Probab. Theory Relat. Fields 96, 521–543 (1993)

Probability Theory and Related Fields

The drift of a one-dimensional self-avoiding random walk

Wolfgang König

Institut für Angewandte Mathematik, Universität Zürich, Rämistrasse 74, CH-8001 Zürich, Switzerland

Received December 18, 1992; in revised form April 21, 1993

Summary. We prove that a self-avoiding random walk on the integers with bounded increments grows linearly. We characterize its drift in terms of the Frobenius eigenvalue of a certain one parameter family of primitive matrices. As an important tool, we express the local times as a two-block functional of a certain Markov chain, which is of independent interest.

Mathematics Subject Classification (1980): 60K35, 58E30, 60F10, 60J15

1 Introduction and statement of results

Consider a random walk $(S_n)_{n\in\mathbb{N}_0}$ on $\mathbb{Z}$, starting at 0, with i.i.d. and bounded increments. Let the steps be given by

$$S_{n+1}-S_n=k\in\{\pm 1, \ldots, \pm r\} \quad \text{with probability} \quad p_{|k|}\, e^{hk}/Z_h$$

for $n\in\mathbb{N}_0$, where $Z_h=\sum_{|k|=1}^{r} p_{|k|}\, e^{hk}$ is a normalization constant, $r\in\mathbb{N}$ an increment bound, $p_1, \ldots, p_r$ are positive numbers and $h\in[0, \infty)$ is a drift parameter. The associated path measure and expectation are denoted by P_h and E_h, respectively. We assume that $Z_0=1$. Under P_h the steps have mean $\frac{d}{dh}\log Z_h$, which is positive if and only if h is.

Let

$$T:=\inf\{n\in\mathbb{N}\colon S_k=S_n \quad \text{for some } k\in\{0, 1, \ldots, n-1\}\}$$

be the first time of a self-intersection of the path. The purpose of this paper is to prove that, conditioned on the event $\{T>n\}$, the path approaches a straight line for large n, and to characterize its slope. (Of course, this is trivial in the nearest-neighbour case $r=1$.) Using linear scaling, we formulate our first main result in the context of the function space $C[0, 1]$, endowed with the Borel

σ-field generated by the supremum norm. For $n \in \mathbb{N}$, let $S^{(n)}: [0, 1] \to [-r, r]$ be the random continuous function defined by

$$S^{(n)}\left(\frac{k}{n}\right) = \frac{S_k}{n} \qquad (k = 0, 1, \ldots, n)$$

and linear interpolation. Define a probability measure $\mathbb{P}_n^h$ on $C[0, 1]$ by

$$\mathbb{P}_n^h(A) = P_h(S^{(n)} \in A \mid T > n) \qquad \text{for Borel subsets } A \text{ of } C[0, 1].$$

Put $t_\theta(x) := \theta x$ for $\theta \in \mathbb{R}$, $x \in [0, 1]$, and let δ_g be the Dirac measure on $g \in C[0, 1]$. We state our law of large numbers for $(\mathbb{P}_n^h)_{n \in \mathbb{N}}$:

Theorem 1.1 *There exists a real-analytic and strictly increasing function Θ from $[0, \infty)$ into $[1, r]$ such that*

$$\mathbb{P}_n^h \xrightarrow{n \to \infty} \begin{cases} \delta_{t_{\Theta(h)}} & \text{if } h > 0, \\ \frac{1}{2}(\delta_{t_{\Theta(0)}} + \delta_{t_{-\Theta(0)}}) & \text{if } h = 0. \end{cases}$$

If $r \geqq 2$, then $\Theta(h) \in (1, r)$ for every $h \in [0, \infty)$, and $\lim_{h \uparrow \infty} \Theta(h) = r$.

We call this number $\Theta(h)$ the *effective drift* of the self-avoiding walk. Theorem 1.1 implies the conjecture (7.6) in [1]. Clearly, $\Theta(h) \leqq E_h(S_1 \mid S_1 > 0)$, but a proof for the rather natural conjecture that $\Theta(h) \geqq E_h(S_1)$ (which is obvious from Theorem 1.1 only for small h) seems to be very hard to derive.

The main work in the proof consists of analyzing the functions J_h, $\tilde{J}_h$: $[0, r] \to [-\infty, 0]$ defined by

$$J_h(\theta) = \lim_{n \to \infty} \frac{1}{n} \log P_h(T > n, S_n = \lfloor \theta n \rfloor), \tag{1.2}$$

$$\tilde{J}_h(\theta) = \lim_{n \to \infty} \frac{1}{n} \log P_h(T > n, S_n = \lfloor \theta n \rfloor, \\ S_1, S_2, \ldots, S_{n-1} \in [1, \lfloor \theta n \rfloor - 1]). \tag{1.3}$$

(Here and in the sequel $\lfloor x \rfloor := \max\{k \in \mathbb{Z}: k \leqq x\}$.) The existence of these limits will be shown in Proposition 2.9 and 4.1, respectively. Note the obvious facts that $\tilde{J}_h \equiv -\infty$ on $[0, 1)$ and that $J_h(\theta) > -\infty$ for every $\theta \in (0, r]$ if $r \geqq 2$. For proving Theorem 1.1, we need that J_h has a strict maximum at $\Theta(h)$. This is established in Sect. 4. In order to study the function J_h, we need an analysis of the function $\tilde{J}_h$, which is given in Sects. 2 and 3. In fact, $\Theta(h)$ is introduced as the strict maximum point of $\tilde{J}_h$ and we show in Sect. 4 that $J_h(\Theta(h)) = \tilde{J}_h(\Theta(h))$.

In Sect. 2 we derive a variational formula for $\tilde{J}_h(\theta)$ in the case $\theta \in (1, r)$, $h > 0$. The main tool is an expression of the local times of $(S_n)_{n \in \mathbb{N}_0}$ as a two-block functional of a certain Markov chain. We analyze this variational formula in Sect. 3 and obtain a characterization of the function $\tilde{J}_h$ and its strict maximum point $\Theta(h)$ in terms of the Frobenius eigenvalue $\lambda_h(b)$ of a certain primitive matrix A_b (with $b \in \mathbb{R}$), which is our second main result:

Theorem 1.4 *For $h>0$ there exists a real-analytic, strictly incresing, and strictly* log-*convex function λ_h:* $\mathbb{R}\to(0,\infty)$ *such that, for* $\theta\in(1,r)$,

$$\tilde{J}_h(\theta)=\theta\log\lambda_h(b_h(\theta))-b_h(\theta),\qquad (1.5)$$
$$\tilde{J}_h'(\theta)=\log\lambda_h(b_h(\theta)),$$
$$\tilde{J}_h''(\theta)=\frac{b_h'(\theta)}{\theta}<0,$$

where b_h: $(1,r)\to\mathbb{R}$ denotes the inverse function of λ_h/λ_h'. In particular,

$$\Theta(h)=\frac{1}{\lambda_h'(\lambda_h^{-1}(1))}\quad \textit{for}\ h>0. \qquad (1.6)$$

We emphasize that the Markov chain introduced at the beginning of Sect. 2 is suited for dealing with any functional of the local times of a random walk on $\mathbb{Z}$ having bounded and nonzero increments and positive drift. We will use this chain in a forthcoming paper to prove the following. Let

$$X_n:=\#\{(i,j)\in\{0,1,\dots,n\}^2\,|\,i<j\quad\text{and}\quad S_i=S_j\}$$

be the number of self-intersections up to time n. If we replace the density $1_{\{T>n\}}/P_h(T>n)$ of $\mathbb{P}_n^h$ by $(1-\alpha)^{X_n}/E_h((1-\alpha)^{X_n})$ for a fixed $\alpha\in(0,1)$, then we obtain a probability measure $\mathbb{P}_n^{h,\alpha}$ which suppresses, but not neglects paths having self-intersections before time n. This polymer measure is a model for a self-repellent random walk with repulsion strength α. In [7] we will prove a result for $(\mathbb{P}_n^{h,\alpha})_n$ which is analogous to Theorems 1.1 and 1.4.

We close this section by mentioning some mathematical works on self-repellent or self-avoiding random walks on the integers (An overview on polymers from a more physical and chemical point of view is given in [5]). The analogous self-repellent model for Brownian motion is considered in [11] and [8], and the long-time behaviour of the endpoint of the path is shown to be linear. In [1] the sequence of successive times of self-intersections of a walk with bounded increments is shown to possess a limit law as the maximal step size tends to infinity. Using variational techniques, upper and lower bounds for the long-time behaviour of a self-repellent walk are derived in [2]. Here no restriction is made concerning the size of the steps, but the method of this paper only works for small repulsion strength. The nearest-neighbour case is investigated in [6] using an approach which is analogous to the one in the present paper. In that work a quite explicit representation of the effective drift of the self-repellent walk is derived.

2 A variational formula for $\tilde{J}_h$

Let $h>0$, $r\geqq 2$, and $\theta\in(1,r)$ be fixed during this section. We regard P_h as a probability measure on

$$\{(\omega_n)_n\in\mathbb{Z}^{\mathbb{N}_0}\colon \omega_0=0,\ |\omega_n-\omega_{n-1}|\in\{1,\dots,r\}\ \text{for all}\ n\in\mathbb{N}\ \text{and}\ \lim_{n\to\infty}\omega_n=+\infty\}$$

with the relative σ-field induced by the sequence $(S_n)_n$.

2.1 Local times

The random walk is transient, so the local time

$$l(x) := \#\{k \in \mathbb{N}_0 : S_k = x\} \qquad (x \in \mathbb{Z})$$

is finite. In order to study self-avoiding walks, we must obtain some information about the distribution of the sequence of local times. Our first aim is to express the local times $l(x)$ for $x \in \mathbb{N}$ as a two-block functional of a Markov chain. The idea is to count the excursions above the line between x and $x+1$ with given positions at the beginning and at the end of the excursion and to register the position after the last jump over this line. Let

$$\begin{aligned}
&E := \mathbb{N}_0^{r \times r} \times \{1, \ldots, r\}, \\
&\eta_{j,k}(x) := \#\{(m,n) \in \mathbb{N}^2 : m < n,\ S_{m-1} \leqq x,\ S_m = x+j, \\
&\qquad\qquad S_{m+1}, \ldots, S_{n-1} > x,\ S_n = x+1-k\}, \\
&\tau(x) := \max\{n \in \mathbb{N} : S_{n-1} \leqq x < S_n\}, \\
&q(x) := S_{\tau(x)} - x,
\end{aligned} \tag{2.1}$$

for $x \in \mathbb{N}_0$ and $j, k \in \{1, \ldots, r\}$. Then we have

Lemma 2.2 *The process* $(\Lambda(x))_{x \in \mathbb{N}_0} := ((\eta_{j,k}(x))_{j,k=1,\ldots,r}, q(x))_{x \in \mathbb{N}_0}$ *is a homogeneous Markov chain on E.*

Proof. Fix $x \in \mathbb{N}$ during the proof. The two intuitive ideas behind the Markov property are: (1) every excursion above $x+1$ occurs during an excursion above x or between the times $\tau(x)$ and $\tau(x+1)$, and (2) the excursions above x with fixed starting and ending points and the path $(S_{\tau(x)}, \ldots, S_{\tau(x+1)})$ are conditionally independent of the parts between them given their number and $S_{\tau(x)}$.

Define the beginning and ending times of the excursions above x with parameters $j, k \in \{1, \ldots, r\}$ by $\tau_{j,k}^{(0)} := 0$ and, for $i \in \mathbb{N}$,

$$\begin{aligned}
\tau_{j,k}^{(2i)} := \inf\{n > \tau_{j,k}^{(2i-2)} &: \text{there exists } m \in \{\tau_{j,k}^{(2i-2)}, \ldots, n\} \text{ such that} \\
&S_m \leqq x,\ S_{m+1} = x+j,\ S_{m+2}, \ldots, S_{n-1} > x,\ S_n = x+1-k\}
\end{aligned}$$

(by convention, $\inf \emptyset = \infty$) and

$$\tau_{j,k}^{(2i-1)} := \begin{cases} \max\{m < \tau_{j,k}^{(2i)}(x) : S_{m-1} \leqq x,\ S_m = x+j\} & \text{if } \tau_{j,k}^{(2i)} < \infty, \\ \infty & \text{else.} \end{cases}$$

Then $(\tau_{j,k}^{(i)})_{i \in \mathbb{N}_0}$ is increasing, $\tau_{j,k}^{(2i)}$ is a stopping time, but $\tau_{j,k}^{(2i-1)}$ not (if $r \geqq 2$). Furthermore, define for $i \in \mathbb{N}$ and $j, k \in \{1, \ldots, r\}$,

$$Y_{j,k}^{(i)} := \begin{cases} (S_{\tau_{j,k}^{(i-1)}}, \ldots, S_{\tau_{j,k}^{(i)}}) & \text{if } \tau_{j,k}^{(i)} < \infty, \\ (S_{\tau_{j,k}^{(i-1)}}, \ldots, S_{\tau(x)}) & \text{if } \tau_{j,k}^{(i-1)} < \infty = \tau_{j,k}^{(i)}, \\ (S_{\tau(x)}, \ldots, S_{\tau(x+1)}) & \text{else.} \end{cases}$$

So $Y_{j,k}^{(2)}, Y_{j,k}^{(4)}, \ldots, Y_{j,k}^{(2\eta_{j,k}(x))}$ are the excursions above x with parameters j, k, starting in $x+j$ and ending in $x+1-k$. For every $j, k \in \{1, \ldots, r\}$, the variables $\Lambda(0), \ldots, \Lambda(x-1)$ are $\sigma(Y_{j,k}^{(2i-1)} : i \in \mathbb{N})$-measurable and $q(x+1)$ is

$\sigma(Y_{j,k}^{(2\eta_{j,k}(x)+2)})$-measurable. For $y \in \mathbb{N}$, $j, k \in \{1, \ldots, r\}$ and $s, t \in \mathbb{N}$ satisfying $s \leqq t$, let

$$N_{\text{exc}}(y, j, k, s, t) := \#\{(m, n) \in \{s, \ldots, t\}^2 \mid m < n \text{ and } S_{m-1} \leqq y, \\ S_m = y + j, S_{m+1} > y, \ldots, S_{n-1} > y, S_n = y + 1 - k\}$$

be the number of excursions above y with parameters j, k during the time interval $\{s, \ldots, t\}$, then we have, for $\tilde{j}, \tilde{k} \in \{1, \ldots, r\}$,

$$\eta_{\tilde{j},\tilde{k}}(x+1) = \sum_{j,k=1}^{r} \sum_{i=1}^{\eta_{j,k}(x)} N_{\text{exc}}(x+1, \tilde{j}, \tilde{k}, \tau_{j,k}^{(2i-1)}, \tau_{j,k}^{(2i)}) \\ + N_{\text{exc}}(x+1, \tilde{j}, \tilde{k}, \tau(x), \tau(x+1)).$$

We will prove that $\bigvee_{j,k=1}^{r} \sigma(Y_{j,k}^{(2i)}: i \in \mathbb{N})$ and $\bigcap_{\tilde{j},\tilde{k}=1}^{r} \sigma(Y_{\tilde{j},\tilde{k}}^{(2i-1)}: i \in \mathbb{N})$ are independent under $P_h(\cdot \mid \Lambda(x))$. (The Markov property of $(\Lambda(y))_y$ follows from this independence since $N_{\text{exc}}(x+1, \tilde{j}, \tilde{k}, \tau_{j,k}^{(2i-1)}, \tau_{j,k}^{(2i)})$ is $\sigma(Y_{j,k}^{(2i)})$-measurable for every $j, k, \tilde{j}, \tilde{k} \in \{1, \ldots, r\}$ and $i \in \{1, \ldots, \eta_{j,k}(x)\}$ (because of $S_{\tau_{j,k}^{(2i-1)}-1} \leqq x$) and since $\bigcap_{\tilde{j},\tilde{k}=1}^{r} \sigma(Y_{\tilde{j},\tilde{k}}^{(2i-1)}: i \in \mathbb{N})$ contains $\sigma(\Lambda(0), \ldots, \Lambda(x-1))$ (see above).) To show this independence, it suffices to show that $(Y_{j,k}^{(2)}, Y_{j,k}^{(4)}, \ldots, Y_{j,k}^{(2\eta_{j,k}(x)+2)})$ and $(Y_{j,k}^{(1)}, Y_{j,k}^{(3)}, \ldots, Y_{j,k}^{(2\eta_{j,k}(x)+1)})$ are independent under $P_h(\cdot \mid \Lambda(x))$ for every $j, k \in \{1, \ldots, r\}$, since, for $(\tilde{j}, \tilde{k}) \neq (j, k)$ and $i \in \mathbb{N}$, the variable $Y_{\tilde{j},\tilde{k}}^{(2i)}$ is $\sigma(Y_{j,k}^{(2\hat{i}-1)}: \hat{i} \in \mathbb{N})$-measurable. For that, it is enough to show that, for $\eta \in \mathbb{N}_0$ and $q \in \{1, \ldots, r\}$, the variables $Y_{j,k}^{(1)}, Y_{j,k}^{(2)}, \ldots, Y_{j,k}^{(2\eta+2)}$ are independent under

$$P^{\eta,q} := P_h(\cdot \mid \eta_{j,k}(x) = \eta, q(x) = q).$$

For $i = 1, \ldots, 2\eta + 2$ let $\alpha_i \in \operatorname{supp}(P^{\eta,q} Y_{j,k}^{(i)-1})$ (this set is countable) be a path in $\mathbb{Z}$ with length $n_i \in \mathbb{N}_0$. Write $\alpha - y$ for the path α which is vertically shifted by $y \in \mathbb{Z}$ and $\alpha\beta$ for the path which arises if the paths α and β are put together. Then we have for $\eta \geqq 1$, by using the strong Markov property at the times $\tau_{j,k}^{(2)}, \tau_{j,k}^{(4)}, \ldots, \tau_{j,k}^{(2\eta)}$:

$$P_h(\eta_{j,k}(x) = \eta, q(x) = q) \\ = P_h(\eta_{j,k}(x) \geqq 1)\, P_h(\eta_{j,k}(k-1) \geqq 1)^{\eta-1}\, P_h(\eta_{j,k}(k-1) = 0, q(k-1) = q)$$

and, for $i = 2, \ldots, \eta$,

$$P_h(Y_{j,k}^{(2i-1)} = \alpha_{2i-1}, Y_{j,k}^{(2i)} = \alpha_{2i}, \eta_{j,k}(x) = \eta, q(x) = q) \\ = P_h(\eta_{j,k}(x) \geqq i-1)\, P_h((S_0, \ldots, S_{n_{2i-1}+n_{2i}}) = \alpha_{2i-1}\alpha_{2i} - (x+1-k)) \\ \cdot P_h(\eta_{j,k}(k-1) = \eta - i, q(k-1) = q)$$

and analogous formulae for $i \in \{1, \eta+1\}$. So one can easily derive that the pairs $(Y_{j,k}^{(2i-1)}, Y_{j,k}^{(2i)})$ (for $i = 1, \ldots, \eta+1$) are independent under $P^{\eta,q}$.

The last step in this proof is to show the independence of $Y_{j,k}^{(2i-1)}$ and $Y_{j,k}^{(2i)}$ under $P^{\eta,q}$ for every $i \in \{1, \ldots, \eta+1\}$. We will do this for $i = 1$ (the other cases are similar).

Write $c_{j,k}=P_h(\exists n\in\mathbb{N}: S_1, S_2, \ldots, S_{n-1}>-j,\ S_n=-j+1-k)\in(0,1)$ and $d_{j,k}=P_h(\eta_{j,k}(k-1)=\eta-1,\ q(k-1)=q)$, then we derive from the Markov property at time n_1 that

$$P_h(Y_{j,k}^{(1)}=\alpha_1,\ \eta_{j,k}(x)=\eta,\ q(x)=q)=P_h((S_0,\ldots,S_{n_1})=\alpha_1)\,c_{j,k}\,d_{j,k}$$

and

$$\begin{aligned}&P_h(Y_{j,k}^{(2)}=\alpha_2,\ \eta_{j,k}(x)=\eta,\ q(x)=q)\\&=\sum_{\tilde\alpha_1\in\operatorname{supp}(P^{\eta,q}Y_{j,k}^{(1)-1})} P_h((S_0,\ldots,S_{\tilde n_1})=\tilde\alpha_1)\\&\quad\cdot P((S_0,\ldots,S_{n_2})=\alpha_2-(x+j))\,d_{j,k}\\&=c_{j,k}^{-1}\,P((S_0,\ldots,S_{n_2})=\alpha_2-(x+j))\,d_{j,k},\end{aligned}$$

since $P_h(Y_{j,k}^{(1)}=\tilde\alpha_1)=P_h((S_0,\ldots,S_{\tilde n_1})=\tilde\alpha_1)\,c_{j,k}$ for $\tilde\alpha_1\in\operatorname{supp}(P^{\eta,q}\,Y_{j,k}^{(1)-1})$. The independence of $Y_{j,k}^{(1)}$ and $Y_{j,k}^{(2)}$ under $P^{\eta,q}$ easily follows from this, so the Markov property of $(A(x))_{x\in\mathbb{N}_0}$ is proved. The homogeneity of this chain follows from the assumption that every step $S_{n+1}-S_n$ does nor depend on the time $n\in\mathbb{N}_0$ neither on the site $S_n\in\mathbb{Z}$. □

We denote the transition kernel of $(A(x))_{x\in\mathbb{N}_0}$ by Q_h: $E\times E\to[0,1]$. Define g: $E\times E\to\mathbb{N}_0$ by

(2.3) $$g(A,\tilde A)=\sum_{k=1}^{r}\eta_{1,k}+1_{q=1}+\sum_{j=1}^{r}\tilde\eta_{j,1}$$

for $A=((\eta_{j,k})_{j,k},q)$ and $\tilde A=((\tilde\eta_{j,k})_{j,k},\tilde q)\in E$. Then we have, for $x\in\mathbb{N}$,

$$\begin{aligned}l(x)&=\#\{\text{jumps from below to }x\}+\#\{\text{jumps from above to }x\}\\&=g(A(x-1),A(x)),\end{aligned}$$

so our first aim is attained.

Note that, for $A=((\eta_{j,k})_{j,k},q)$ and $\tilde A=((\tilde\eta_{j,k})_{j,k},\tilde q)\in E$, we have

(2.4) $$Q_h(A,\tilde A)>0\Rightarrow\begin{cases}\sum_{j=1}^{r}\tilde\eta_{j,k+1}\leqq\sum_{j=1}^{r}\eta_{j,k}\\ \quad\text{for every }k=1,\ldots,r-1\quad\text{and}\\ \sum_{k=1}^{r}\tilde\eta_{j,k}+1_{\tilde q=j}\geqq\sum_{k=1}^{r}\eta_{j+1,k}+1_{q=j+1}\\ \quad\text{for every }j=1,\ldots,r-1\end{cases}$$

since $Q_h(A,\tilde A)$ is positive if and only if $P_h(A(0)=A,\ A(1)=\tilde A)$ is positive (note that $P_h(A(0)=A)>0$ for every $A\in E$), and $\sum_{j=1}^{r}\eta_{j,k}(x)$ is the number of jumps from $\{x+1,\ldots,x+m\}$ to $x+1-k$ and $\sum_{k=1}^{r}\eta_{j,k}(x)+1_{q(x)=j}$ is the number of jumps from $\{x+1-r,\ldots,x\}$ to $x+j$ for $x\in\mathbb{N}$ and $j,k\in\{1,\ldots,r\}$, respectively.

For paths without any self-intersection we have $\Lambda(x)\in\tilde{E}$ for all x, where

$$\tilde{E}:=\{\Lambda\in E \mid P_h(\Lambda(0)=\Lambda,\, l(x)\leqq 1 \quad \text{for } x=1-r, 2-r, \ldots, r)>0\}. \tag{2.5}$$

Note that $\tilde{E}$ is finite and that $g(\Lambda, \tilde{\Lambda})\leqq 2$ for $\Lambda, \tilde{\Lambda}\in\tilde{E}$. We prove now an irreducibility property of $(\Lambda(x))_{x\in\mathbb{N}_0}$ on $\tilde{E}$ which will be important in Sect. 3:

Lemma 2.6 *For every pair $(\Lambda, \tilde{\Lambda})\in\tilde{E}^2$ there are $\Lambda_1, \ldots, \Lambda_{3r-1}\in\tilde{E}$ such that $Q_h(\Lambda_{i-1}, \Lambda_i)>0$ and $g(\Lambda_{i-1}, \Lambda_i)\leqq 1$ for $i=1, \ldots, 3r$, where we put $\Lambda_0:=\Lambda$ and $\Lambda_{3r}:=\tilde{\Lambda}$.*

Proof. Given $\Lambda_i=((\eta_{j,k}^{(i)})_{j,k}, q^{(i)})\in\tilde{E}$ (for $i\in\{0, 3r\}$), we construct two paths $S^i=(S_0^i, \ldots, S_{k_i}^i)$ (with suited $k_0, k_{3r}\in\mathbb{N}$) without self-intersection, running within $\{1-r, \ldots, r\}$, performing exactly $\eta_{j,k}^{(i)}$ excursions with parameters $j, k\in\{1, \ldots, r\}$ over the line between 0 and 1, and ending in $q^{(i)}$. Now lengthen S^i by adding three jumps of height $r+1-q^{(i)}$, $r-1$, r (in this order) and put the lengthened path S^{3r} (after lifting by $3r$ sites) on the end of the lengthened path S^0 to obtain a path of length $k_0+k_{3r}+6$ without self-intersections, ending in $6r$, performing exactly $\eta_{j,k}^{(i)}$ excursions with parameters $j, k\in\{1, \ldots, r\}$ on the lines between i and $i+1$ (for $i\in\{0, 3r\}$) and the last jump over this line ending in $i+q^{(i)}$. So we see that

$$P_h(\Lambda(0)=\Lambda_0, \Lambda(3r)=\Lambda_{3r}, l(x)\leqq 1 \text{ for } x=1-r, \ldots, 3r)$$

is positive, which implies the assertion. □

2.2 *Formulation of the variational formula*

Now we formulate the main result of this section. The set of probability measures on an at most countable set X is denoted by $\mathcal{M}_1(X)$. Let

$$M_\theta:=\Big\{\nu\in\mathcal{M}_1(\tilde{E}\times\tilde{E}) \,\Big|\, \sum_{\Lambda,\tilde{\Lambda}\in\tilde{E}} g(\Lambda, \tilde{\Lambda})\, \nu(\Lambda, \tilde{\Lambda})=\frac{1}{\theta}, \tag{2.7}$$
$$\sum_{\tilde{\Lambda}\in\tilde{E}} \nu(\Lambda, \tilde{\Lambda})=\sum_{\tilde{\Lambda}\in\tilde{E}} \nu(\tilde{\Lambda}, \Lambda) \quad \text{for } \Lambda\in\tilde{E},$$
$$\text{and } \nu(\Lambda, \tilde{\Lambda})=0 \text{ if } g(\Lambda, \tilde{\Lambda})\geqq 2\Big\},$$

$$I_h(\nu):=\sum_{\Lambda,\tilde{\Lambda}\in E} \nu(\Lambda, \tilde{\Lambda}) \log\frac{\nu(\Lambda, \tilde{\Lambda})}{\bar{\nu}(\Lambda)\, Q_h(\Lambda, \tilde{\Lambda})} \tag{2.8}$$

for $\nu\in\mathcal{M}_1(E\times E)$ satisfying $\sum_{\tilde{\Lambda}\in E}\nu(\Lambda, \tilde{\Lambda})=\sum_{\tilde{\Lambda}\in E}\nu(\tilde{\Lambda}, \Lambda)=:\bar{\nu}(\Lambda)$ for every $\Lambda\in E$. (Define $I_h(\nu)$ to be $+\infty$ if there is some $(\Lambda, \tilde{\Lambda})$ satisfying $Q_h(\Lambda, \tilde{\Lambda})=0<\nu(\Lambda, \tilde{\Lambda})$.) View M_θ as a subset of $\mathcal{M}_1(E\times E)$. Then we have

Proposition 2.9 *The limit* (1.3) *exists, and*

$$\tilde{J}_h(\theta) = -\theta \inf_{\nu \in M_\theta} I_h(\nu). \tag{2.10}$$

The proof of this proposition requires several steps.

2.3 Reformulation in terms of the Markov chain

Define, for $n, k \in \mathbb{N}$,

$$V(n, k) = \{T > n, S_n = k, S_i > k \text{ for } i > n, \\ S_i \in [1, k-1] \text{ for } i = 1, \ldots, n-1\}, \tag{2.11}$$

so we have, if we define $\tilde{J}_h(\theta)$ to be the limit superior in (1.3) instead of the limit,

$$\tilde{J}_h(\theta) = \limsup_{n \to \infty} \frac{1}{n} \log P_h(V(n, \lfloor \theta n \rfloor)), \tag{2.12}$$

since escaping to the right has positive probability. Let, for $n, k \in \mathbb{N}$,

$$\tilde{V}(n, k) := \left\{ (A_0, \ldots, A_{k-1}) \in \tilde{E}^k \,\middle|\, A_0 = ((0), r),\ A_{k-1} = ((0), 1), \right.$$
$$\left. g(A_{x-1}, A_x) \leqq 1 \quad \text{for } x = 1, \ldots, k \quad \text{and} \quad \sum_{x=1}^{k} g(A_{x-1}, A_x) = n \right\},$$

where (0) denotes the $r \times r$-matrix consisting of zeros and $A_k := A_0$. So we have, by inserting the Markov property in (2.12),

$$\tilde{J}_h(\theta) = \limsup_{n \to \infty} \frac{1}{n} \log \sum_{(A_0, \ldots, A_{\lfloor \theta n \rfloor - 1}) \in \tilde{V}(n, \lfloor \theta n \rfloor)} \prod_{x=1}^{\lfloor \theta n \rfloor} Q_h(A_{x-1}, A_x) \tag{2.13}$$
$$\text{(always } A_{\lfloor \theta n \rfloor} := A_0\text{)}.$$

(Paths in $\{(A(0), \ldots, A(\lfloor \theta n \rfloor - 1)) \in \tilde{V}(n, \lfloor \theta n \rfloor)\}$ are allowed to spend some time in $-\mathbb{N}_0$ before they jump from 0 to r and stay in $\{1, \ldots, \lfloor \theta n \rfloor\}$ for n successive times. This little manipulation causes a factor which vanishes on the exponential scale.)

2.4 Empirical pair measures and large deviations

Next, we will write this expression in terms of empirical pair distributions which are driven by an independent uniformly on $\tilde{E}$ distributed sequence $(A_n)_{n \in \mathbb{N}_0}$: Define a random probability measure ν_n on $\tilde{E} \times \tilde{E}$ by

$$\nu_n = \frac{1}{n} \left(\sum_{k=1}^{n-1} \delta_{(A_{k-1}, A_k)} + \delta_{(A_{n-1}, A_0)} \right) \quad (n \in \mathbb{N}).$$

With the denotations $\hat{V}(n,k)=M_{\frac{k}{n}}$ (for $k, n\in\mathbb{N}$) and

(2.14) $$Q(\nu)=\sum_{\Lambda,\tilde{\Lambda}\in\tilde{E}}\nu(\Lambda,\tilde{\Lambda})\log Q_h(\Lambda,\tilde{\Lambda})\in[-\infty,0]\quad\text{for } \nu\in\mathcal{M}_1(\tilde{E}^2)$$

we have

(2.15) $$\tilde{J}_h(\theta)=\theta\log\#\tilde{E}+\limsup_{n\to\infty}\frac{1}{n}\log E_u\left(e^{\lfloor\theta n\rfloor Q(\nu_{\lfloor\theta n\rfloor})}1_{\{\nu_{\lfloor\theta n\rfloor}\in\hat{V}(n,\lfloor\theta n\rfloor)\}}\right),$$

where E_u denotes expectation with respect to $(\Lambda_n)_n$. The key point leading to this latter reformulation is the fact that we have $\nu_k\in\hat{V}(n,k)$ if and only if $(\Lambda_0,\ldots,\Lambda_{k-1})\in\tilde{V}(n,k)$, if we assume that $\Lambda_0=((0),r)$, $\Lambda_{k-1}=((0),1)$, and $Q(\nu_k)\in\mathbb{R}$.

The sequence $(\nu_n)_n$ obeys a large deviation principle on $\tilde{E}$ with rate function $\tilde{I}$, given by

$$\tilde{I}(\nu):=\sum_{\Lambda,\tilde{\Lambda}\in\tilde{E}}\nu(\Lambda,\tilde{\Lambda})\log\frac{\nu(\Lambda,\tilde{\Lambda})\#\tilde{E}}{\bar{\nu}(\Lambda)}$$

for $\nu\in M:=\{\mu\in\mathcal{M}_1(\tilde{E}\times\tilde{E})\mid\sum_{\tilde{\Lambda}\in\tilde{E}}\mu(\Lambda,\tilde{\Lambda})=\sum_{\tilde{\Lambda}\in\tilde{E}}\mu(\tilde{\Lambda},\Lambda)\text{ for }\Lambda\in\tilde{E}\}$ (Th. IX.4.3 in [4]). For small $\varepsilon>0$, the set

$$M_\theta^\varepsilon:=\bigcup_{|\delta|\leq\varepsilon}M_{\theta+\delta}$$

is compact, and we have $\hat{V}(n,\lfloor\theta n\rfloor)\subset M_\theta^\varepsilon$ for large $n\in\mathbb{N}$.

2.5 *Approximative variational formula*

We prove first that we have in fact

(2.16) $$\frac{1}{n}\log E_u\left(e^{\lfloor\theta n\rfloor Q(\nu_{\lfloor\theta n\rfloor})}1_{\{\nu_{\lfloor\theta n\rfloor}\in M_\theta^\varepsilon\}}\right)\xrightarrow{n\to\infty}\theta\sup_{\nu\in M_\theta^\varepsilon}(Q(\nu)-\tilde{I}(\nu)).$$

From Lemma 2.1.8 in [3] it follows that the limit superior of the l.h.s. is not bigger than the r.h.s. It is a little bit care needed to conclude the remaining part of (2.16), since M_θ^ε has no interior in M. For removing this technical problem, we will construct a Markov chain on

$$\Sigma:=\bigcup_{\Lambda\in\tilde{E}}\{\Lambda\}\times E_\Lambda$$

where

$$E_\Lambda:=\{\tilde{\Lambda}\in\tilde{E}: Q_h(\Lambda,\tilde{\Lambda})>0\quad\text{and}\quad g(\Lambda,\tilde{\Lambda})\leq 1\}.$$

We write $\sigma=(\sigma^{(1)}, \sigma^{(2)})$ for elements σ of Σ. Define a transition probability function $\Pi\colon \Sigma\times\Sigma\to[0,1]$ by

$$\Pi(\sigma,\tau):=\begin{cases}1/\#E_{\tau^{(1)}} & \text{if } \sigma^{(2)}=\tau^{(1)}\\ 0 & \text{else.}\end{cases}$$

From Lemma 2.6 it follows that Π is irreducible, so it satisfies condition (U) in [3, p. 100]. For $\sigma\in\Sigma$, let $P_\sigma\in\mathcal{M}_1(\Sigma^{\mathbb{N}_0})$ be the distribution of a Markov chain $(\sigma_n)_{n\in\mathbb{N}_0}$ with transition kernel Π and starting in σ. Then Theorem 4.1.43 in [3] states that the distributions of the empirical measures $L_n:=\frac{1}{n}\sum_{k=0}^{n-1}\delta_{\sigma_k}$ under P_σ satisfy a large deviation principle for every $\sigma\in\Sigma$ with good rate function $J_\Pi\colon\mathcal{M}_1(\Sigma)\to\bar{\mathbb{R}}$, given by

$$J_\Pi(\nu)=\sup\left\{-\sum_{\sigma\in\Sigma}\nu(\sigma)\log\frac{\Pi u(\sigma)}{u(\sigma)}\,\middle|\, u\colon\Sigma\to[1,\infty)\right\}$$

(here $\Pi u(\sigma)=\sum_{\tau\in\Sigma}u(\tau)\,\Pi(\sigma,\tau)$). Now define $\tilde{Q}, F\colon\mathcal{M}_1(\Sigma)\to\mathbb{R}$ by

$$\tilde{Q}(\nu)=\sum_{\sigma\in\Sigma}\nu(\sigma)\log Q_h(\sigma),$$

$$F(\nu)=\sum_{\sigma\in\Sigma}\nu(\sigma)\log\#E_{\sigma^{(2)}}$$

and

$$\tilde{M}_\theta^\varepsilon=\left\{\nu\in\mathcal{M}_1(\Sigma)\,\middle|\,\sum_{\sigma\in\Sigma}g(\sigma)\,\nu(\sigma)\in\left[\frac{1}{\theta+\varepsilon},\frac{1}{\theta-\varepsilon}\right]\right\}.$$

By translation into terms of the Markov chain $(\sigma_n)_n$ we conclude that the l.h.s. of (2.16) is, for every $\sigma\in\Sigma$, equal to

$$\theta\log\#\tilde{E}+o(1)+\frac{1}{n}\log E_{P_\sigma}\big(e^{\lfloor\theta n\rfloor(\tilde{Q}+F)(L_{\lfloor\theta n\rfloor})}\,1_{\{L_{\lfloor\theta n\rfloor}\in\tilde{M}_\theta^\varepsilon\}}\big).$$

(Here little changes have been made concerning the values of Λ_0, Λ_1 and $\Lambda_{\lfloor\theta n\rfloor-1}$. Their influence vanishes in the limit for n to infinity.) Since $\tilde{Q}$ and F are continuous and bounded on $\mathcal{M}_1(\Sigma)$ and since $\tilde{M}_\theta^\varepsilon$ is equal to its interior's closure in $\mathcal{M}_1(\Sigma)$, we arrived in a setting where standard large deviation arguments can be applied to deduce that the limit inferior of the last display is not smaller than

$$\theta\log\#\tilde{E}+\theta\sup_{\nu\in(\tilde{M}_\theta^\varepsilon)^\circ}(\tilde{Q}+F-J_\Pi)(\nu)$$
$$\ge\theta\log\#\tilde{E}-\theta\inf_{\nu\in\bigcup_{\delta<\varepsilon}M_\theta^\delta}\ \sup_{u\in[1,\infty)^\Sigma}\ \sum_{(\Lambda,\tilde{\Lambda})\in\Sigma}\nu(\Lambda,\tilde{\Lambda})\log\frac{u(\Lambda,\tilde{\Lambda})}{Q_h(\Lambda,\tilde{\Lambda})\sum_{\Lambda'\in E_{\tilde{\Lambda}}}u(\tilde{\Lambda},\Lambda')}.$$

Using the marginal property of ν, one sees that $u_1(\tilde{A}) := \sum_{A' \in E_{\tilde{A}}} u(\tilde{A}, A')$ can be replaced by $u_1(A)$ in this last display. Since $(\bar{\nu}(A)\, u(A, \tilde{A})/u_1(A))_{(A,\tilde{A}) \in \Sigma}$ is a probability measure, an application of Jensen's inequality shows that this supremum is not bigger than $I_h(\nu)$, and by the continuity of I_h, (2.16) follows.

2.6 Finish of the proof and perturbation invariance

Since the r.h.s. of (2.16) equals $-\theta \inf_{\nu \in M_\theta^\varepsilon} I_h(\nu) - \theta \log \# \tilde{E}$, the proof of Proposition 2.9 will follow at once from the following statements:

$$\lim_{\varepsilon \downarrow 0} \overline{\lim_{n \to \infty}} \frac{1}{n} \left| \log \frac{E_u(e^{\lfloor \theta n \rfloor Q(\nu_{\lfloor \theta n \rfloor})} 1_{\{\nu_{\lfloor \theta n \rfloor} \in M_\theta^\varepsilon\}})}{E_u(e^{\lfloor \theta n \rfloor Q(\nu_{\lfloor \theta n \rfloor})} 1_{\{\nu_{\lfloor \theta n \rfloor} \in \hat{V}(n, \lfloor \theta n \rfloor)\}})} \right| = 0 \tag{2.17}$$

and

$$\lim_{\varepsilon \downarrow 0} \inf_{\nu \in M_\theta^\varepsilon} I_h(\nu) = \inf_{\nu \in M_\theta} I_h(\nu). \tag{2.18}$$

(Note that every reformulation step until yet did not affect the existence or the value of the limit (1.3).) The proof of (2.18) is easy and left to the reader. We prove now the first assertion. For large $n \in \mathbb{N}$ the following implications hold:

$$\nu_{\lfloor \theta n \rfloor} \in \bigcup_{|i| \leq \frac{\varepsilon n}{2(\theta - \varepsilon)}} \hat{V}(n+i, \lfloor \theta n \rfloor) \Rightarrow \nu_{\lfloor \theta n \rfloor} \in M_\theta^\varepsilon \Rightarrow \nu_{\lfloor \theta n \rfloor} \in \bigcup_{|i| \leq \frac{2\varepsilon n}{\theta - \varepsilon}} \hat{V}(n+i, \lfloor \theta n \rfloor).$$

So it suffices to show that there is a function $f: (0, \infty) \to (0, \infty)$ satisfying $f(\varepsilon) \downarrow 0$ (for $\varepsilon \downarrow 0$) and

$$\left| \log \frac{E_u(e^{\lfloor \theta n \rfloor Q(\nu_{\lfloor \theta n \rfloor})} 1_{\{\nu_{\lfloor \theta n \rfloor} \in \hat{V}(n+\varepsilon_n, \lfloor \theta n \rfloor)\}})}{E_u(e^{\lfloor \theta n \rfloor Q(\nu_{\lfloor \theta n \rfloor})} 1_{\{\nu_{\lfloor \theta n \rfloor} \in \hat{V}(n, \lfloor \theta n \rfloor)\}})} \right| \leq n f(\varepsilon) \qquad (n \in \mathbb{N}) \tag{2.19}$$

for small $\varepsilon > 0$ and every sequence $(\varepsilon_n)_n$ in $\mathbb{Z}$ satisfying $|\varepsilon_n| \leq \varepsilon n$. Let such ε, $(\varepsilon_n)_n$ be given. Assume that $\varepsilon_n > 0$ (the other case is similar). Remember that our manipulations of path classes to get from the r.h.s. to the reformulation in (2.19) essentially consist of forcing the path to stay above $\lfloor \theta n \rfloor$ after time n and to allow him to stay some time in $-\mathbb{N}_0$ before running within $\{1, \ldots, \lfloor \theta n \rfloor\}$. So we may handle this problem in terms of self-avoiding walks instead of empirical pair distributions and will construct two maps Γ: $V(n, \lfloor \theta n \rfloor) \to V(n+\varepsilon_n, \lfloor \theta n \rfloor)$ and $\tilde{\Gamma}$: $V(n+\varepsilon_n, \lfloor \theta n \rfloor) \to V(n, \lfloor \theta n \rfloor)$ having not too much entropy, i.e. satisfying

$$\overline{\lim_{\varepsilon \downarrow 0}}\, \overline{\lim_{n \to \infty}} \frac{1}{n} \log \sup_{S^* \in V(n+\varepsilon_n, \lfloor \theta n \rfloor)} \# \Gamma^{-1}(S^*) = 0 \tag{2.20}$$

and

$$\overline{\lim_{\varepsilon \downarrow 0}}\, \overline{\lim_{n \to \infty}} \frac{1}{n} \log \sup_{S^* \in V(n, \lfloor \theta n \rfloor)} \# \tilde{\Gamma}^{-1}(S^*) = 0, \tag{2.21}$$

respectively. Then (2.19) will follow from the facts

$$\sup_{S\in V(n,\lfloor\theta n\rfloor)} \frac{P_h(S)}{P_h(\Gamma(S))} \leqq e^{nO(\varepsilon)} \quad \text{and} \quad \sup_{S\in V(n+\varepsilon_n,\lfloor\theta n\rfloor)} \frac{P_h(S)}{P_h(\tilde{\Gamma}(S))} \leqq e^{nO(\varepsilon)}$$

which we also will show.

Construction of Γ. Let a path $S=(S_k)_k\in V(n,\lfloor\theta n\rfloor)$ be given and consider the times $1<t_1<t_2<\ldots<t_{n^*}\leqq n-1$ (with some $n^*<n-1$) such that $S_{t_i}>S_{t_i-1}$ and $k_i:=\#\{x\in\{S_{t_i-1}+1,\ldots,S_{t_i}-1\}: l(x)=0\}$ is not zero for $i=1,2,\ldots,n^*$. Now delete the t_i-th step of the path and replace it by k_i+1 suited connecting upward steps such that the new path hits every site in $\{S_{t_i-1}+1,\ldots,S_{t_i}-1\}$. Perform this procedure for $i=1,2,\ldots,i_\varepsilon$ until the length of the path is increased by exactly ε_n steps (the new path needs not hit every site in $\{S_{t_{i_\varepsilon}-1}+1,\ldots,S_{t_{i_\varepsilon}}\}$). The local times of the arising path $\Gamma(S)\in V(n+\varepsilon_n,\lfloor\theta n\rfloor)$ satisfy $l(x,\Gamma(S))\geqq l(x,S)$ for every $x\in\mathbb{Z}$. There are less than $\binom{n+\varepsilon_n}{\varepsilon_n}$ sites $x_1,x_2,\ldots,x_{\varepsilon_n}\in\{r+1,\ldots,\lfloor\theta n\rfloor-r\}$ such that some path $\tilde{S}\in V(n,\lfloor\theta n\rfloor)$ exists satisfying $\Gamma(\tilde{S})=\Gamma(S)$ and $l(x_i,\Gamma(\tilde{S}))=1>l(x_i,\tilde{S})$ for $i=1,\ldots,\varepsilon_n$. For a given choice of $x_1,\ldots,x_{\varepsilon_n}$, this path $\tilde{S}$ is uniquely determined. Now Stirling's formula yields (2.20). Since Γ removes at least ε_n/r steps and creates at last $2\varepsilon_n$ steps, we obtain a bound $P_h(S)/P_h(\Gamma(S))\leqq Z_h^{\varepsilon_n} p_{\max}^{\varepsilon_n/r} p_{\min}^{-2\varepsilon_n}$, which is not bigger than $e^{nO(\varepsilon)}$. (Here $p_{\max}$ and $p_{\min}$ denote the maximal and minimal value in $\{p_1,\ldots,p_r\}$, respectively.)

Construction of $\tilde{\Gamma}$. Consider a path $S^0=(S_k^0)_k\in V(n+\varepsilon_n,\lfloor\theta n\rfloor)$. We are going to describe an algorithm such that the j-th step produces a path $S^j\in V(n+\varepsilon_n-j,\lfloor\theta n\rfloor)$. Afterwards we put $\tilde{\Gamma}(S^0):=S^{\varepsilon_n}$. For $j=1,\ldots,\varepsilon_n$ we perform the following procedure. If there is a smallest $x\in\{r+1,\ldots,\lfloor\theta n\rfloor-r\}$ such that there is some $k\in\mathbb{N}$ satisfying $x=S_k^{j-1}<\min\{S_{k-1}^{j-1},S_{k+1}^{j-1}\}$, then replace the k-th and the $(k+1)$-th step by one connecting step and call the resulting path $S^j=(S_k^j)_{k\in\mathbb{N}}$. If there is no such x, then S^{j-1} is strictly increasing up to time $n+\varepsilon_n-j$. In this case choose the smallest $k\in\{2,\ldots,n+\varepsilon_n-j-1\}$ such that $S_{k+1}^{j-1}-S_{k-1}^{j-1}\leqq r$ and again replace the k-th and $(k+1)$-th step by one suited jump such that the resulting path S^j is in $V(n+\varepsilon_n-j,\lfloor\theta n\rfloor)$. If there is no such k, then choose the smallest k satisfying $S_{k+2}^{j-1}-S_{k-1}^{j-1}\leqq 2r$ (or, next, $S_{k+3}^{j-1}-S_{k-1}^{j-1}\leqq 3r$, $S_{k+4}^{j-1}-S_{k-1}^{j-1}\leqq 4r$ and so on) and replace the three steps k, $k+1$, $k+2$ (or the four steps $k,\ldots,k+3$, respectively, and so on) by two (or three, or four, respectively) suited connecting upward steps.

The ε_n-th step of this algorithm replaces $c(S)$ jumps by $c(S)-1$ jumps, where $c(S)\in\{2,3,\ldots,\lfloor r/(r-\theta)\rfloor\}$ is a suitable constant. (If $c(S)$ would be larger than $c_\theta:=\lfloor r/(r-\theta)\rfloor$, then we would have $S_{k+c_\theta-1}^{\varepsilon_n-1}-S_{k-1}^{\varepsilon_n-1}>(c_\theta-1)r$ for every $k=1,\ldots,n-c_\theta$, and it would follow that $S_n^{\varepsilon_n-1}>\lfloor\theta n\rfloor$.)

We may write $\tilde{\Gamma}=\tilde{\Gamma}_{c_\theta}\circ\tilde{\Gamma}_{c_\theta-1}\circ\ldots\circ\tilde{\Gamma}_1$, where $\tilde{\Gamma}_1$ replaces (zero or several times) two certain steps by one jump and, for $i\geqq 2$, the map $\tilde{\Gamma}_i$ replaces (zero or several times) certain i upward steps by suited $i-1$ upward steps. For every $i=1,\ldots,c_\theta$ there are less than $(c_\theta-1)\,\varepsilon_n\binom{n}{(c_\theta-1)\,\varepsilon_n}$ subsets $\{y_1,\ldots,y_l\}$ of $\{r+1,\ldots,\lfloor\theta n\rfloor-r\}$ (with $l\leqq(c_\theta-1)\,\varepsilon_n$) and less than $c_\theta\,\varepsilon_n\binom{\lfloor\theta n\rfloor}{c_\theta\,\varepsilon_n}$ subsets $\{x_1,\ldots,x_k\}$ of

$\{r+1, \ldots, \lfloor \theta n \rfloor - r\}$ (with $k \leqq c_\theta \varepsilon_n$) such that there is a path $\tilde{S}$ in $\bigcup_{\gamma=1}^{\varepsilon_n} V(n+\gamma, \lfloor \theta n \rfloor)$ satisfying $\tilde{\Gamma}_i(\tilde{S}) = \tilde{\Gamma}_i(\tilde{\Gamma}_{i-1}(\ldots(\tilde{\Gamma}_1(S^0)\ldots)$ and $l(x_\gamma, \tilde{S}) = 0 < 1 = l(x_\gamma, \tilde{\Gamma}_i(\tilde{S}))$ (for $\gamma = 1, \ldots, k$) and $l(y_\gamma, \tilde{S}) = 1 > 0 = l(y_\gamma, \tilde{\Gamma}_i(\tilde{S}))$ (for $\gamma = 1, \ldots, l$). If k, l, $\{x_1, \ldots, x_k\}$ and $\{y_1, \ldots, y_l\}$ are given with this property, then the path $\tilde{S}$ is uniquely determined since the algorithm $\tilde{\Gamma}_i$ works in a strictly increasing way and can be followed back. Now (2.21) is obvious by using Stirling's formula. The fact that $\tilde{\Gamma}$ changes the probability not too much is analogous to this property of Γ. □

The claim (2.19) implies the invariance of the limit (2.15) under perturbations of size $o(n)$ in the first argument of $\hat{V}$. A similar proof applies to the perturbation invariance in the second argument. Since the manipulations which we performed to get from (1.3) to the reformulation (2.15) do not affect this invariance, we can state, for future reference, the following corollary.

Remark 2.22 The limits (1.3), (2.12) and (2.15) are invariant under perturbations of size $o(n)$ in n and $\lfloor \theta n \rfloor$. In particular, for sequences $(\varepsilon_n)_n$ and $(\delta_n)_n$ in $\mathbb{Z}$ which are $o(n)$, for every $h \geqq 0$, $c > 0$ and $\theta \in [c, rc]$, we have:

$$c\,\tilde{J}_h\left(\frac{\theta}{c}\right) = \lim_{n\to\infty} \frac{1}{n} \log P_h(T > \lfloor cn \rfloor + \delta_n,$$
$$0 < S_1, \ldots, S_{\lfloor cn \rfloor + \delta_n - 1} < S_{\lfloor cn \rfloor + \delta_n} = \lfloor \theta n \rfloor + \varepsilon_n).$$

3 Analysis of the variational formula

Let further on $h > 0$, $r \geqq 2$ and $\theta \in (1, r)$ be fixed. We will show in this section that the infimum in (2.10) is a uniquely attained minimum, and we will analyze the minimizer $v^\theta \in M_\theta$ by variational techniques. This will lead to a characterization of $\tilde{J}_h(\theta)$ in terms of the Frobenius eigenvalue $\lambda_h(b)$ of a certain one-parameter family of non-negative primitive matrices A_b which we will introduce now. Recall the definitions (2.3) and (2.5).

3.1 Eigenvalue properties

For $b \in \mathbb{R}$ and $\Lambda, \tilde{\Lambda} \in \tilde{E}$ define

$$A_b(\Lambda, \tilde{\Lambda}) = \begin{cases} Q_h(\Lambda, \tilde{\Lambda})\, e^{b g(\Lambda, \tilde{\Lambda})} & \text{if } g(\Lambda, \tilde{\Lambda}) \leqq 1, \\ 0 & \text{else.} \end{cases} \tag{3.1}$$

The matrix $A_b := (A_b(\Lambda, \tilde{\Lambda}))_{\Lambda, \tilde{\Lambda} \in \tilde{E}}$ has non-negative components and is primitive by Lemma 2.6. By $\lambda_h(b)$ we denote the Frobenius eigenvalue of A_b, i.e. the only eigenvalue with positive left and right eigenvectors which we denote by τ_b^l, $\tau_b^r \in (0, \infty)^{\tilde{E}}$, respectively, and we assume them to be normed, i.e. $\langle \tau_b^l, \tau_b^r \rangle = 1$ (where $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $\mathbb{R}^{\tilde{E}}$). Since the representation

$$\lambda_h(b) = \lim_{n\to\infty} (A_b^n(\Lambda, \Lambda))^{1/n} \qquad (b \in \mathbb{R}) \tag{3.2}$$

holds independently of $\Lambda \in \tilde{E}$ ([9, p. 200–201]; note that A_b is aperiodic since it is primitive), the function λ_h is analytic in a neighbourhood of $\mathbb{R}$ (use Vitali's theorem). The eigenvectors τ_b^l and τ_b^r are unique if we determine some fixed $\tau_b^r(\Lambda)$ to be 1. Their coefficients are differentiable with respect to the variable b as can be derived from elementary considerations.

We compute the derivation λ_h' of λ_h: For $b \in \mathbb{R}$ we have

$$\begin{aligned}
\lambda_h'(b) &= \sum_{\Lambda, \tilde{\Lambda} \in \tilde{E}} \frac{d}{db} (\tau_b^l(\Lambda)\, A_b(\Lambda, \tilde{\Lambda})\, \tau_b^r(\tilde{\Lambda})) \\
&= \langle \tau_b^l, A_b' \tau_b^r \rangle + \sum_{\Lambda, \tilde{\Lambda} \in \tilde{E}} A_b(\Lambda, \tilde{\Lambda}) \left(\tau_b^r(\tilde{\Lambda}) \frac{d}{db} \tau_b^l(\Lambda) + \tau_b^l(\Lambda) \frac{d}{db} \tau_b^r(\tilde{\Lambda}) \right) \\
&= \langle \tau_b^l, A_b' \tau_b^r \rangle + \lambda_h(b) \sum_{\Lambda \in \tilde{E}} \left(\tau_b^r(\Lambda) \frac{d}{db} \tau_b^l(\Lambda) + \tau_b^l(\Lambda) \frac{d}{db} \tau_b^r(\Lambda) \right) \\
&= \langle \tau_b^l, A_b' \tau_b^r \rangle + \lambda_h(b) \frac{d}{db} \langle \tau_b^l, \tau_b^r \rangle \\
&= \langle \tau_b^l, A_b' \tau_b^r \rangle,
\end{aligned}$$

where the matrix A_b' is given by

$$A_b'(\Lambda, \tilde{\Lambda}) := \frac{d}{db} A_b(\Lambda, \tilde{\Lambda}) = \begin{cases} A_b(\Lambda, \tilde{\Lambda}) & \text{if } g(\Lambda, \tilde{\Lambda}) = 1, \\ 0 & \text{if } g(\Lambda, \tilde{\Lambda}) \in \{0, 2\}. \end{cases}$$

By [9, Ex. 1.11], λ_h is strictly increasing with

$$\lim_{b \to +\infty} \lambda_h(b) = +\infty, \qquad \lim_{b \to -\infty} \lambda_h(b) < 1. \tag{3.3}$$

We will show now the log-convexity of λ_h, i.e. the monotonicity of λ_h'/λ_h: Because of (3.2), it suffices to show the convexity of the mapping $b \mapsto \log A_b^n(\Lambda, \Lambda)$ for every $\Lambda \in \tilde{E}$ and $n \in \mathbb{N}$. The second derivative of this map is easily seen to be positive by writing out the n-fold matrix product and symmetrizing the sum which arises. We will see in the sequel that λ_h'/λ_h is not constant, and, by analyticity, the strictness of the monotonicity will follow.

3.2 *Minimizer and positivity*

Since I_h is finite and continuous on the compact set

$$M_\theta^Q := \{ \nu \in M_\theta \mid Q_h(\Lambda, \tilde{\Lambda}) = 0 \Rightarrow \nu(\Lambda, \tilde{\Lambda}) = 0 \},$$

the existence of a minimizer in (2.10) is obvious. Note that $I_h = +\infty$ on $M_\theta \setminus M_\theta^Q$. For applying variational techniques to any such minimizer, the following is important:

Lemma 3.4 *Every minimizer $\nu^\theta \in M_\theta$ of I_h on M_θ has the following property: For every $(\Lambda, \tilde{\Lambda})$ in $\tilde{E}^2$ satisfying $Q_h(\Lambda, \tilde{\Lambda}) > 0$ and $g(\Lambda, \Lambda) \leqq 1$ it holds $\nu^\theta(\Lambda, \tilde{\Lambda}) > 0$.*

Proof. We write v instead of v^θ. The lemma will follow by induction from the following fact which we will prove now: For every $\Lambda_1, \Lambda_2, \tilde{\Lambda}_1, \tilde{\Lambda}_2 \in \tilde{E}$ satisfying $v(\Lambda_1, \tilde{\Lambda}_1)\, Q_h(\Lambda_2, \tilde{\Lambda}_1)\, Q_h(\Lambda_1, \tilde{\Lambda}_2) > 0$, $g(\Lambda_2, \tilde{\Lambda}_1) \leqq 1$, and $g(\Lambda_1, \tilde{\Lambda}_2) \leqq 1$ it holds $v(\Lambda_2, \tilde{\Lambda}_1) > 0$ and $v(\Lambda_1, \tilde{\Lambda}_2) > 0$.

Let Λ_1, $\tilde{\Lambda}_1$, Λ_2, $\tilde{\Lambda}_2$ be given as above, and we assume that $v(\Lambda_1, \tilde{\Lambda}_2)\, v(\Lambda_2, \tilde{\Lambda}_1) = 0$. We will construct a function $t\colon \tilde{E}^2 \to \mathbb{R}$ such that

$$\sum_{\Lambda, \tilde{\Lambda} \in \tilde{E}} t(\Lambda, \tilde{\Lambda}) = 0 = \sum_{\Lambda, \tilde{\Lambda} \in \tilde{E}} g(\Lambda, \tilde{\Lambda})\, t(\Lambda, \tilde{\Lambda}), \tag{3.5}$$

$$\sum_{\tilde{\Lambda} \in \tilde{E}} t(\Lambda, \tilde{\Lambda}) = \sum_{\tilde{\Lambda} \in \tilde{E}} t(\tilde{\Lambda}, \Lambda) \qquad \text{for every } \Lambda \in \tilde{E}, \tag{3.6}$$

$$t(\Lambda, \tilde{\Lambda}) > 0 \Rightarrow (Q_h(\Lambda, \tilde{\Lambda}) > 0 \text{ and } g(\Lambda, \tilde{\Lambda}) \leqq 1), \tag{3.7}$$

$$t(\Lambda, \tilde{\Lambda}) < 0 \Rightarrow v(\Lambda, \tilde{\Lambda}) > 0, \tag{3.8}$$

$$t(\Lambda_1, \tilde{\Lambda}_2) > 0 \quad \text{and} \quad t(\Lambda_2, \tilde{\Lambda}_1) > 0. \tag{3.9}$$

Then, for small $\varepsilon > 0$, the measure $v_\varepsilon := v + \varepsilon t$ is in M_θ and satisfies $v_\varepsilon(\Lambda_1, \tilde{\Lambda}_2)\, v_\varepsilon(\Lambda_2, \tilde{\Lambda}_1) > 0$ and $I_h(v_\varepsilon) < I_h(v)$, as we will see now. We write $f(x) = x \log x$ and use the inequalities $(x-y)(1+\log y) \leqq f(x) - f(y) \leqq (x-y)(1+\log x)$ for $0 < y < x$. For sufficiently small $\varepsilon > 0$ it holds with suitable constants $c_1, c_2, c_3, c_4, c_5 \in \mathbb{R}$ (without loss of generality we assume $v(\Lambda_1, \tilde{\Lambda}_2) = 0$):

$$\begin{aligned}
I_h(v_\varepsilon) - I_h(v) = {} & -\varepsilon \sum_{\Lambda, \tilde{\Lambda}} t(\Lambda, \tilde{\Lambda}) \log Q_h(\Lambda, \tilde{\Lambda}) \\
& + \sum_{\Lambda, \tilde{\Lambda} : v(\Lambda, \tilde{\Lambda}) > 0} [f((v + \varepsilon t)(\Lambda, \tilde{\Lambda})) - f(v(\Lambda, \tilde{\Lambda})) \\
& - (f((\bar{v} + \varepsilon \bar{t})(\Lambda)) - f(\bar{v}(\Lambda)))] \\
& + \sum_{\Lambda, \tilde{\Lambda} : \bar{v}(\Lambda) = 0} (f(\varepsilon t(\Lambda, \tilde{\Lambda})) - f(\varepsilon \bar{t}(\Lambda))) \\
& + \sum_{\Lambda, \tilde{\Lambda} : \bar{v}(\Lambda) > 0 = v(\Lambda, \tilde{\Lambda})} [f(\varepsilon t(\Lambda, \tilde{\Lambda})) - f((\bar{v} + \varepsilon \bar{t})(\Lambda)) + f(\bar{v}(\Lambda))] \\
\leqq {} & \varepsilon c_1 + \varepsilon c_2 + \varepsilon c_3 + \varepsilon c_4 + \sum_{\Lambda, \tilde{\Lambda} : \bar{v}(\Lambda) > 0 = v(\Lambda, \tilde{\Lambda})} f(\varepsilon t(\Lambda, \tilde{\Lambda})) \\
\leqq {} & \varepsilon (c_1 + c_2 + c_3 + c_4) + c_5\, \varepsilon \log \varepsilon \\
< {} & 0.
\end{aligned}$$

(Note that c_5 is positive since $\{(\Lambda, \tilde{\Lambda}) : \bar{v}(\Lambda) > 0 = v(\Lambda, \tilde{\Lambda})\}$ is not empty.)

Construction of t. First we show that there are $n \in \mathbb{N}$ and $\Lambda_1^*, \ldots, \Lambda_n^* \in \tilde{E}$ satisfying $v(\Lambda_i^*, \Lambda_{i+1}^*) > 0$ for $i = 1, \ldots, n$, where we put $\Lambda_i^* := \Lambda_{i(\text{mod}\, n)}^*$ for $i \in \mathbb{Z}$. Take some pair $(\mu_1, \mu_2) \in \tilde{E}^2$ such that $v(\mu_1, \mu_2) > 0$. If $\mu_1, \ldots, \mu_i$ are already constructed, then choose $\mu_{i+1} \in \tilde{E}$ such that $v(\mu_i, \mu_{i+1}) > 0$ (this is possible since $0 < \bar{v}(\mu_i) = \sum_{\mu \in \tilde{E}} v(\mu_i, \mu)$). Continue until some μ_i is repeated in the j-th step. Then $(\mu_i, \ldots, \mu_j)$ has the demanded property.

It is important noting that $\sum_{i=1}^{n} g(\Lambda_i^*, \Lambda_{i+1}^*)$ is positive. This follows from the fact that $Q_h(\Lambda_i^*, \Lambda_{i+1}^*)>0$ (otherwise $I_h(v)=+\infty$) which leads, by (2.4), to

$$\sum_{j=1}^{r} \eta_{j,k+1}^{(i+1)} \leqq \sum_{j=1}^{r} \eta_{j,k}^{(i)} \qquad (k=1, \ldots, r-1)$$

and

$$\sum_{k=1}^{r} \eta_{j,k}^{(i+1)} + 1_{q^{(i+1)}=j} \geqq \sum_{k=1}^{r} \eta_{j+1,k}^{(i)} + 1_{q^{(i)}=j+1} \qquad (j=1, \ldots, r-1)$$

if $\Lambda_i^* = ((\eta_{j,k}^{(i)})_{j,k=1,\ldots,r}, q^{(i)})$. In the case $g(\Lambda_i^*, \Lambda_{i+1}^*)=0$ for all i we would get, by induction, $\eta_{j,k}^{(i)}=0$ for all j, $k=1, \ldots, r$ and $i\in\mathbb{Z}$, and then a contradiction would follow by considering $q^{(i)}$.

Next, by Lemma 2.6, we may choose $\Lambda_1^i, \ldots, \Lambda_{3r-1}^i \in \tilde{E}$ for $i=1, 2$ such that $Q_h(\Lambda_k^i, \Lambda_{k+1}^i)>0$ and $g(\Lambda_k^i, \Lambda_{k+1}^i)\leqq 1$ for $k=0, 1, \ldots, 3r-1$ where we put $\Lambda_0^1 := \tilde{\Lambda}_2$, $\Lambda_{3r}^1 := \Lambda_1$, $\Lambda_0^2 := \tilde{\Lambda}_1$, and $\Lambda_{3r}^2 := \Lambda_2$. Now take positive numbers a, b, and c and define

$$\begin{aligned} -a &= t(\Lambda_i^*, \Lambda_{i+1}^*) && (i=1, \ldots, n), \\ b &= t(\Lambda_i^1, \Lambda_{i+1}^1) && (i=0, \ldots, 3r, \text{ where } \Lambda_{3r+1}^1 := \Lambda_0^1) \\ c &= t(\Lambda_i^2, \Lambda_{i+1}^2) && (i=0, \ldots, 3r, \text{ where } \Lambda_{3r+1}^2 := \Lambda_0^2). \end{aligned}$$

(If $t(\Lambda, \tilde{\Lambda})$ is doubly defined for some $(\Lambda, \tilde{\Lambda})\in\tilde{E}^2$, then this is to be understood to add up as often as it is defined.)

For every other pair $(\Lambda, \tilde{\Lambda})\in\tilde{E}^2$ set $t(\Lambda, \tilde{\Lambda})=0$. The conditions (3.6) through (3.9) are satisfied automatically, and condition (3.5) reads as follows:

$$a\,n = (b+c)(3r+1)$$

and

$$a \sum_{i=1}^{n} g(\Lambda_i^*, \Lambda_{i+1}^*) = b \sum_{i=0}^{3r} g(\Lambda_i^1, \Lambda_{i+1}^1) + c \sum_{i=0}^{3r} g(\Lambda_i^2, \Lambda_{i+1}^2).$$

Of course, we can choose $\Lambda_1^1, \ldots, \Lambda_{3r-1}^1$ and $\Lambda_1^2, \ldots, \Lambda_{3r-1}^2$ in such a manner that the coefficients for b and c in the second equation are different. So it is possible to fulfil condition (3.5) by choosing suitable $a, b, c>0$. □

3.3 Variations and identification of the minimizer

Now we can apply variational techniques to any minimizer v^θ of I_h on M_θ: Let $t\in\mathbb{R}^{\tilde{E}^2}$ satisfy the conditions

(3.10) $$t(\Lambda, \tilde{\Lambda})=0 \quad \text{if } (Q_h(\Lambda, \tilde{\Lambda})=0 \quad \text{or} \quad g(\Lambda, \tilde{\Lambda})\geqq 2),$$

(3.11) $$t\perp 1, \quad t\perp g, \quad t\perp B_\Lambda \quad (\Lambda\in\tilde{E}),$$

where $\perp$ refers to the standard inner product $\langle\cdot|\cdot\rangle$ on $\mathbb{R}^{\tilde{E}^2}$, $1\in\mathbb{R}^{\tilde{E}^2}$ denotes the constant function and $B_\Lambda\in\mathbb{R}^{\tilde{E}^2}$ is defined by $B_\Lambda(\Lambda_1, \Lambda_2)=\delta_{\Lambda\Lambda_1}-\delta_{\Lambda\Lambda_2}$ for

$\Lambda_1, \Lambda_2 \in \tilde{E}$. For such a function t the measure $\nu_\varepsilon := \nu^\theta + \varepsilon t$ is in M_θ for $\varepsilon \in \mathbb{R}$ with small $|\varepsilon|$, and it follows

$$0 = \frac{d}{d\varepsilon} I_h(\nu_\varepsilon)\bigg|_{\varepsilon=0} = \langle t \mid u \rangle$$

where

$$(3.12) \qquad u(\Lambda, \tilde{\Lambda}) := \begin{cases} \log \dfrac{\nu^\theta(\Lambda, \tilde{\Lambda})}{\bar{\nu}^\theta(\Lambda)\, Q_h(\Lambda, \tilde{\Lambda})} & \text{if } Q_h(\Lambda, \tilde{\Lambda}) > 0 \\ 0 & \text{else.} \end{cases}$$

Hence, there exist constants a, b, c_Λ (for $\Lambda \in \tilde{E}$) such that

$$u = a + b\,g + \sum_{\Lambda \in \tilde{E}} c_\Lambda B_\Lambda,$$

i.e.

$$(3.13) \qquad u(\Lambda, \tilde{\Lambda}) = a + b\,g(\Lambda, \tilde{\Lambda}) + c_\Lambda - c_{\tilde{\Lambda}}$$
$$\text{for all } \Lambda, \tilde{\Lambda} \in \tilde{E} \text{ satisfying } Q_h(\Lambda, \tilde{\Lambda}) > 0 \text{ and } g(\Lambda, \tilde{\Lambda}) \leqq 1$$

or, equivalently,

$$\nu^\theta(\Lambda, \tilde{\Lambda}) = \bar{\nu}^\theta(\Lambda)\, e^{c_\Lambda} A_b(\Lambda, \tilde{\Lambda})\, e^{-c_{\tilde{\Lambda}}} e^{a} \qquad \text{for all } \Lambda, \tilde{\Lambda} \in \tilde{E}.$$

The conditions $\sum_{\tilde{\Lambda} \in \tilde{E}} \nu^\theta(\tilde{\Lambda}, \Lambda) = \bar{\nu}^\theta(\Lambda) = \sum_{\tilde{\Lambda} \in \tilde{E}} \nu(\Lambda, \tilde{\Lambda})$ for $\Lambda \in \tilde{E}$ mean that

$$\tau^l := (\bar{\nu}^\theta(\Lambda)\, e^{c_\Lambda})_{\Lambda \in \tilde{E}} \quad \text{and} \quad \tau^r := (e^{-c_\Lambda})_{\Lambda \in \tilde{E}}$$

are positive left and right eigenvectors of the matrix A_b belonging to the eigenvalue e^{-a}, respectively. By the fact $\nu^\theta \in \mathcal{M}_1(\tilde{E}^2)$, they are normed, i.e. $\langle \tau^l, \tau^r \rangle = 1$. So we have $a = -\log \lambda_h(b)$ and can assume $\tau^l = \tau^l_b$ and $\tau^r = \tau^r_b$.

The condition $\sum_{\Lambda, \tilde{\Lambda} \in \tilde{E}} g(\Lambda, \tilde{\Lambda})\, \nu^\theta(\Lambda, \tilde{\Lambda}) = \frac{1}{\theta}$ reads as

$$(3.14) \qquad \frac{1}{\theta} = \frac{\langle \tau^l_b, A'_b \tau^r_b \rangle}{\langle \tau^l_b, A_b \tau^r_b \rangle} = \frac{\lambda'_h(b)}{\lambda_h(b)}.$$

So we see that the increasing and analytic function λ'_h/λ_h is not constant, hence it is strictly increasing, i.e. the function λ_h is strictly log-convex. The last display determines a strictly decreasing function

$$(3.15) \qquad b_h := \left(\frac{\lambda_h}{\lambda'_h}\right)^{-1} : (1, r) \to \mathbb{R}, \qquad \frac{\lambda'_h}{\lambda_h}(b_h(\theta)) = \frac{1}{\theta}.$$

(Now it is clear that the minimizer is unique.) From (3.12) and (3.13) one deduces easily that (1.5) holds. In particular, $\tilde{J}_h$ is strictly concave and analytic. Since the r.h.s. of (1.6) exists in $(1, r)$ by (3.3) and since it is obviously a zero of $\tilde{J}'_h$, the strict maximum point $\Theta(h)$ of $\tilde{J}_h$ in $(1, r)$ satisfies (1.6). As can be seen from the definition (1.3), it is maximal for $\tilde{J}_h$ on the whole interval $[0, r]$. From the fact

$$(3.16) \qquad \tilde{J}_h(\theta) = \tilde{J}_0(\theta) + h\theta - \log Z_h \qquad (h \geqq 0,\ \theta \in [0, r]),$$

which is implied by our choose of step distribution of the path, it follows that $\tilde{J}_0$ is analytic and strictly concave, too. The equation $\tilde{J}_0'(\theta) = 0$ has a solution $\theta = \Theta(0)$ in $(1, r)$ as can be derived from (3.16) and the second line of (1.5) using the fact that $\lambda_h' < \lambda_h$. Furthermore, Θ: $[0, \infty) \to (1, r)$ is strictly increasing and analytic, as is seen from the representation

$$(3.17) \qquad \Theta(h) = (\tilde{J}_0')^{-1}(-h) \qquad (h \geqq 0).$$

From this last display it also follows the fact that $\lim_{h\uparrow\infty} \Theta(h) = r$. (Note that, in particular, it follows $\lim_{\theta\uparrow r} \tilde{J}_0'(\theta) = -\infty$.) So we proved Theorem 1.4 if we define $\Theta(h)$ to be the maximum point of $\tilde{J}_h$.

4 Proof of Theorem 1.1

In this section we will prove

Proposition 4.1 *For* $h \geqq 0$, $\theta \in [0, r]$ *the limit* (1.2) *exists and for* $\varepsilon > 0$ *it holds* $\max\{J_h(\theta) \colon |\theta - \Theta(h)| \geqq \varepsilon\} < \tilde{J}_h(\Theta(h))$.

During the proof, it will become clear that the limit (1.2) is invariant under perturbations of size $o(n)$, like the limit (1.3). So it follows

Corollary 4.2 *For* $h \geqq 0$ *it holds*

$$\lim_{n\to\infty} \frac{1}{n} \log P_h(T > n)$$
$$= \sup_{\theta\in[0,r]} J_h(\theta) = \tilde{J}_h(\Theta(h)) \qquad (= -\lambda_h^{-1}(1)\ \textit{for}\ h > 0\ \textit{by}\ (1.5)).$$

Theorem 1.1 is implied as follows. For $\varepsilon > 0$ and $h > 0$ one obtains from scaling arguments like those which are performed in the end of the proof of Proposition 4.1 that

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log P_h\left(T > n, \max_{k=1}^{n} \left|\frac{S_k}{k} - \Theta(h)\right| \geqq \varepsilon\right)$$
$$\leqq \sup\{c\, J_h(\theta) + (1-c)\, \tilde{J}_h(\Theta(h)) \colon c \in [\varepsilon/(r - \Theta(h)), 1],\ |\theta - \Theta(h)| \geqq \varepsilon\},$$

which is strictly smaller than $\tilde{J}_h(\Theta(h))$ by Proposition 4.1. So we see that $\mathbb{P}_n^h(\{f \in C[0,1] \mid \|f - t_{\Theta(h)}\|_\infty \geqq \varepsilon\})$ decays exponentially fast towards zero for $n \to \infty$. Taking the symmetry of $\mathbb{P}_n^0$ into account, we see that the same fact is true for $\mathbb{P}_n^0(\{f \in C[0,1] \mid \|f - t_{\Theta(0)}\|_\infty + \|f - t_{-\Theta(0)}\|_\infty \geqq \varepsilon\})$, and this implies Theorem 1.1.

Proof of Proposition 4.1 The notion path is used in the sequel for any random or nonrandom, finite or infinite sequence in $\mathbb{Z}$ with increments in $\{\pm 1, \ldots, \pm r\}$, not necessarily starting in 0. Where no ambiguity is possible, we use also the symbol $(S_0, \ldots, S_n)$ for a nonrandom path.

We will clap parts of self-avoiding paths which arrive at $\lfloor \theta n \rfloor$ at time n in such a way that the clapped path runs within a box and that we can control the size of this box.

Because of (3.16) (which is, of course, valid for J_h, too), it suffices to handle only the symmetric case, i.e. we assume $h=0$. The first step is to show that the main contribution to the self-avoiding paths $(S_0, \ldots, S_n)$ satisfying $S_0=0$ and $S_n=\lfloor \theta n \rfloor$ comes from those paths which cross the two lines between 0 and 1 and $\lfloor \theta n \rfloor - 1$ and $\lfloor \theta n \rfloor$, respectively, exactly once.

Lemma 4.3 *For $\theta \in [0, r]$ and every two sequences $(\delta_n)_n$ and $(\varepsilon_n)_n$ in $\mathbb{Z}$ which are $o(n)$ it holds*

$$P_0(T > n+\delta_n, S_{n+\delta_n} = \lfloor \theta n \rfloor + \varepsilon_n) = e^{o(n)} P_0\left(\bigcup_{i=0}^{2} U(n+\delta_n+i, \lfloor \theta n \rfloor + \varepsilon_n)\right),$$

where, for $n, k \in \mathbb{N}$,

$$U(n, k) := \{T > n, \exists n_1, n_2 \in \{1, \ldots, n\}: S_1, \ldots, S_{n_1-1} < 0, \\ S_{n_1}, S_{n_1+1}, \ldots, S_{n_2} \in \{1, \ldots, k\}, S_{n_2+1}, \ldots, S_{n-1} > S_n = k\}.$$

Proof. We perform the proof only for $\delta_n = \varepsilon_n = 0$; it is clear that the insertion of the perturbations changes nothing. For a self-avoiding path $(S_0, \ldots, S_n)$ satisfying $S_0 = 0$ and $S_n = \lfloor \theta n \rfloor$ we consider the times

$$\sigma_0 = 0 < \tau_1 < \sigma_1 < \tau_2 < \sigma_2 < \ldots < \tau_k < n \quad \text{with some} \quad k \in \{1, \ldots, r-1\},$$

satisfying $S_{\tau_i - 1} \leqq 0 < S_{\tau_i}$ and $S_{\sigma_i - 1} > 0 \geqq S_{\sigma_i}$. We will construct a path $(\tilde{S}_0, \ldots, \tilde{S}_n)$ or $(\tilde{S}_0, \ldots, \tilde{S}_{n+1})$ by reorganizing the order of the $2k-1$ parts $(S_{\sigma_i}, \ldots, S_{\tau_{i+1}-1})$ $(i=0, \ldots, k-1)$ and $(S_{\tau_i}, \ldots, S_{\sigma_i - 1})$ $(i=1, \ldots, k-1)$: We delete the connecting steps between these $2k$ parts (i.e. including the last part), bring the $2k-1$ parts mentioned above into a new order (first the excursions below 0, then the excursions above 0) and insert new connecting steps. In case $k \geqq 3$ it is possible to leave the length n of the new part unchanged by choosing a suitable order, in case $k=2$ the length perhaps must be increased by one by inserting two suitable steps between $(S_{\sigma_1}, \ldots, S_{\tau_2 - 1})$ and $(S_{\tau_1}, \ldots, S_{\sigma_1 - 1})$.

An analogous procedure is performed on the line between $\lfloor \theta n \rfloor - 1$ and $\lfloor \theta n \rfloor$, so it arises a path $S^* \in \bigcup_{i=0}^{2} U(n+i, \lfloor \theta n \rfloor)$. Of course, the probabilities of $(S_0, \ldots, S_n)$ and $(S_0^*, \ldots, S_{n+i}^*)$ (with $i=0, 1, 2$, respectively) are equal on the exponential scale, and the number of paths which can be mapped onto a given $S^* \in \bigcup_{i=0}^{2} U(n+i, \lfloor \theta n \rfloor)$ is of polynomial order in n. □

For $n, k \in \mathbb{N}$ define

$$W(n, k) = \{T > n, S_1, \ldots, S_{n-1} < 0 < S_n, \quad \text{and} \quad \min_{i=1}^{n} S_i = -k\}.$$

Then the first part of a path lying in $U(n, \lfloor\theta n\rfloor)$ looks like the beginning of a path from $W(\tilde{n}, \tilde{k})$ with suitable $\tilde{n}$, $\tilde{k}$. We introduce the function

(4.4) $$R\colon \left[0, \frac{r}{2}\right] \to [-\infty, 0], \qquad R(\theta) := \overline{\lim_{n\to\infty}} \frac{1}{n} \log P_0(W(n, \lfloor\theta n\rfloor)),$$

and it will become clear along the way that this limit superior is a limit indead and unchanged by inserting perturbations of size $o(n)$ in the two arguments of W. The intuitive idea for realizing the fact that $R(\theta) \leqq \tilde{J}_0(2\theta)$ is the following: If you clap down the first part of a path being in $W(n, \lfloor\theta n\rfloor)$ (more exactly: the part between time 0 and the minimal point), then you obtain (after suitable lifting) a path lying in $V(n, \lfloor 2\theta n\rfloor + o(n))$. Note that this treatment does not change the path's probability under P_0. Since the clapped part of the path avoides the rest of the path, one can expect that this clapping-lifting map is far from being surjective, so we should get a strict estimation (Note that both R and $\tilde{J}_0$ take the value $-\infty$ on $[0, 1)$):

Lemma 4.5 *We have* $R(\theta) < \tilde{J}_0(2\theta)$ *for* $\theta \in \left[\frac{1}{2}, \frac{r}{2}\right)$.

Proof. Let $S = (S_k)_{k\in\mathbb{N}_0}$ be in $W(n, \lfloor\theta n\rfloor)$ and $k^* \in \{1, \dots, n\}$ be the index with $S_{k^*} = \min_{k=1}^{n} S_k = -\lfloor\theta n\rfloor$. We first assume that $k^* = \left\lceil\frac{n}{2}\right\rceil$ and explain the general case later. The paths $(S_0^1, \dots, S^1_{n-\lfloor n/2\rfloor}) := (S_{k^*} - S_{k^*}, S_{k^*+1} - S_{k^*}, \dots, S_n - S_{k^*})$ and $(S_0^2, \dots, S^2_{\lfloor n/2\rfloor}) := (S_{k^*} - S_{k^*}, S_{k^*-1} - S_{k^*}, \dots, S_0 - S_{k^*})$ can be seen as elements of $V\left(\left\lceil\frac{n}{2}\right\rceil + o(n), \lfloor\theta n\rfloor + o(n)\right)$ and have the same probabilities as $(S_{k^*}, \dots, S_n)$ and $(S_0, \dots, S_{k^*})$, respectively. By regarding these paths as two copies of one random walk, we will derive a variational formula for

$$R_{1/2}(\theta) := \overline{\lim_{n\to\infty}} \frac{1}{n} \log P_0\left(W(n, \lfloor\theta n\rfloor) \cap \left\{S_{\lfloor n/2\rfloor} = \min_{i=1}^{n} S_i\right\}\right)$$

which acts on $\tilde{E}^2 \times \tilde{E}^2$ rather than on $\tilde{E} \times \tilde{E}$. Let, for $n, k \in \mathbb{N}$,

$$V^{\otimes 2}(n, k) := \{((S_i^1)_i, (S_i^2)_i) \in V(n, k)^2 \mid S_i^1 \neq S_j^2 \quad \text{for } i, j = 1, \dots, n-1\}$$

$$\tilde{V}^{\otimes 2}(n, k) := \{((A_0^1, \dots, A_{k-1}^1), (A_0^2, \dots, A_{k-1}^2)) \in \tilde{V}(n, k)^2 \mid g(A_{x-1}^1, A_x^1) + g(A_{x-1}^2, A_x^2) \leqq 1 \quad \text{for } x = 1, \dots, n-1\},$$

$$M_\theta^{\otimes 2} := \{\hat{v} \in \mathcal{M}_1(\tilde{E}^2 \times \tilde{E}^2) \mid \hat{v}^{(1)} \in M_{2\theta}, \hat{v}^{(2)} \in M_{2\theta}, \sum_{\tilde{Y}\in\tilde{E}^2} \hat{v}(Y, \tilde{Y}) = \sum_{\tilde{Y}\in\tilde{E}^2} \hat{v}(\tilde{Y}, Y) \quad \text{for } Y \in \tilde{E}^2, \hat{v}((A^1, A^2), (\tilde{A}^1, \tilde{A}^2)) = 0 \quad \text{if } g(A^1, \tilde{A}^1) + g(A^2, \tilde{A}^2) \geqq 2\},$$

where

$$\hat{v}^{(1)}(A^1, \tilde{A}^1) := \sum_{A^2, \tilde{A}^2 \in \tilde{E}} \hat{v}((A^1, A^2), (\tilde{A}^1, \tilde{A}^2)),$$

$$\hat{v}^{(2)}(A^2, \tilde{A}^2) := \sum_{A^1, \tilde{A}^1 \in \tilde{E}} \hat{v}((A^1, A^2), (\tilde{A}^1, \tilde{A}^2)),$$

and

$$\hat{I}_h(\hat{v}) := \sum_{A^1, \tilde{A}^2, A^2, \tilde{A}^2 \in \tilde{E}} \hat{v}((A^1, A^2), (\tilde{A}^1, \tilde{A}^2)) \log \frac{\hat{v}((A^1, A^2), (\tilde{A}^1, \tilde{A}^2))}{\bar{v}(A^1, A^2)\, Q_h(A^1, \tilde{A}^1)\, Q_h(A^2, \tilde{A}^2)}.$$

So we can perform the same procedure as in the second section and obtain (recall definition (2.14) and compare (2.12), (2.13), (2.15), and (2.10); take a fixed $h>0$):

$$\begin{aligned} R_{1/2}(\theta) &= -2\theta h + \log Z_h + \overline{\lim_{n\to\infty}} \frac{1}{n} \log P_h \otimes P_h(V^{\otimes 2}(\lfloor n/2 \rfloor, \lfloor \theta n \rfloor)) \\ &= -2\theta h + \log Z_h + \overline{\lim_{n\to\infty}} \frac{1}{n} \log \sum_{(A^1, A^2)\in \tilde{V}^{\otimes 2}(\lfloor \frac{n}{2} \rfloor, \lfloor \theta n \rfloor)} e^{Q(A^1)+Q(A^2)} \\ &= -2\theta h + \log Z_h - \theta \inf_{\hat{v}\in M_\theta^{\otimes 2}} \hat{I}_h(\hat{v}). \end{aligned}$$

(The first two terms in the last line stem from inserting the drift and clapping the first part of the path.) So we see in particular that the limit superior in the definition of $R_{1/2}(\theta)$ is a limit. An application of Jensen's inequality to the function $x \mapsto x \log x$ shows

$$\hat{I}_h(\hat{v}) \geqq I_h(\hat{v}^{(1)}) + I_h(\hat{v}^{(2)}) \quad \text{for } \hat{v}\in M_\theta^{\otimes 2} \tag{4.6}$$

with equality if and only if

$$\hat{v}((A^1, A^2), (\tilde{A}^1, \tilde{A}^2)) = \frac{\hat{v}^{(1)}(A^1, \tilde{A}^1)}{\overline{\hat{v}^{(1)}}(A^1)} \frac{\hat{v}^{(2)}(A^2, \tilde{A}^2)}{\overline{\hat{v}^{(2)}}(A^2)} \bar{\hat{v}}(A^1, A^2) \tag{4.7}$$

for every A^1, $\tilde{A}^1$, A^2, $\tilde{A}^2 \in \tilde{E}$. Because of the continuity of $\hat{I}_h$ on the compact set $M_\theta^{\otimes 2} \cap \{\hat{v} \mid \hat{I}_h(\hat{v}) < \infty\}$ there is some $\hat{v}_\theta \in M_\theta^{\otimes 2}$ satisfying $\hat{I}_h(\hat{v}_\theta) = \inf_{\hat{v}\in M_\theta^{\otimes 2}} I_h(\hat{v})$. We will show now that either the inequality (4.6) is strict for $\hat{v}_\theta$ or that $\hat{v}_\theta^{(1)}$ or $\hat{v}_\theta^{(2)}$ are different from $v^{2\theta}$, the unique minimizer of I_h on $M_{2\theta}$. (Then Lemma 4.5 follows via (3.16) from the estimation

$$\begin{aligned} R_{1/2}(\theta) &\leqq -2\theta h + \log Z_h - \theta(I_h(\hat{v}_\theta^{(1)}) + I_h(\hat{v}_\theta^{(2)})) \\ &\leqq -2\theta h + \log Z_h + \tilde{J}_h(2\theta) \\ &= \tilde{J}_0(2\theta).) \end{aligned}$$

For every $\hat{v}\in M_\theta^{\otimes 2}$ it holds

$$0 = \sum_{A^1, \tilde{A}^1, A^2, \tilde{A}^2} \hat{v}((A^1, A^2), (\tilde{A}^1, \tilde{A}^2))\, g(A^1, \tilde{A}^1)\, g(A^2, \tilde{A}^2). \tag{4.8}$$

If $v^{2\theta} = \hat{v}_\theta^{(1)} = \hat{v}_\theta^{(2)}$ and if we would have equality in (4.6) for $\hat{v}_\theta$, then, by (4.7) and (4.8), we would get

$$0 = \sum_{A^1, A^2} \left(\sum_{\tilde{A}^1} g(A^1, \tilde{A}^1) \frac{v^{2\theta}(A^1, \tilde{A}^1)}{\bar{v}^{2\theta}(A^1)} \right) \left(\sum_{\tilde{A}^2} g(A^2, \tilde{A}^2) \frac{v^{2\theta}(A^2, \tilde{A}^2)}{\bar{v}^{2\theta}(A^2)} \right) \bar{\hat{v}}_\theta(A^1, A^2).$$

But, by Lemma 3.4, the two terms in the brackets are positive for every A^1 and A^2, respectively, which is a contradiction.

Now we remove the assumption $k^*=\left[\frac{n}{2}\right]$ in the beginning of the proof. In the same way as above, one derives a variational formula for

$$R_c(\theta):=\lim_{n\to\infty}\frac{1}{n}\log P_0\left(W(n,\lfloor\theta n\rfloor)\cap\left\{S_{\lfloor cn\rfloor}=\min_{i=1}^{n} S_i\right\}\right)$$

for $c\in[\theta/r, 1-\theta/r]$. This limit is unchanged under perturbations of size $o(n)$ in $\lfloor cn\rfloor$ and $\lfloor\theta n\rfloor$ as can be shown by adapting the proof of (2.19). So it follows, by estimating the arising variational formula in the same way as above,

$$R(\theta)\leqq\sup_{c\in[\theta/r,1-\theta/r]} R_c(\theta)<\max_{c\in[\theta/r,1-\theta/r]}\left(c\tilde{J}_0\left(\frac{\theta}{c}\right)+(1-c)\tilde{J}_0\left(\frac{\theta}{1-c}\right)\right),$$

which is not bigger than $\tilde{J}_0(2\theta)$ by concavity. Lemma 4.5 is proved now. □

We finish now the proof of Proposition 4.1.

Let two sequences $(\delta_n)_n$ and $(\varepsilon_n)_n$ in $\mathbb{Z}$ be given which are $o(n)$ and let $(S_k)_{k\in\mathbb{N}_0}$ be a path in $U(n+\delta_n, \lfloor\theta n\rfloor+\varepsilon_n)$. Then, up to lifting and lengthening by a suited infinite path, the parts $(S_0, \ldots, S_{k_1})$ and $(-S_{k_2}, \ldots, -S_{n+\delta_n})$ are elements of $W(k_1, \lfloor\theta_1^{(n)}k_1\rfloor)$ and $W(k_2, \lfloor\theta_2^{(n)}k_2\rfloor)$, respectively, and the part $(S_{k_1}, \ldots, S_{k_2})$ is an element of $V(k_2-k_1, \lfloor\theta n\rfloor+o(n))$ with suitable k_1, $k_2\in\{1, \ldots, n+\delta_n\}$ and $\theta_1^{(n)}, \theta_2^{(n)}\in\left[0, \frac{r}{2}\right]$. By passing to a subsequence, we may assume that $k_i=\lfloor c_i n\rfloor+o(n)$ $(i=1, 2)$ for suitable c_1, $c_2\in[0, 1]$ with $c_1+c_2\leqq 1-\theta/r$ and $\theta_i^{(n)}\to\theta_i\in\left[0, \frac{r}{2}\right]$ $(i=1, 2)$. Using the invariance of the limits (4.4) and (1.3) under $o(n)$-perturbations, these considerations imply that

$$\begin{aligned}&\overline{\lim_{n\to\infty}}\frac{1}{n}\log P_0(U(n+\delta_n, \lfloor\theta n\rfloor+\varepsilon_n))\\ &\quad\leqq\max\left\{c_1 R\left(\frac{\theta_1}{c_1}\right)+(1-c_1-c_2)\tilde{J}_0\left(\frac{\theta}{1-c_1-c_2}\right)+c_2 R\left(\frac{\theta_2}{c_2}\right)\right|\\ &\qquad c_1, c_2\in[0,1], c_1+c_2\leqq 1-\theta/r, \theta_i\in[c_i, c_i r/2]\ (i=1,2)\Big\},\end{aligned}$$

where we define $0R\left(\frac{0}{0}\right)$ to be zero. The r.h.s. of the last display is not bigger than $\underline{\lim}_{n\to\infty}\frac{1}{n}\log P_0(U(n+\delta_n, \lfloor\theta n\rfloor+\varepsilon_n))$, since this maximum ranges over the exponential rates of the P_0-probability of certain path classes which are included in $U(n+\delta_n, \lfloor\theta n\rfloor+\varepsilon_n)$. Using Lemma 4.3, the existence of the limit (1.2) and its invariance under perturbations of size $o(n)$ are proved and its value is shown to be the r.h.s. of the last display. The strict estimate in Proposition 4.1 follows from this using Lemma 4.5 and the strict concavity of $\tilde{J}_0$. □

Acknowledgement. I am grateful to Prof. Dr. E. Bolthausen, who brought this problem to my attention and gave many helpful advices. Furthermore, I would like to thank a referee and the managing editor for their suggestions.

References

1. Aldous, D.J.: Self-intersections of 1-dimensional random walks. Probab. Theory Relat. Fields **72**, 559–587 (1986)
2. Bolthausen, E.: On self-repellent one-dimensional random walks. Probab. Theory Relat. Fields **86**, 423–441 (1990)
3. Deuschel, J.-D., Stroock, D.W.: Large deviations. Boston: Academic Press 1989
4. Ellis, R.: Entropy, large deviations, and statistical mechanics. Grundlehren der mathematischen Wissenschaften vol. 271. New York: Springer 1984
5. Freed, K.F.: Polymers as self-avoiding walks. Ann. Appl. Probab. **9**, 537–556 (1981)
6. Greven, A., den Hollander, F.: A variational characterization of the speed of a one-dimensional self-repellent random walk. Ann. Appl. Probab. (to appear)
7. König, W.: The drift of a one-dimensional self-repellent random walk with bounded increments (in preparation)
8. Kusuoka, S.: Asymptotics of the polymer measure in one dimension. Infinite dimensional analysis and stochastic processes, Sem. Meet. Bielefeld 1983. Albeverio, S. (ed.). Res. Notes Math. vol. 124, pp. 66–82. Pitman Advanced Publishing Program. Boston: Pitman 1985
9. Seneta, E.: Non-negative matrices and Markov chains. New York: Springer 1981
10. Varadhan, S.R.S.: Asymptotic probabilities and differential equations. Commun. Pure Appl. Math. **XIX**, 261–286 (1966)
11. Westwater, J.: On Edward's model for polymer chains. Trends and developments in the eighties. Albeverio, S., Blanchard, P. (eds.). Bielefeld encounters in Math./Phys. Singapore: World Scientific 1984