

# The Pattern of Escape from Metastability of a Stochastic Ising Model\*

Roberto H. Schonmann

Mathematics Department, University of California at Los Angeles, Los Angeles, CA 90024, USA

Received March 27, 1991; in revised form December 16, 1991

**Abstract.** We study further the metastable behavior of Metropolis dynamics for the two-dimensional nearest neighbor ferromagnetic Ising model, with positive and small external field, in the limit as the temperature vanishes (see [NS]). We focus on the typical features of the escape (nucleation) from the (metastable) configuration with all spins  $-1$ , to the (stable) configuration with all spins  $+1$ . Using the reversibility of the process as the main tool, we prove (for the discrete time version of the model) that the first step of a typical escaping path is the time reverse of a typical time evolution of a shrinking subcritical rectangular droplet, which is one slice smaller than a critical droplet. This subcritical droplet then evolves in a time of order 1 to a critical droplet, which finally grows with features described in [NS].

## I. The Model

In this paper we consider a discrete time version of Metropolis dynamics for the two-dimensional nearest neighbor ferromagnetic Ising model. The state space of the process is the set  $\{-1, +1\}^{A_N}$ , where  $A_N = \{1, \dots, N\}^2$  is taken with periodic boundary conditions (meaning that two sites  $x, y \in A_N$  will be said to be neighbors if they have one coordinate in common and the other one differing by 1 or  $N-1$ ).  $\sigma(x) = \pm 1$  is called the spin at the site  $x$ . To each configuration  $\sigma \in \{-1, +1\}^{A_N}$  an energy is assigned by the expression

$$H(\sigma) = -\frac{1}{2} \sum_{x,y \in \mathbb{B}_N} \sigma(x)\sigma(y) - \frac{h}{2} \sum_{z \in A_N} \sigma(z),$$

where  $\mathbb{B}_N$  is the set of (unordered) pairs of neighbors in  $A_N$  and  $h$  is called the external field. The discrete time Metropolis dynamics at inverse temperature  $\beta > 0$  is the Markov chain defined by taking uniformly at random at each integer unit of time a site  $x \in A_N$  and then flipping the spin there with probability

\* Work partially supported by the Brazilian CNPq and by the American NSF, under grant DMS91-00725

$\exp(-\beta(\Delta_x H(\sigma))^+)$ , where  $\sigma$  is the present configuration,  $\Delta_x H(\sigma) = H(\sigma^x) - H(\sigma)$ , with  $\sigma^x(y) = \sigma(y)$  for  $y \neq x$  and  $\sigma^x(x) = -\sigma(x)$ , and  $(\cdot)^+ = (\cdot) \vee 0$  is the positive part function. The corresponding transition probabilities when  $\sigma \neq \eta$  are given by

$$p(\sigma, \eta) = \begin{cases} N^{-2} \exp(-\beta(\Delta_x H(\sigma))^+), & \text{if } \eta = \sigma^x \text{ for some } x \in \Lambda_N, \\ 0, & \text{otherwise.} \end{cases}$$

We will denote by  $(\sigma_t^\eta)_{t=0,1,2,\dots}$  the corresponding Markov chain started at time 0 from the configuration  $\eta$ . Let  $\mu$  be the Gibbs probability measure on  $\{-1, +1\}^{\Lambda_N}$  at inverse temperature  $\beta$  given by

$$\mu(\sigma) = Z^{-1} \exp(-\beta H(\sigma)),$$

where  $Z^{-1}$  is a normalization constant. Metropolis dynamics is reversible with respect to  $\mu$ , in the sense that for all  $\sigma, \eta \in \{-1, +1\}^{\Lambda_N}$ ,

$$\mu(\sigma)p(\sigma, \eta) = \mu(\eta)p(\eta, \sigma). \quad (\text{I.1})$$

In particular  $\mu$  is the unique invariant probability measure for this process.

We will denote the configuration with all spins  $-1$  (resp.  $+1$ ) by  $\underline{-1}$  (resp.  $\underline{+1}$ ).

In [NS] we considered a continuous time version of this process, which can for instance be obtained from the discrete time version by considering a Poisson process with rate  $N^2$  and then at each occurrence time of this process choosing one of the  $N^2$  sites at random and proceeding as above to decide to flip or not the spin at this site. The law of large numbers for the Poisson process allows one then to translate easily results from the continuous time version into the discrete time version and back. In particular all the results in [NS] can be verified to hold for the discrete time version (but one can also adapt the proofs from [NS] directly to discrete time). We switched in this paper to the discrete time version to avoid technicalities, but one can translate our results back to continuous time if desired.

## II. Results

In [NS] we studied the metastable behavior of the continuous time version of this process with  $0 < h < 4$  and large arbitrary  $N$  started from  $\underline{-1}$ , in the limit in which  $\beta \rightarrow \infty$ . Roughly speaking we proved that the system stays close to  $\underline{-1}$  for a long time, until it eventually goes to  $\underline{+1}$  at an unpredictable time. In this paper we will look in detail to the first excursion from  $\underline{-1}$  to  $\underline{+1}$ , i.e., to the way the system eventually goes from  $\underline{-1}$  to  $\underline{+1}$ . For the motivation of this line of research and for more background the reader is referred to [NS] and to the expository paper [Sch], as well as to papers quoted therein. The idea of characterizing metastability by considering the typical paths of the process was introduced in the paper [CGOV], and further studied in a series of papers quoted in [NS] and [Sch]. The first passage from  $\underline{-1}$  to  $\underline{+1}$  may be thought of as the escape from an energy well. In this sense this problem is related to the Freidlin–Wentzell problem of the escape from the domain of an attractor in the context of weakly perturbed dynamical system (see [FW], pp. 115–117). But in our case the abundance of local minima of the energy (configurations with rectangular droplets of spins  $+1$  in a sea of spins  $-1$ ) produces an interesting pattern for the typical escape route. Our analysis is

based on elementary methods and relies mostly on the reversibility of the process; we do not use the Freidlin–Wentzell machinery. It is even possible that the methods used here may be also used to provide an alternative derivation of similar results for weakly perturbed dynamical systems. We observe also that even if we consider our process with large  $\beta$  as a perturbation of a corresponding reference process with  $\beta = +\infty$ , this reference system is not completely deterministic, as opposed to what occurs in the Freidlin–Wentzell setting.

Now we recall some of the results proven in [NS] in continuous time, but which hold also in discrete time, as discussed in the previous section. To concentrate on the most interesting cases we will always assume

$$0 < h < 2, 2/h \text{ is not an integer, and } N > ((2/h) + 1)^2 + 1. \tag{II.1}$$

*Remark.* This is essentially what was called the “standard case” in [NS].

$\mathcal{R}$  will denote the set of configurations in  $\{-1, +1\}^{A_N}$  with all spins  $-1$  except for those in a rectangle  $l_1 \times l_2$  which are  $+1$ . With  $l_1$  and  $l_2$  less than  $N - 1$ . For  $\eta \in \mathcal{R}$  define  $l_1(\eta)$  and  $l_2(\eta)$  as the  $l_1$  and  $l_2$  above and  $l(\eta) = \min\{l_1(\eta), l_2(\eta)\}$ .

$$T^n(A) = \inf\{t \geq 0: \sigma_t^n \in A\}$$

will denote hitting times. We use the abbreviation  $T^n(\xi)$  for  $T^n(\{\xi\})$ . Sometimes we omit  $\eta$  when it is  $-\underline{1}$  and  $\xi$  when it is  $+\underline{1}$ ; so

$$T = \inf\{t \geq 0: \sigma_t^{-1} = +\underline{1}\}.$$

Theorem 1 in [NS] states that small rectangular droplets of  $+1$  tend to shrink while large rectangular droplets tend to grow. Small or large being decided by the comparison between  $l(\eta)$  and  $2/h$ . Many details about the way the droplet shrinks or grows can also be obtained from the analysis done in [NS]. We repeat now in detail the aspects already explicitly stated in [NS], because they will be important later on. First we define for each  $\eta \in \mathcal{R}$  and  $\varepsilon > 0$  the events

$$S_\varepsilon(\eta) = \{T^n(-\underline{1}) < T^n(+\underline{1}); \text{ from time } 0 \text{ up to time } T^n(-\underline{1}) \\ \text{all spins which were } -1 \text{ at time } 0 \text{ are still } -1 \text{ and} \\ \text{all spins } +1 \text{ form a single cluster; also} \\ \exp(\beta((l(\eta) - 1)h - \varepsilon)) < T^n(-\underline{1}) < \exp(\beta((l(\eta) - 1)h + \varepsilon))\}.$$

$$G_\varepsilon(\eta) = \{T^n(+\underline{1}) < T^n(-\underline{1}); \text{ from time } 0 \text{ up to time } T^n(+\underline{1}) \\ \text{the number of spins } +1 \text{ is at least } l_1(\eta)l_2(\eta) - l(\eta) + 1 \\ \text{and they form a single cluster with the property that the} \\ \text{smallest rectangle which contains all spins } +1 \text{ contains} \\ \text{the rectangle where the original spins } +1 \text{ were; also} \\ \exp(\beta((2 - h) - \varepsilon)) < T^n(+\underline{1}) < \exp(\beta((2 - h) + \varepsilon))\}.$$

In [NS] it was proven that for  $0 \leq h < 2$  and  $\eta \in \mathcal{R}$ ,

$$\text{If } l(\eta) < 2/h \text{ and } \varepsilon > 0 \text{ then } \lim_{\beta \rightarrow \infty} P(S_\varepsilon(\eta)) = 1, \tag{II.2}$$

$$\text{If } l(\eta) > 2/h \text{ and } \varepsilon > 0 \text{ then } \lim_{\beta \rightarrow \infty} P(G_\varepsilon(\eta)) = 1. \tag{II.3}$$

Once the typical behavior of rectangular droplets was established in [NS], these results were used to show the metastable character of the evolution starting

from  $-\underline{1}$  and described in Theorem 2 there (omitted here since it is not used below). Finally some details of the typical way the system escapes from  $-\underline{1}$  to  $+\underline{1}$  were obtained in Theorem 3 in [NS]. Our goal here is to strengthen parts (c) and (d) of this theorem by describing the law of the first excursion from  $-\underline{1}$  to  $+\underline{1}$  and in particular proving the statements in the abstract of this paper.

Next we recall more definitions from [NS] and introduce new ones.  $L = \lceil 2/h \rceil =$  (smallest integer not smaller than  $2/h$ ).  $L$  is the side of the critical droplet. For  $k = 0, 1, \dots, L$ ,  $\mathcal{P}^{(k)}$  will denote the set of configurations with all spins  $-1$  except for a droplet of spins  $+1$  with the following shape: the union of a rectangle of sides  $L$  and  $L - 1$  with a disjoint rectangle of sides  $1$  and  $k$  (the empty set if  $k = 0$ ) such that all the sites in the latter rectangle are neighbors of a site in the former one, and in case  $k \neq 0$  the smallest rectangle which contains the droplet of spins  $+1$  is a square  $L \times L$ . (Think of the smaller rectangle as a protuberance.)  $\mathcal{P}^{(1)}$  is the set  $\mathcal{P}$  of [NS] (protocritical droplets) and  $\mathcal{P}^{(L)}$  is the set  $\mathcal{G}$  of [NS] (critical droplets).

Let  $\mathcal{S}$  be the set of finite sequences of configurations (including  $\emptyset$ ),

$$\mathcal{S} = \bigcup_{k=0}^{\infty} (\{-1, +1\}^A)^k.$$

Define the random time

$$\Theta = \sup\{t \in [0, T]: \sigma_t^{-1} = -\underline{1}\}.$$

The path of the first excursion from  $-\underline{1}$  to  $+\underline{1}$  is

$$S = (\sigma_{\Theta}^{-1}, \sigma_{\Theta+1}^{-1}, \dots, \sigma_{T-1}^{-1}, \sigma_T^{-1}),$$

which is a random element of  $\mathcal{S}$ . We are concerned with the law of  $S$ , i.e., with  $P(S \in \mathcal{V})$  for each  $\mathcal{V} \subseteq \mathcal{S}$ . To state our result about this law we need more notation. Given  $s = (s_1, \dots, s_k) \in \mathcal{S}$  we define  $K(s) = k$  ( $K(\emptyset) = 0$ ),  $\alpha(s) = s_1$  and  $\zeta(s) = s_k$  ( $\alpha(\emptyset)$  and  $\zeta(\emptyset)$  are not defined).  $R(s) = (s_k, s_{k-1}, \dots, s_2, s_1)$  will denote the time reversal of  $s$  ( $R(\emptyset) = \emptyset$ ). For  $\mathcal{V} \subseteq \mathcal{S}$ ,  $R(\mathcal{V}) = \{R(s): s \in \mathcal{V}\}$ . Given also  $u = (u_1, \dots, u_l) \in \mathcal{S}$  we write  $s * u = (s_1, \dots, s_k, u_1, \dots, u_l) \in \mathcal{S}$  ( $s * \emptyset = \emptyset * s = s$ ). Given  $s, u \in \mathcal{S}$ , we say that “ $u$  is part of  $s$ ” iff there are  $v, w \in \mathcal{S}$  such that  $s = v * u * w$  (observe that  $v$  and  $w$  may be  $\emptyset$ ). Given  $s, u \in \mathcal{S}$  such that  $\zeta(s) = \alpha(u)$  we say that “ $u$  may follow  $s$ ” and define  $s \oplus u = (s_1, s_2, \dots, s_{K(s)-1}, s_{K(s)}, u_2, \dots, u_{K(u)})$ .

Our approach for obtaining information about  $S$  will be based in part on the observation that in its way from  $-\underline{1}$  to  $+\underline{1}$  the system is likely to visit  $\mathcal{P}^{(0)}$ , jump from there to  $\mathcal{P} = \mathcal{P}^{(1)}$  and from there to  $\mathcal{P}^{(2)}$ . With this in mind we define

$$\mathcal{S}_2 = \{s \in \mathcal{S}: \alpha(s) \in \mathcal{P}^{(0)}, s_i \in \mathcal{P} \text{ for } i = 2, \dots, K(s) - 1, \zeta(s) \in \mathcal{P}^{(2)}\},$$

$$\mathcal{S}_1 = \{s \in \mathcal{S}: \alpha(s) = -\underline{1}, \zeta(s) \in \mathcal{P}^{(0)}, s_i \notin \{-\underline{1}, +\underline{1}\} \text{ for } i = 2, \dots, K(s) - 1, \text{ there is no } u \in \mathcal{S}_2 \text{ which is a part of } s\},$$

$$\mathcal{S}_3 = \{s \in \mathcal{S}: \alpha(s) \in \mathcal{P}^{(2)}, \zeta(s) = +\underline{1}, s_i \notin \{-\underline{1}, +\underline{1}\} \text{ for } i = 2, \dots, K(s) - 1\},$$

$$\mathcal{S}_0 = \{s \in \mathcal{S}: \text{there exists } u \in \mathcal{S}_1, v \in \mathcal{S}_2 \text{ and } w \in \mathcal{S}_3 \text{ such that } v \text{ may follow } u, w \text{ may follow } v \text{ and } s = u \oplus v \oplus w\}.$$

Observe that the decomposition above of  $s \in \mathcal{S}_0$  into  $s = u \oplus v \oplus w$  is unique. We write in this case  $s = [u, v, w]$ . Define now the events

$$\begin{aligned} \{(\sigma_t^\eta) \text{ starts as } s\} &= \{(\sigma_t^\eta: t = 0, \dots, K(s) - 1) = s\}, \\ \{(\sigma_t^\eta) \text{ starts in } \mathcal{V}'\} &= \bigcup_{s \in \mathcal{V}'} \{(\sigma_t^\eta) \text{ starts as } s\}. \end{aligned}$$

We define now various probability measures on  $\mathcal{S}$ :

$$v_1(s) = \begin{cases} |\mathcal{P}^{(0)}|^{-1} P((\sigma_t^{\xi(s)}) \text{ starts as } R(s) | (\sigma_t^{\xi(s)}) \text{ starts in } R(\mathcal{S}_1)), & \text{if } s \in \mathcal{S}_1, \\ 0, & \text{otherwise.} \end{cases}$$

For  $\eta \in \mathcal{P}^{(0)}$ ,

$$v_2^\eta(s) = \begin{cases} P((\sigma_t^\eta) \text{ starts as } s | (\sigma_t^\eta) \text{ starts in } \mathcal{S}_2), & \text{if } s \in \mathcal{S}_2, \\ 0, & \text{otherwise.} \end{cases}$$

For  $\xi \in \mathcal{P}^{(2)}$ ,

$$v_3^\xi(s) = \begin{cases} P((\sigma_t^\xi) \text{ starts as } s | (\sigma_t^\eta) \text{ starts in } \mathcal{S}_3), & \text{if } s \in \mathcal{S}_3, \\ 0, & \text{otherwise.} \end{cases}$$

$$v(s) = \begin{cases} v_1(u) \cdot v_2^{\xi(u)}(v) \cdot v_3^{\xi(v)}(w), & \text{if } s \in \mathcal{S}_0 \text{ and } s = [u, v, w], \\ 0, & \text{otherwise.} \end{cases}$$

As  $\beta \rightarrow \infty$ , (II.2) shows that the conditioning in the definition of  $v_1$  becomes irrelevant. So for large  $\beta$ ,  $v_1$  may be thought of as the distribution of the time reverse of the path along which a rectangular droplet of sides  $L$  and  $L - 1$  taken uniformly from all the possible ones (configurations in  $\mathcal{P}^{(0)}$ ) shrinks.

To be able to say something similar about  $v_3^\xi$ , we have first to observe that if  $\xi \in \mathcal{P}^{(2)}$  the initial evolution of  $(\sigma_t^\xi)$  is very predictable. Let  $A_\varepsilon(\xi)$  be the event  $\{(\sigma_t^\xi)$  jumps from  $\mathcal{P}^{(2)}$  to  $\mathcal{P}^{(3)}$ , then to  $\mathcal{P}^{(4)}$ , and so on, until it hits  $\mathcal{P}^{(L)} = \mathcal{G}$  at a time  $t < e^{\beta\varepsilon}\}$ . Then one easily sees that for all  $\xi \in \mathcal{P}^{(2)}$  and  $\varepsilon > 0$ ,

$$\lim_{\beta \rightarrow \infty} P(A_\varepsilon(\xi)) = 1. \tag{II.4}$$

From (II.4) and (II.3) we see that as  $\beta \rightarrow \infty$  the conditioning in the definition of  $v_3^\xi$  becomes irrelevant and for large  $\beta$  this law may be thought of as the distribution of the path along which a droplet in a configuration  $\xi$  of  $\mathcal{P}^{(2)}$  grows.

In contrast, the conditioning in the definition of  $v_2^\eta$  becomes crucial as  $\beta \rightarrow \infty$ . But the basic asymptotic properties of this measure can be easily obtained; for instance the fact that it concentrates its mass on paths in  $\mathcal{S}_2$  with lengths which are bounded above by (for example)  $\exp(\beta\varepsilon)$  as  $\beta \rightarrow \infty$ .

Our main result in this paper is

**Theorem.** *Under condition (I.1)*

$$\lim_{\beta \rightarrow \infty} \sup_{\mathcal{V}' \subset \mathcal{S}} |P(S \in \mathcal{V}') - v(\mathcal{V}')| = 0,$$

i.e., the total variation distance between  $P(S \in \cdot)$  and  $v$  vanishes as  $\beta \rightarrow \infty$ .

*Remark.* As observed in [NS] such a result implies that one can also take a double limit, with  $N \rightarrow \infty$  as  $\beta \rightarrow \infty$  but slowly enough so that the same result holds for this limit.

Before going to the proof of the Theorem, we comment on its meaning and some of its consequences. Essentially it says that for large  $\beta$  the passage from  $-\underline{1}$  to  $+\underline{1}$  is likely to occur by the formation of a rectangular droplet of sides  $L$  and  $L - 1$  via the time reverse of the way such a droplet would shrink (described to some extent by (II.2)). This is followed by the growing on such a droplet of a protuberance first of one spin, then of a second neighboring spin, and so on, until a critical droplet  $L \times L$  is formed, which then grows to cover the whole system (with features described in part by (II.3)).

As corollaries to the Theorem, we can prove various statements about the typical  $S$ . For instance (details of the proofs left to the reader)

$$\lim_{\beta \rightarrow \infty} P(\text{In all configurations in } S \text{ the spins } +1 \text{ form a single cluster}) = 1 .$$

Or, say that the first passage from  $-\underline{1}$  to  $+\underline{1}$  is  $\varepsilon$ -regular if “a configuration in  $\mathcal{P}^{(0)}$  is reached in a time between  $\exp(\beta(h(L - 2) - \varepsilon))$  and  $\exp(\beta(h(L - 2) + \varepsilon))$ , after which a droplet in  $\mathcal{G}$  is reached in a time shorter than  $\exp(\beta\varepsilon)$ , after which  $+\underline{1}$  is reached in a time between  $\exp(\beta(2 - h - \varepsilon))$  and  $\exp(\beta(2 - h + \varepsilon))$ .” Then

$$\lim_{\beta \rightarrow \infty} P(S \text{ is } \varepsilon\text{-regular}) = 1 .$$

Observe that  $h(L - 2) < h((2/h + 1) - 2) = 2 - h$ , so that the time to form the critical droplet is much shorter than the time it needs to grow. Recall that in Theorem 3 (parts (a) and (b)) of [NS] we proved that for the total time  $T$  (from  $t = 0$  until  $+\underline{1}$  is reached) we have as  $\beta \rightarrow \infty$

$$\begin{aligned} \beta^{-1} \log T &\rightarrow \Gamma(h) \text{ in probability ,} \\ \beta^{-1} \log ET &\rightarrow \Gamma(h) , \end{aligned}$$

where  $\Gamma(h) = H(\eta) - H(-\underline{1})$  for  $\eta \in \mathcal{P}$ , i.e.,  $\Gamma(h) = 4L - (L^2 - L + 1)h > 2 - h$ . So the first excursion from  $-\underline{1}$  to  $+\underline{1}$  takes much less time than the waiting time for it to happen.

### III. Proofs

Let  $S_0$  be the first element of  $\mathcal{S}_0$  which is part of  $(\sigma_t^\eta: t = 0, \dots, t')$  for some  $t'$ . (By this we mean that  $S_0$  is an element of  $\mathcal{S}_0$  which is part of  $(\sigma_t^\eta: t = 0, \dots, t')$  for some  $t'$ , and such that for any  $t'' < t'$  there is no element of  $\mathcal{S}_0$  which is part of  $(\sigma_t^\eta: t = 0, \dots, t'')$ ; with probability one  $S_0$  is well defined.) In this section we will prove

**Proposition 1.** *The law of  $S_0$  is  $\nu$ , i.e., for all  $\mathcal{V} \subset \mathcal{S}$*

$$P(S_0 \in \mathcal{V}) = \nu(\mathcal{V}) .$$

**Proposition 2.** *Under condition (II.1)*

$$\lim_{\beta \rightarrow \infty} P(S_0 \neq S) = 0 .$$

The two propositions above clearly imply the Theorem, since

$$|P(S \in \mathcal{V}) - P(S_0 \in \mathcal{V})| \leq 2P(S_0 \neq S) .$$

Before we prove Proposition 1, we will prove a similar fact in a more general setting (Proposition 3 below). This will illustrate the main argument that we will then use to prove Proposition 1. This result will also have applications for Metropolis dynamics, as discussed after its proof. And will be used in the proof of Proposition 2.

Let  $(X_t)_{t=0,1,\dots}$  be an irreducible reversible Markov chain on a finite state space  $\mathcal{X}$ , having transition probabilities  $p(\cdot, \cdot)$ . Single out two states,  $\eta, \xi \in \mathcal{X}$ . Let  $\mathcal{S}$  be as before the set of finite sequences of elements of  $\mathcal{X}$  and set

$$\mathcal{U} = \{s \in \mathcal{X}: \alpha(s) = \eta, \zeta(s) = \xi, s_i \notin \{\eta, \xi\} \text{ for } i = 2, \dots, K(s) - 1\} .$$

Choose an initial state  $\gamma$  and let  $U$  (resp.  $\bar{U}$ ) be the first element of  $\mathcal{U}$  (resp.  $R(\mathcal{U})$ ) which is part of  $(X_t^\gamma: t = 0, \dots, t')$  for some  $t'$ . Observe that the laws of  $U$  and  $\bar{U}$  do not depend on the initial state  $\gamma$ . Define the probability measures  $\rho$  and  $\bar{\rho}$  on  $\mathcal{S}$  by

$$\rho(s) = \begin{cases} P((X_t^\eta) \text{ starts as } s | (X_t^\eta) \text{ starts in } \mathcal{U}), & \text{if } s \in \mathcal{U}, \\ 0, & \text{otherwise .} \end{cases}$$

$$\bar{\rho}(s) = \begin{cases} P((X_t^\xi) \text{ starts as } R(s) | (X_t^\xi) \text{ starts in } R(\mathcal{U})), & \text{if } s \in \mathcal{U}, \\ 0, & \text{otherwise .} \end{cases}$$

**Proposition 3.** *For all  $\mathcal{V} \subset \mathcal{S}$ ,*

$$P(U \in \mathcal{V}) = \rho(\mathcal{V}) = \bar{\rho}(\mathcal{V}) = P(\bar{U} \in R(\mathcal{V})) . \tag{III.1}$$

The proof of this proposition is simple and possibly known. We present it here because we did not find it in the literature, and the arguments that we use to prove it will be also used later to prove Proposition 1.

*Proof of Proposition 3.* Define the events  $E_n = \{X_n^\eta = \eta, \text{ there is no } u \in \mathcal{U} \text{ which is part of } (X_t^\eta: t = 0, \dots, n)\}$ . Now for  $s \notin \mathcal{U}$ ,  $P(U = s) = 0$  and for  $s \in \mathcal{U}$ ,

$$P(U = s) = \sum_{n=0}^{\infty} P(E_n, X_{n+1}^\eta = s_2, X_{n+2}^\eta = s_3, \dots, \\ X_{n+K(s)-2}^\eta = s_{K(s)-1}, X_{n+K(s)-1}^\eta = s_{K(s)} = \xi) \\ = P((X_t^\eta) \text{ starts as } s) \sum_{n=0}^{\infty} P(E_n) . \tag{III.2}$$

Summing over  $s \in \mathcal{U}$  gives

$$\sum_{n=0}^{\infty} P(E_n) = (P((X_t^\eta) \text{ starts in } \mathcal{U}))^{-1} . \tag{III.3}$$

The first equality in (III.1) follows from (III.2) and (III.3). The last equality in (III.1) is the same as the first one with  $\eta$  and  $\xi$  interchanged. Finally to prove the second

equality let  $\mu$  be the invariant reversible measure for  $(X_t)$  and using (I.1) write ( $k = K(s)$ )

$$\begin{aligned}
 &P((X_t^n) \text{ starts as } s) \\
 &= p(\eta, s_2)p(s_2, s_3) \dots p(s_{k-2}, s_{k-1})p(s_{k-1}, \xi) \\
 &= p(\eta, s_2)p(s_2, s_3) \dots p(s_{k-2}, s_{k-1})p(s_{k-1}, \xi)(\mu(\xi))^{-1}\mu(\xi) \\
 &= p(\eta, s_2)p(s_2, s_3) \dots p(s_{k-2}, s_{k-1})(\mu(s_{k-1}))^{-1}p(\xi, s_{k-1})\mu(\xi) \\
 &= p(\eta, s_2)p(s_2, s_3) \dots (\mu(s_{k-2}))^{-1}p(s_{k-1}, s_{k-2})p(\xi, s_{k-1})\mu(\xi) \\
 &= \dots \dots \dots \\
 &= (\mu(\eta))^{-1}p(s_2, \eta)p(s_3, s_2) \dots p(s_{k-1}, s_{k-2})p(\xi, s_{k-1})\mu(\xi) \\
 &= (\mu(\xi)/\mu(\eta)) \cdot P((X_t^\xi) \text{ starts as } R(s)).
 \end{aligned}
 \tag{III.4}$$

The second equality in (III.1) follows from (III.4). ■

Back to Metropolis dynamics, Proposition 3 tells us that the law of the first excursion of  $(\sigma_t^{-1})$  from  $-\underline{1}$  to  $+\underline{1}$  is the time reverse of the law of the first excursion (which occurs later) from  $+\underline{1}$  to  $-\underline{1}$ . This is not very useful (at this stage) to extract information about the former, but in combination with the Theorem proven in this paper, gives us in fact information about the latter. In the proof of Proposition 2 we will in fact exploit this observation.

*Proof of Proposition 1.* Define the events  $F_n = \{\sigma_n^{-1} = -\underline{1}, \text{ there is no } u \in \mathcal{S}_0 \text{ which is part of } (\sigma_t^{-1}: t = 0, \dots, n)\}$ . Now for  $s \notin \mathcal{S}_0$ ,  $P(S_0 = s) = 0$  and for  $s \in \mathcal{S}_0$ , proceeding as in (III.2) and (III.3), with  $F_n$  replacing  $E_n$ , we obtain

$$P(S_0 = s) = \frac{P(\sigma^{-1} \text{ starts as } s)}{P(\sigma_t^{-1} \text{ starts in } \mathcal{S}_0)}.
 \tag{III.5}$$

As mentioned after the definition of  $\mathcal{S}_0$ , each  $s \in \mathcal{S}_0$  can be written in a unique form as  $s = u \oplus v \oplus w = [u, v, w]$ , with  $u \in \mathcal{S}_1, v \in \mathcal{S}_2$  and  $w \in \mathcal{S}_3$ . For the random  $S_0$ , let

$$S_0 = [U, V, W]$$

be the corresponding decomposition. Equation (III.5) may now be rewritten as

$$\begin{aligned}
 &P(U = u, V = v, W = w) \\
 &= \frac{P((\sigma_t^{-1}) \text{ starts as } u) \cdot P((\sigma_t^{\zeta(u)}) \text{ starts as } v) \cdot P((\sigma_t^{\zeta(v)}) \text{ starts as } w)}{\sum_{\substack{u' \in \mathcal{S}_1 \\ v' \in \mathcal{S}_2 \\ w' \in \mathcal{S}_3}} P((\sigma_t^{-1}) \text{ starts as } u') \cdot P((\sigma_t^{\zeta(u')}) \text{ starts as } v') \cdot P((\sigma_t^{\zeta(v')}) \text{ starts as } w')},
 \end{aligned}
 \tag{III.6}$$

for  $u \in \mathcal{S}_1, v \in \mathcal{S}_2$  and  $w \in \mathcal{S}_3$ . By symmetry  $\sum_{w' \in \mathcal{S}_3} P((\sigma_t^{\zeta(v')}) \text{ starts as } w')$  does not depend on  $v' \in \mathcal{S}_2$ , and  $\sum_{v' \in \mathcal{S}_2} P((\sigma_t^{\zeta(u')}) \text{ starts as } v')$  does not depend on  $u' \in \mathcal{S}_1$ .

Therefore the r.h.s. of (III.6) can be written as the product of the following three factors:

$$P(W = w | U = u, V = v) = \frac{P((\sigma_t^{\zeta(v)}) \text{ starts as } w)}{P((\sigma_t^{\zeta(v)}) \text{ starts in } \mathcal{S}_3)} = v_3^{\zeta(v)}(w), \quad \text{(III.7a)}$$

$$P(V = v | U = u) = \frac{P((\sigma_t^{\zeta(u)}) \text{ starts as } v)}{P((\sigma_t^{\zeta(u)}) \text{ starts in } \mathcal{S}_2)} = v_2^{\zeta(u)}(v), \quad \text{(III.7b)}$$

$$P(U = u) = \frac{P((\sigma_t^{-1}) \text{ starts as } u)}{\sum_{u' \in \mathcal{S}_1} P((\sigma_t^{-1}) \text{ starts as } u')}. \quad \text{(III.7c)}$$

To finish the proof now it suffices to show that the r.h.s. of (III.7c) is equal to  $v_1(u)$ . This follows from a computation analogous to (III.4), using (I.1). The factor  $|\mathcal{P}^{(0)}|^{-1}$  appears here because the sum in the denominator can be partitioned according to what  $\zeta(u')$  is, and doing so one obtains  $|\mathcal{P}^{(0)}|$  terms identical to

$$(\mu(\zeta(u))/\mu(-1)) \cdot P((\sigma_t^{\zeta(u)}) \text{ starts in } R(\mathcal{S}_1)). \quad \blacksquare$$

*Proof of Proposition 2.* Theorem 3 in [NS] assures that the system is likely to visit  $\mathcal{P}$  before it hits  $+1$ ,

$$\lim_{\beta \rightarrow \infty} P(T(\mathcal{P}) < T) = 1. \quad \text{(III.8)}$$

But this does not exclude the possibility that after visiting  $\mathcal{P}$  it returns to  $-1$  and eventually goes to  $+1$  without visiting  $\mathcal{P}$  again. To strengthen the result we can make rigorous an idea mentioned at the end of [NS]. Divide the path  $(\sigma_t^{-1}; t = 0, 1, \dots)$  into cycles. The  $n^{\text{th}}$  cycle starts when the  $(n - 1)^{\text{th}}$  cycle was completed and it will be completed when the system hits  $-1$  for the first time after having visited  $\mathcal{P}$  during this cycle. Say that a cycle is “acceptable” if  $+1$  is not hit during this cycle. Say that a cycle is “good” if  $+1$  is not hit before of  $\mathcal{P}$  in this cycle and the system after first hitting  $\mathcal{P}$  during this cycle has as first change of state a transition to  $\mathcal{P}^{(2)}$  which is followed by a sequence in  $\mathcal{S}_3$  leading to  $+1$ . Say that a cycle is “bad” if it is neither acceptable nor good.

When the system is in a configuration in  $\mathcal{P}$ , it jumps to configurations in  $\mathcal{P}^{(0)}$  or  $\mathcal{P}^{(2)}$  with probabilities which do not vanish as  $\beta \rightarrow \infty$ , while the probabilities of all other jumps go to zero as  $\beta \rightarrow \infty$  (since they cause increase in energy). Consider the time of the first entrance in  $\mathcal{P}$  in each cycle as a stopping time. The strong Markov property assures then that the probabilities of first moving to  $\mathcal{P}^{(0)}$ , or first moving to  $\mathcal{P}^{(2)}$ , after this stopping time are both bounded away from 0 as  $\beta \rightarrow \infty$ , while all other possibilities have probabilities which vanish as  $\beta \rightarrow \infty$ . Now consider the time of this first change of state after the first visit to  $\mathcal{P}$  in a cycle also as a stopping time. From (III.8), (II.2), (II.3) and (II.4) it follows that the probabilities that a cycle is acceptable or good are both bounded away from 0 as  $\beta \rightarrow \infty$  while the probability that it is bad goes to zero as  $\beta \rightarrow \infty$ . Therefore

$$\lim_{\beta \rightarrow \infty} P(\Omega_1) = 1,$$

where

$$\Omega_1 = \{\text{a good cycle occurs before any bad one}\}.$$

On  $\Omega_1$   $S$  has almost all the features required to guaranty that  $S = S_0$ , but we still have to prove that in the sequence  $S$   $\mathcal{P}$  is likely to be reached coming from  $\mathcal{P}^{(0)}$ . To prove this we use Proposition 3. For the process  $(\sigma_t^{+\underline{1}})$  the law of the first excursion  $\bar{S}$  from  $+\underline{1}$  to  $-\underline{1}$  is the time reverse of the law of  $S$ . So by now we know that  $(\sigma_t^{+\underline{1}})$  is likely to visit  $\mathcal{P}$  before it hits  $-\underline{1}$ ,

$$\lim_{\beta \rightarrow \infty} P(T^{+\underline{1}}(\mathcal{P}) < T^{+\underline{1}}(-\underline{1})) = 1 .$$

Now we can reproduce the arguments above, decomposing the path  $(\sigma_t^{+\underline{1}}: t = 0, 1, \dots)$  into cycles as before, but with  $-\underline{1}$  and  $+\underline{1}$  interchanged and one different detail. The cycles of  $(\sigma_t^{+\underline{1}})$  are completed after a visit to  $\mathcal{P}$  followed by a return to  $+\underline{1}$ . We say that a cycle is “acceptable” if  $-\underline{1}$  is not visited during this cycle. A cycle is “good” if  $-\underline{1}$  is not visited before of  $\mathcal{P}$  in this cycle and after visiting  $\mathcal{P}$  the system has as first change of state a transition to  $\mathcal{P}^{(0)}$  followed by a sequence which leads to  $-\underline{1}$  without returning to  $\mathcal{P}$  or  $+\underline{1}$ . Define

$$\Omega_2 = \{ \text{a good cycle occurs for the process } (\sigma_t^{+\underline{1}}) \text{ before any bad cycle} \} .$$

Then as before

$$\lim_{\beta \rightarrow \infty} P(\Omega_2) = 1 .$$

And on  $\Omega_2$   $\bar{S}$  is a sequence which enters in  $\mathcal{P}$  only once, stays there for some time and then jumps to  $\mathcal{P}^{(0)}$ .

To complete the proof now write, using the fact that the law of  $\bar{S}$  is the time reverse of that of  $S$ ,

$$\begin{aligned} P(S \neq S_0) &\leq P((\Omega_1)^c) + P(\text{In the sequence } S \mathcal{P} \text{ is visited and } \mathcal{P} \text{ is first reached} \\ &\quad \text{with a jump from a configuration not in } \mathcal{P}^{(0)}) \\ &\leq P((\Omega_1)^c) + P(\text{In the sequence } \bar{S} \mathcal{P} \text{ is visited and the last visit to } \mathcal{P} \\ &\quad \text{is followed by a transition not to } \mathcal{P}^{(0)}) \\ &\leq P((\Omega_1)^c) + P((\Omega_2)^c). \quad \blacksquare \end{aligned}$$

*Note.* R. Kotecky and E. Olivieri are studying questions similar to those treated in [NS] and in the present paper for different Hamiltonians. It is a pleasure to thank Enzo Olivieri for discussions on these projects.

**References**

[CGOV] Cassandro, M., Galves, A., Olivieri, E., Vares, M.E.: Metastable behavior of stochastic dynamics: A pathwise approach. *J. Stat. Phys.* **35**, 603–634 (1984)  
 [FW] Freidlin, M.I., Wentzell, A.D.: *Random Perturbations of Dynamical Systems*. Berlin Heidelberg New York: Springer 1984  
 [NS] Neves, E.J., Schonmann, R.H.: Critical droplets and metastability for a Glauber dynamics at very low temperature. *Commun. Math. Phys.* **137**, 209–230 (1991)  
 [Sch] Schonmann, R.H.: An approach to characterize metastability and critical droplets in stochastic Ising models. *Ann. Inst. H. Poincaré (Phys. Theor.)* **55**, 591–600 (1991)