\mathcal{C}'_n(\varepsilon)} e^{-\sum_{k=0}^{n-1} V(\Delta\varphi_k)} \prod_{k=1}^{n-1} d\varphi_k \right)^2.$$

Restricting the first integration on  $[0, n^{5/2}]$  and applying Jensen's inequality we get

$$\begin{aligned} w_{0,0}(2n+1) &\geq \frac{e^{-(V(0)+1)}}{n^{5/2}} \left( \int_0^{n^{5/2}} d\varphi_n \int_{\mathcal{C}'_n(\varepsilon)} e^{-\sum_{k=0}^{n-1} V(\Delta\varphi_k)} \prod_{k=1}^{n-1} d\varphi_k \right)^2 \\ (C.1) \quad &= \frac{e^{-(V(0)+1)}}{n^{5/2}} \{ \mathbf{P}(\Omega_n^+, Z_n \leq n^{5/2}, |Z_n - Z_{n-1}| \leq \varepsilon) \}^2. \end{aligned}$$

We are thus left with giving a polynomial lower bound for the probability appearing inside the braces. We observe that by definition  $Z_n - Z_{n-1} = Y_n$  and that we have the inclusion  $\Omega_n^+ \supset \{Y_1 \geq 0, \dots, Y_n \geq 0\} =: \Lambda_n^+$ . Therefore

$$\begin{aligned} (C.2) \quad \mathbf{P}(\Omega_n^+, Z_n \leq n^{5/2}, |Z_n - Z_{n-1}| \leq \varepsilon) &\geq \mathbf{P}(\Lambda_n^+, Z_n \leq n^{5/2}, Y_n \leq \varepsilon) \\ &\geq \mathbf{P}(\Lambda_n^+, Y_n \leq \varepsilon) - \mathbf{P}(Z_n > n^{5/2}). \end{aligned}$$

For the second term, Chebyshev's inequality and (2.2) yield

$$(C.3) \quad \mathbf{P}(Z_n > n^{5/2}) \leq \frac{\mathbf{E}(Z_n^2)}{n^5} \leq \frac{(const.)n^3}{n^5} = \frac{(const.)}{n^2}.$$

For the first term, we are going to use some results from fluctuation theory. We denote by  $\{(T_k, H_k)\}_{k \geq 0}$  the weak ascending ladder process associated to the random walk  $\{Y_k\}_k$ , that is,  $(T_0, H_0) = (0, 0)$  and  $T_{k+1} := \inf\{n > T_k : Y_n \geq Y_{T_k}\}$ ,  $H_k := Y_{T_k}$ . The celebrated duality lemma [12], Chapter XII, gives

$$\mathbf{P}(\Lambda_n^+, Y_n \leq \varepsilon) = \sum_{k=0}^{\infty} \mathbf{P}(T_k = n, Y_n \leq \varepsilon),$$

and applying Alili and Doney's combinatorial identity [3] we get

$$\begin{aligned} \mathbf{P}(\Lambda_n^+, Y_n \leq \varepsilon) &= \sum_{k=1}^{\infty} \frac{k}{n} \mathbf{P}(H_{k-1} \leq Y_n < H_k, Y_n \leq \varepsilon) \\ &= \frac{1}{n} \sum_{k=1}^{\infty} \mathbf{P}(H_{k-1} \leq Y_n, Y_n \leq \varepsilon). \end{aligned}$$

Considering only the  $k = 1$  term in the sum gives

$$\mathbf{P}(\Lambda_n^+, Y_n \leq \varepsilon) \geq \frac{1}{n} \mathbf{P}(Y_n \in [0, \varepsilon]) \sim \frac{1}{n} \cdot \frac{(const.)}{\sqrt{n}} \quad (n \rightarrow \infty).$$

Putting together these bounds with (C.1) and (C.2), we have shown that  $w_{0,0}(2n+1) \geq (const.)/n^{11/2}$ , hence the proof of (2.13) is complete.

**C.2. Proof of the upper bound in 1.23.** Throughout this section  $\delta$  will denote a small positive parameter. More precisely, let us set

$$(C.4) \quad c_+(\delta) := \frac{\log 1/((1/2) + 2\delta)}{\log(3/2 + \delta)}.$$

Since  $c_+(0) = \log 2 / \log \frac{3}{2} > 1$ , by continuity  $c_+(\delta) > 1$  for small  $\delta > 0$ . In the sequel, we are free to fix any positive  $\delta$  such that  $c_+(\delta) > 1$ ; in fact we are going to prove that (2.12) holds true with  $c_+ = c_+(\delta)$ . For convenience, we split the proof into four steps.

*First step.* We introduce the integer-valued sequence  $g_n$  defined for  $n \in \mathbb{N}$  by

$$(C.5) \quad g_n := \lfloor \exp\left\{\left(\frac{3}{2} + \delta\right)^n\right\} \rfloor.$$

We claim that for every  $D > 0$  there exists  $n_0 \in \mathbb{N}$  such that for every  $n \geq n_0$  we have

$$(C.6) \quad \mathbf{P}(Z_{g_{n+1}-g_n} \geq -D(g_n)^{7/4} - (g_{n+1} - g_n)D(g_n)^{3/4}) \leq \frac{1}{2} + \delta.$$

It is easily checked that for large  $n$

$$\frac{1}{2}(g_n)^{3/2+\delta} \leq g_{n+1} - g_n \leq 2(g_n)^{3/2+\delta},$$

which implies that

$$\frac{-D(g_n)^{7/4} - (g_{n+1} - g_n)D(g_n)^{3/4}}{(g_{n+1} - g_n)^{3/2}} \rightarrow 0 \quad (n \rightarrow \infty).$$

We have already observed in Section 2.3 that  $Z_k/(\sigma k^{3/2})$  converges in distribution toward  $\int_0^1 B_s ds$ , where  $\{B_s\}_s$  is a standard Brownian motion, hence

$$\mathbf{P}(Z_{g_{n+1}-g_n} \geq -D(g_n)^{7/4} - (g_{n+1} - g_n)D(g_n)^{3/4}) \xrightarrow{n \rightarrow \infty} P\left(\int_0^1 B_s ds \geq 0\right) = \frac{1}{2},$$

and (C.6) follows.

*Second step.* We claim that we can fix the positive constant  $D$  such that for every  $n \in \mathbb{N}$

$$(C.7) \quad \mathbf{P}(\{Z_{g_n} > D(g_n)^{7/4}\} \cup \{Y_{g_n} > D(g_n)^{3/4}\} | Z_{g_k} \geq 0, \forall k = 1, \dots, n) \leq \delta.$$

By Chebyshev's inequality and (2.2) we can write for every  $m \in \mathbb{N}$

$$\mathbf{P}(Z_m > Dm^{7/4}) \leq \frac{\mathbf{E}(Z_m^2)}{D^2 m^{7/2}} \leq \frac{(\text{const.}) m^3}{D^2 m^{7/2}} = \frac{(\text{const.})}{D^2 \sqrt{m}},$$

and analogously

$$\mathbf{P}(Y_m > Dm^{3/4}) \leq \frac{\mathbf{E}(Y_m^2)}{D^2 m^{3/2}} \leq \frac{(\text{const.}') m}{D^2 m^{3/2}} = \frac{(\text{const.}')}{D^2 \sqrt{m}}.$$

The inclusion bound yields

$$\begin{aligned} \mathbf{P}(Z_{g_k} \geq 0, \forall k = 1, \dots, n) &\geq \mathbf{P}(Y_m \geq 0, \forall m = 1, \dots, g_n) \\ &\geq \frac{(const.'')}{\sqrt{g_n}} \quad (n \rightarrow \infty), \end{aligned}$$

where the last bound is a well-known result of fluctuation theory; cf. [12], Section XII.7. Then

$$\begin{aligned} &\mathbf{P}(\{Z_{g_n} > D(g_n)^{7/4}\} \cup \{Y_{g_n} > D(g_n)^{3/4}\} \mid Z_{g_k} \geq 0, \forall k = 1, \dots, n) \\ &\leq \frac{\mathbf{P}(Z_{g_n} > D(g_n)^{7/4}) + \mathbf{P}(Y_{g_n} > D(g_n)^{3/4})}{\mathbf{P}(Z_{g_k} \geq 0, \forall k = 1, \dots, n)} \leq \frac{(const.) + (const.')}{(const.'')} \frac{1}{D^2}, \end{aligned}$$

and (C.7) follows.

*Third step.* We claim that for every  $n \geq n_0$  (defined in the preceding step) the following relation holds:

$$(C.8) \quad \mathbf{P}(Z_{g_k} \geq 0, \forall k = 1, \dots, n+1) \leq \left(\frac{1}{2} + 2\delta\right) \cdot \mathbf{P}(Z_{g_k} \geq 0, \forall k = 1, \dots, n).$$

Conditioning on the  $\sigma$ -algebra  $\sigma(Z_1, \dots, Z_{g_n}) = \sigma(X_1, \dots, X_{g_n})$  we have

$$\begin{aligned} (C.9) \quad &\mathbf{P}(Z_{g_k} \geq 0, \forall k = 1, \dots, n+1) \\ &= \mathbf{E}(\mathbf{P}(Z_{g_{n+1}} \geq 0 \mid Y_{g_n}, Z_{g_n}) \cdot \mathbf{1}_{\{Z_{g_k} \geq 0, \forall k = 1, \dots, n\}}). \end{aligned}$$

Let us introduce the event  $\mathcal{A} := \{Z_{g_n} \leq D(g_n)^{7/4}, Y_{g_n} \leq D(g_n)^{3/4}\}$ , the constant  $D$  being fixed in the preceding step. By (2.3) we can write

$$\mathbf{P}(Z_{g_{n+1}} \geq 0 \mid Y_{g_n}, Z_{g_n}) = \mathbf{P}(Z_{g_{n+1}-g_n} \geq -z - (g_{n+1} - g_n)y) \mid_{y=Y_{g_n}, z=Z_{g_n}},$$

hence on the event  $\mathcal{A}$  we have

$$\mathbf{P}(Z_{g_{n+1}} \geq 0 \mid Y_{g_n}, Z_{g_n}) \leq \mathbf{P}(Z_{g_{n+1}} \geq -D(g_n)^{7/4} - (g_{n+1} - g_n)D(g_n)^{3/4}) \leq \frac{1}{2} + \delta,$$

having applied (C.6). Coming back to (C.9), we get

$$\begin{aligned} \mathbf{P}(Z_{g_k} \geq 0, \forall k = 1, \dots, n+1) &\leq \left(\frac{1}{2} + \delta\right) \mathbf{P}(Z_{g_k} \geq 0, \forall k = 1, \dots, n) \\ &\quad + \mathbf{P}(\mathcal{A}^c, Z_{g_k} \geq 0, \forall k = 1, \dots, n), \end{aligned}$$

and thanks to (C.7) we can bound the second term in the r.h.s. by

$$\mathbf{P}(\mathcal{A}^c, Z_{g_k} \geq 0, \forall k = 1, \dots, n) \leq \delta \cdot \mathbf{P}(Z_{g_k} \geq 0, \forall k = 1, \dots, n),$$

hence (C.8) follows.

*Fourth step.* We are finally ready to complete the proof of (2.12). Assume first that  $N = g_n$  for some  $n \in \mathbb{N}$ . Using the inclusion bound and iterating (C.8) we get

$$\mathbf{P}(\Omega_N^+) \leq \mathbf{P}(Z_{g_k} \geq 0, \forall k = 1, \dots, n) \leq C\left(\frac{1}{2} + 2\delta\right)^n,$$

where  $C > 0$  is an absolute constant. By (C.5) we have

$$g_n = N \implies \exp\left\{\left(\frac{3}{2} + \delta\right)^n\right\} \geq N \implies n \geq \frac{\log \log N}{\log(3/2 + \delta)},$$

hence

$$\mathbf{P}(\Omega_N^+) \leq C\left(\frac{1}{2} + 2\delta\right)^{\log \log N / \log(3/2 + \delta)} = \frac{C}{(\log N)^{c_+(\delta)}},$$

where we recall that  $c_+(\delta)$  has been defined in (C.4). Since  $\delta > 0$  has been chosen sufficiently small such that  $c_+(\delta) > 1$ , (2.12) is proven. In the general case, let

$$n^* := \max\{n \in \mathbb{N} : g_n \leq N\}.$$

Since  $g_{k+1} \approx (g_k)^{3/2+\delta}$ , it is easily checked that  $g_{n^*} \geq \sqrt{N}$  for large  $N$  (provided  $\delta < \frac{1}{2}$ , which is no harm). Therefore we can repeat the above arguments just replacing  $n$  by  $n^*$  and  $N$  by  $\sqrt{N}$ , thus getting

$$\mathbf{P}(\Omega_N^+) \leq \frac{C}{(\log \sqrt{N})^{c_+(\delta)}} = \frac{2^{c_+(\delta)} C}{(\log N)^{c_+(\delta)}},$$

and (2.12) is proven in full generality.

**Acknowledgments.** We are very grateful to Yvan Velenik and Ostap Hrynniv for suggesting the present problem to us, and to Giambattista Giacomin and Fabio Lucio Toninelli for enlightening discussions on the content of Section 1.6.

## REFERENCES

- [1] ALEXANDER, K. S. (2008). The effect of disorder on polymer depinning transitions. *Comm. Math. Phys.* **279** 117–146. MR2377630
- [2] ALEXANDER, K. S. and SIDORAVICIUS, V. (2006). Pinning of polymers and interfaces by random potentials. *Ann. Appl. Probab.* **16** 636–669. MR2244428
- [3] ALILI, L. and DONEY, R. A. (1999). Wiener-Hopf factorization revisited and some applications. *Stochastics Stochastics Rep.* **66** 87–102. MR1687803
- [4] ASMUSSEN, S. (2003). *Applied Probability and Queues*, 2nd ed. *Applications of Mathematics (New York)* **51**. Springer, New York. MR1978607
- [5] BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. (1989). *Regular Variation. Encyclopedia of Mathematics and its Applications* **27**. Cambridge Univ. Press, Cambridge. MR1015093
- [6] BUNDSCUH, R., LÄSSIG, M. and LIPOWSKY, R. (2000). Semiflexible polymers with attractive Interactions. *Eur. Phys. J. E* **3** 295–306.
- [7] CARAVENNA, F. and DEUSCHEL, J.-D. (2008) Scaling limits of (1+1)-dimensional pinning models with Laplacian interaction. *Ann. Probab.* To appear..

- [8] CARAVENNA, F., GIACOMIN, G. and ZAMBOTTI, L. (2007). A renewal theory approach to periodic copolymers with adsorption. *Ann. Appl. Probab.* **17** 1362–1398. MR2344310
- [9] CARAVENNA, F., GIACOMIN, G. and ZAMBOTTI, L. (2006). Sharp asymptotic behavior for wetting models in  $(1 + 1)$ -dimension. *Electron. J. Probab.* **11** 345–362 (electronic). MR2217821
- [10] DEUSCHEL, J.-D., GIACOMIN, G. and ZAMBOTTI, L. (2005). Scaling limits of equilibrium wetting models in  $(1 + 1)$ -dimension. *Probab. Theory Related Fields* **132** 471–500. MR2198199
- [11] FISHER, M. E. (1984). Walks, walls, wetting, and melting. *J. Statist. Phys.* **34** 667–729. MR751710
- [12] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*. II, 2nd ed. Wiley, New York. MR0270403
- [13] GIACOMIN, G. (2007). *Random Polymer Models*. Imperial College Press, London. MR2380992
- [14] GIACOMIN, G. and TONINELLI, F. L. (2006). Smoothing of depinning transitions for directed polymers with quenched disorder. *Phys. Rev. Lett.* **96** 070602.
- [15] GIACOMIN, G. and TONINELLI, F. L. (2006). Smoothing effect of quenched disorder on polymer depinning transitions. *Comm. Math. Phys.* **266** 1–16. MR2231963
- [16] GIACOMIN, G. and TONINELLI, F. L. (2007). On the irrelevant disorder regime of pinning models. Preprint. arXiv.org:0707.3340v1 [math.PR].
- [17] HRYNIV, O. and VELENIK, Y. (2008). Some rigorous results about semiflexible polymers. In preparation.
- [18] ISOZAKI, Y. and YOSHIDA, N. (2001). Weakly pinned random walk on the wall: Pathwise descriptions of the phase transition. *Stochastic Process. Appl.* **96** 261–284. MR1865758
- [19] KAFRI, Y., MUKAMEL, D. and PELITI, L. (2000). Why is the DNA denaturation transition first order? *Phys. Rev. Lett.* **85** 4988–4991.
- [20] KATO, T. (1976). *Perturbation Theory for Linear Operators*, 2nd ed. Springer, Berlin. MR0407617
- [21] KIERFELD, J. and LIPOWSKY, R. (2003). Unbundling and desorption of semiflexible polymers. *Europhys. Lett.* **62** 285–291.
- [22] KURT, N. (2007). Entropic repulsion for a class of Gaussian interface models in high dimensions. *Stochastic Process. Appl.* **117** 23–34. MR2287101
- [23] LACHAL, A. (1991). Sur le premier instant de passage de l'intégrale du mouvement brownien. *Ann. Inst. H. Poincaré Probab. Statist.* **27** 385–405. MR1131839
- [24] LI, W. V. and SHAO, Q.-M. (2004). Lower tail probabilities for Gaussian processes. *Ann. Probab.* **32** 216–242. MR2040781
- [25] MCKEAN, H. P. JR. (1963). A winding problem for a resonator driven by a white noise. *J. Math. Kyoto Univ.* **2** 227–235. MR0156389
- [26] NUMMELIN, E. (1984). *General Irreducible Markov Chains and Nonnegative Operators*. Cambridge Tracts in Mathematics **83**. Cambridge Univ. Press, Cambridge. MR776608
- [27] RICHARD, C. and GUTTMANN, A. J. (2004). Poland-Scheraga models and the DNA denaturation transition. *J. Statist. Phys.* **115** 925–947. MR2054168
- [28] SAKAGAWA, H. (2003). Entropic repulsion for a Gaussian lattice field with certain finite range interaction. *J. Math. Phys.* **44** 2939–2951. MR1982781
- [29] SINAI, Y. G. (1992). Distribution of some functionals of the integral of a random walk. *Teoret. Mat. Fiz.* **90** 323–353. MR1182301
- [30] TONINELLI, F. L. (2008). A replica-coupling approach to disordered pinning models. *Comm. Math. Phys.* **280** 389–401. MR2395475
- [31] TONINELLI, F. L. (2008). Disordered pinning models and copolymers: Beyond annealed bounds. *Ann. Appl. Probab.* **18** 1569–1587.

- [32] VELENIK, Y. (2006). Localization and delocalization of random interfaces. *Probab. Surv.* **3** 112–169 (electronic). MR2216964
- [33] ZERNER, M. (1987). Quelques propriétés spectrales des opérateurs positifs. *J. Funct. Anal.* **72** 381–417. MR886819

DIPARTIMENTO DI MATEMATICA PURA E APPLICATA  
UNIVERSITÀ DEGLI STUDI DI PADOVA  
VIA TRIESTE 63  
35121 PADOVA  
ITALY  
E-MAIL: francesco.caravenna@math.unipd.it

TECHNISCHE UNIVERSITÄT BERLIN  
INSTITUT FÜR MATHEMATIK  
STRASSE DES 17. JUNI 136  
10623 BERLIN  
GERMANY  
E-MAIL: deuschel@math.tu-berlin.de