Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Tunneling time for an Ising spin system with Glauber dynamics

by

Tobias Kuna, and Guido Manzi

Preprint no.: 93 2006

Tunneling time for an Ising spin system with Glauber dynamics

T. Kuna G. Manzi

Abstract

The time spent by a ferromagnetic system, ruled by a Kac potential J_γ, in one of the two phases has the order of magnitude of 1 divided by the probability of observing so a strong fluctuation that an interface is produced and move until the opposite phase is reached. That probability is exponentially small in the intensity γ of the potential and the free energy gap between one of the phase and the next stationary state, where both phases are present in the same quantity separated by an interface.

1 Introduction

We want to prove an estimate for the time needed by an Ising spin system on $[-L, L]$ with Kac type potential, Neumann boundary conditions (the profiles are extended to the whole line by reflecting them around the points kL, k odd) and Glauber dynamics starting in a neighbourhood of the plus phase to reverse to the minus one, due only to fluctuations.

The picture we have in mind is a double well, whose minima are the $\pm$ phases and the saddle is the instanton.

The time for the tunneling can not be small, because large fluctuations are improbable and after a normal one has occurred, the system flows again around the initial phase according to the behaviour of the deterministic evolution. But it can not be too large, too. In fact one can exhibit explicitly paths performing the tunneling whose probability is not zero. Basically the system tries a lot of times to tunnel until it finds the right path to follow.

2 Notation and definitions

2.1 Ising system, Kac potentials and Glauber dynamics

The Ising system is a model for magnetic materials. For any point x on the grid $\mathcal{S} := [-L, L] \cap \gamma\mathbb{Z}$ we consider a spin $\sigma(x)$, namely a variable that can take the values 1 or -1. We call configuration the collection of spin $\sigma = \{\sigma(x)\}_{x\in\mathcal{S}}$. Let $H(\sigma)$ be the energy of the configuration σ:

$$H(\sigma) = -\sum_{x,y\in\mathcal{S},\, x<y} J_\gamma(x, y)\sigma(x)\sigma(y)$$

where J_γ is a Kac potential, namely $J_\gamma(x, y) = \gamma J(|x - y|)$, and J is positive (ferromagnetic), $J(x) = 0$ if $|x| > 1$ and $\int J = 1$.

On a macroscopic scale we disregard the details of each configuration; we only measure local averages. Since a macroscopic system in equilibrium is characterized by well defined quantities (e.g. the magnetization), not all configurations can correspond to them. Then there are configurations that can be considered typical. We formalize this notion by introducing a weight for each configuration:

$$\mu_\gamma(\{\sigma\}) = \frac{e^{-\beta H(\sigma)}}{Z_\gamma}$$

where Z_γ is a normalization factor and β is the reciprocal of the temperature. μ_γ is a probability measure (Gibbs measure) on the space of all possible configurations. Equipping that space with a suitable topology (see below) it is possible to prove a large deviation principle for μ_γ with speed γ^{-1} and rate functional given by the free energy defined in (1). It means basically that the equilibrium of the system is characterized by configurations with low energy, but also that are quite common (high entropy). When the temperature is high (β small) the disorder prevails on the energy and the average magnetization is zero; when $\beta > 1$, the average magnetization is different from zero and it can be either positive (m_β) or negative ($-m_\beta$). The choice between the two values can be modulated by a vanishing external magnetic field.

Of course we do not expect that microscopically the system stays forever in the same configuration, even though it is typical. Rather we imagine some mechanism through which the system switches from one configuration to another, staying typical if it was at the beginning or converging to some equilibrium in the other case. We then define a Markov process that leaves μ_γ invariant; an example is the so called Glauber dynamics (non conservative dynamics), defined by the following generator:

$$Lf(\sigma) = \sum_{x\in\mathcal{S}} c(x,\sigma)[f(\sigma^x) - f(\sigma)]$$

where f depends on finitely many spins and σ^x is the configuration obtained by σ flipping the spin in x. $c(x,\sigma)$ are the jump rates:

$$c(x,\sigma) = \frac{e^{-\sigma(x)\sum_{y\in\mathcal{S}\setminus\{x\}} J_\gamma(x,y)\sigma(y)}}{e^{\sum_{y\in\mathcal{S}\setminus\{x\}} J_\gamma(x,y)\sigma(y)} + e^{-\sum_{y\in\mathcal{S}\setminus\{x\}} J_\gamma(x,y)\sigma(y)}}$$

On each site of $\mathcal{S}$ there is a Poissonian clock with intensity $c(x,\sigma)$; when the first clock rings we flip the corresponding spin and restart all the clocks.

The Glauber dynamics induces a measure on the path space of the realizations of the process. Again it is possible to prove an l.d.p. for that measure with speed γ^{-1} and rate functional given by the Comets cost functional $\mathcal{I}_T$, see [3]. It implies that the process will try to follow in average the solution of the mean field equation (2).

When $\beta > 1$, the constant profiles $\pm m_\beta$ are stationary points of (2) and global minimizer of the free energy; a lot of configurations are attracted by them. The next stationary point in the scale of free energy is the so called instanton, $\hat{m}$, that is the unique antisymmetric, strictly increasing stationary solution of (2) on $[-L, L]$ and Neumann boundary conditions.

A tunneling is a large deviation from the average behaviour.

2.2 Local averages topology

Let $\mathcal{B}$ be the set of profiles whose absolute value is not greater than 1, $\mathcal{B} := L^\infty([-L, L]; [-1, 1])$. The weak topology on $\mathcal{B}$ is defined by the requirement that a sequence $\{m_n\} \subset \mathcal{B}$ converges to $m \in \mathcal{B}$ if and only if for any $f \in \mathcal{C}([-L, L])$

$$< f, m_n > \to < f, m >$$

where $< f, g >= \int_{[-L,L]} fg$. It can be shown that it is equivalent to the following topology. Let us consider a partition $\mathcal{D}^{(\ell)}$ of $[-L, L]$ in intervals Δ whose length is ℓ. Then we introduce the local average $\mathcal{P}^{(\ell)}m$ of the profile m through

$$\mathcal{P}^{(\ell)}m(x) := \sum_{\Delta\in\mathcal{D}^{(\ell)}} \left(\frac{1}{|\Delta|} \int_\Delta dx'\, m(x') \right) 1\!\!1_\Delta(x)$$

We say that m_n converges to m with respect to the local averages if $\mathcal{P}^{(\ell)}m_n \to \mathcal{P}^{(\ell)}m$ for any ℓ.

We denote by ρ a metric compatible with the weak topology (see below); $W_r(\phi)$ is a ball of radius r around the profile ϕ: $W_r(\phi) = \{m \in \mathcal{B} : \rho(m, \phi) < r\}$.

We define stopping times as $\tau(A) = \inf\{t \geq 0 : \sigma_t \in A\}$, where σ_t is a realization of the Glauber dynamics.

We call *profiles* functions defined on the line periodic of period $4L$ and symmetric around kL, K odd.

2.3 Energy landscape

We call D_+ and D_- the basins of attraction of m_β and $-m_\beta$ respectively with respect to the mean field equation (2), the convergence is in L_∞. $D_\pm$ are open domains in the weak topology (the flow defined by (2) is continuous with respect to the initial data, see below).

The free energy is defined by

$$\mathcal{F}(m) = \int_{-L}^{L} \phi_\beta(m)dx + \frac{1}{4}\int_{-L}^{L}\int_{-L}^{L} J^{\mathrm{neum}}(x,x')(m(x) - m(x'))^2 dxdx' \tag{1}$$

where $\beta > 1$,

$$\phi_\beta(m) = \tilde{\phi}_\beta(m) - \min_{|s|\leq 1} \tilde{\phi}_\beta(s), \quad \tilde{\phi}_\beta(m) = -\frac{m^2}{2} - \frac{1}{\beta}\mathcal{E}(m)$$

$$\mathcal{E}(m) = -\frac{1-m}{2}\ln\frac{1-m}{2} - \frac{1+m}{2}\ln\frac{1+m}{2}$$

and

$$J^{\mathrm{neum}}(x,y) = J(x,y) + J(x, R_L(y)) + J(x, R_{-L}(y))$$

with $R_\xi(y) = 2\xi - y$. We also notice that a system with von Neumann boundary conditions on the interval $[-L, L]$ can be seen as a system with periodic boundary conditions on $[-L, 3L]$ with the additional constrain of reflection symmetry around L, i.e. $m(x + 4L) = m(x)$ and $m(2L - x) = m(x)$. For this type of function the set $\left\{ \cos\left(\frac{2\pi k}{4L}(x - L)\right)\right\}_{k\in\mathbb{N}}$ is a basis. Moreover

$$\int_{-L}^{L} dy J^{\mathrm{neum}}(x,y)m(y) = J * m(x)$$

If we consider the set of all profiles whose free energy is less than that of the instanton plus some positive (small) constant, i.e. $G := \{m \in \mathcal{B} : \mathcal{F}(m) \leq \mathcal{F}(\hat{m}) + \sigma\}$, then we know that there are only three stationary points in this set: the two phases and the instanton, citeBDP. Moreover this set is compact in the weak topology (it is a level set for a good rate functional) and it is invariant under the deterministic evolution described by

$$\partial_t m = -m + \tanh(\beta J * m) \tag{2}$$

because $\mathcal{F}$ is a Lyapunov functional for that evolution.

Notice that $\overline{D_-} \subset D_+^c$ because D_+^c is closed and $D_+ \cap D_- = \emptyset$.

3 Main result

Theorem 1. *For any $\delta > 0$, $\sigma \in D_+$ and $r_- > 0$ such that $W_{r_-}(-m_\beta) \subset D_-$ we have*

$$\lim_{\gamma\to\infty} P_{\gamma,\sigma}(T^- < \tau_\gamma < T^+) = 1$$

where $P_{\gamma,\sigma}$ is the path measure on the space of realizations of the Glauber dynamics starting from the configuration σ; $T^\pm = \exp\{\gamma^{-1}(\mathcal{F}(\hat{m}) \pm \delta)\}$ and $\tau_\gamma = \tau(W_{r_-}(-m_\beta))$.

The global strategy is standard (see [5]), but there are several technical points that needed some fresh new ideas, especially in taking care of the lack of continuity of $\mathcal{F}$ and $\mathcal{I}_T$ in the weak topology.

4 D_+ and D_- are open

Here we prove that the basins of attraction of the two pure phases are open sets in the weak topology. The key ingredient is the continuity of the flow with respect to the initial data; to be precise we will use a stronger property of the evolution, namely that after a suitable long time, if the evolution starts from two points close in a weak sense, then the flows will be close in L_∞ sense. The strategy is the following: given $m \in D_+$ (for D_- the argument is the same), we follow its evolution until a time $\bar{t}$ when it enters an L_∞ ball $B_{\delta/2}(m_\beta)$ around m_β such that the free energy in $B_\delta(m_\beta)$ is strictly smaller than the one of the instanton and the magnetization is positive and also $2e^{-\bar{t}} < \delta/4$. This is possible because the free energy is continuous in L_∞. Then we can choose a weak ball $W_{\delta'}(m)$ such that for any $\tilde{m} \in W_{\delta'}(m)$ $\bar{m} := T_{\bar{t}}\tilde{m} \in B_{\delta/2}(T_{\bar{t}}m)$ (see below). Now $T_t\bar{m}$ will converge to m_β because the only available stationary points are the two pure phases ([1]), but the evolution cannot reach $-m_\beta$ because otherwise there would be a time when the magnetization is zero, but in this class of profiles, the least free energy is attained at the instanton ([4]).

4.1 Continuity of the flow w.r.t. the initial data

We want to show that for any $m_0^{(1)}$, $m_0^{(2)}$ and any time t and $\epsilon > 0$, if $\rho(m_0^{(1)}, m_0^{(2)}) < \delta$, then $\rho(m_t^{(1)}, m_t^{(2)}) < \epsilon$ for δ sufficiently small, where $m_t^{(i)} = T_t m_0^{(i)}$. We write (2) in the following integral form:

$$m_t = e^{-t}m_0 + \int_0^t ds\, e^{-(t-s)} \tanh(\beta J * m_s)$$

The whole point is to show that $|\tanh(\beta J * m_s^{(1)}) - \tanh(\beta J * m_s^{(2)})|$ is small for all $s \leq t$. Notice that $\beta J * m_s$ is a solution of the following system

$$\partial_t F = -F + \beta J * \tanh F, \quad F_0 = \beta J * m_0$$

whose flow is clearly continuous w.r.t. the initial data even in L_∞ sense. This is sufficient to conclude, because $J * (m_0^{(1)} - m_0^{(2)})$ is small if $m_0^{(1)}$ and $m_0^{(2)}$ are close in a weak sense.

From the proof it is clear that if the time t is such that $2e^{-t} < \epsilon/2$ and $J * (m_0^{(1)} - m_0^{(2)})$ is so small that $|\tanh(\beta J * m_s^{(1)}) - \tanh(\beta J * m_s^{(2)})|$ is smaller than $\epsilon/2$, then $|m_t^{(1)} - m_t^{(2)}| < \epsilon$. This is the stronger property of the flow that we are using in the previous section.

5 Lower bound

In this section we will show

Theorem 2. *For all $\delta > 0$ and all $\sigma \in D_+$*

$$\lim_{\gamma \to 0} P_{\gamma,\sigma}\big(\tau_\gamma < T^-\big) = 0$$

where T^-, τ_γ and r_- are as in theorem (1).

We perform the proof in three steps: a) we show the lower bound for an initial data distributed according to the stationary Gibbs measure restricted to a sufficiently small ball around m_β using the l.d.p. that holds for this measure; b) then we extend the result to any initial data in a possibly smaller ball through a coupling method and, finally, c) we complete the proof thanks to the closeness of the process to the deterministic path.

5.1 Step a)

Given a set $A \subset \mathcal{B}$, $A^c := \mathcal{B} \setminus A$, we call A_θ and $\partial_\theta A$ the sets

$$\begin{aligned} A_\theta &= \{m \in \mathcal{B} : \rho(A, m) < \theta\} \\ \partial_\theta A &= \{m \in A^c : \rho(A, m) < \theta\} \end{aligned}$$

We introduce $U = \{m \in \mathcal{B} : \mathcal{F}(m) > \mathcal{F}(\hat{m}) - \epsilon\}$, $\epsilon < \delta$; being U^c a level set, it is compact. It follows by a general result (the distance between two disjoint sets, one closed, the other compact, is strictly positive) that there exists $\theta > 0$ such that $2\theta = \rho(U^c \cap D_+, \overline{D_-})$ because $U^c \cap D_+$ is a subset of the compact set $U^c \cap \overline{D_+}$ and we only need to prove that $(U^c \cap \overline{D_+}) \cap \overline{D_-} = \emptyset$. We argue by contradiction: let us suppose that there is $m \in U^c \cap \overline{D_+} \cap \overline{D_-}$. Then $m \notin D_+$ and $m \notin D_-$ because otherwise $D_+ \cap \overline{D_-} \neq \emptyset$ or $D_- \cap \overline{D_+} \neq \emptyset$, which is not possible. But we know that there exists u such that the flow starting from m converges to u (by subsequences) and u is a stationary solution of (2) and $\mathcal{F}(u) \leq \mathcal{F}(m)$ (see Presutti et al.). Then u cannot be $\hat{m}$ because $m \in U^c$, the set of profiles with free energy less or equal to $\mathcal{F}(\hat{m}) - \epsilon$ and obviously it cannot be either m_β or $-m_\beta$, because otherwise m would belong to D_+ or to D_-. This leads to the desired contradiction because according to a recent result by Bellettini, De Masi and Presutti ([1]) the only stationary points below the instanton are the two pure phases.

Now we can prove that $\partial_\theta(U^c \cap D_+) \subset U$. In fact, since $\rho(U^c \cap D_+, \overline{D_-}) = 2\theta$, $\partial_\theta(U^c \cap D_+) \cap \overline{D_-} = \emptyset$; but then $\partial_\theta(U^c \cap D_+) \subset (U \cup D_+^c) \cap \overline{D_-}^c = [U \cup (D_+^c \cap U^c)] \cap \overline{D_-}^c$. By the above mentioned result about the energy levels of the free energy we know that $U^c \subset D_+ \cup D_-$. It follows that $D_+^c \cap U^c \subset D_-$ and then $D_+^c \cap U^c \cap \overline{D_-}^c = \emptyset$. We are left with $\partial_\theta(U^c \cap D_+) \subset U \cap \overline{D_-}^c$ which is sufficient to conclude.

Let $\mu_{\gamma,W}$ be the stationary Gibbs measure restricted to the set W: $\mu_{\gamma,W} = 1\!\!1_W \mu_\gamma / \mu_\gamma(W)$ (provided $\mu_\gamma(W) > 0$). We will choose W as a set containing m_β and contained in $(U^c \cap D_+)_\theta$. Actually, since D_+ is open we take a sufficiently small weak ball around m_β contained in $(U^c)_\theta$, e.g. $W = W_{\theta/2}(m_\beta)$.

Let $P_{\mu_{\gamma,W}}$ be the law of the process whose initial state is distributed according to $\mu_{\gamma,W}$. We have for any given δ' and for γ small enough

$$P_{\mu_{\gamma,W}}(\sigma_T \in \partial_\theta(U^c \cap D_+)) \leq \frac{P_{\mu_\gamma}(\sigma_T \in \partial_\theta(U^c \cap D_+))}{\mu_\gamma(W)}$$
$$\leq \frac{e^{-\gamma^{-1}(\mathcal{F}(\hat{m}) - \epsilon - \delta')}}{e^{-\gamma^{-1}(\inf_{m \in W} \mathcal{F}(m) + \delta')}} = e^{-\gamma^{-1}(\mathcal{F}(\hat{m}) - \epsilon - 2\delta')}$$

and we note that $\inf_{m \in W} \mathcal{F}(m) = 0$ since $m_\beta \in W$; for the numerator we exploited the fact that $\partial_\theta(U^c \cap D_+) \subset U$. Here we have used the stationarity of the Gibbs measure w.r.t. the Glauber dynamics and the large deviation principle for μ_γ.

Given $\theta/2$ there is a partition of $[-L, L]$ in subintervals I and a number θ' such that if $\sum_I |\int_I m_1 - m_2| < \theta'$ then $\rho(m_1, m_2) < \theta/2$. Let Δt be a small time interval, say $\Delta t = \gamma^a$, $a > 0$. Since the jump rates are bounded by a constant c_M, we can estimate the probability of having more than N jumps in a time interval Δt by

$$\frac{(2L\gamma^{-1} c_M \Delta t)^N}{N!} \approx \exp\{-N \ln \frac{N}{2L c_M \gamma^{-1} \Delta t} + \ldots\}$$

If we choose $N = \gamma^{-1}/(\ln \gamma^{-1})^b$, $0 < b < 1$, then the above probability goes to zero, when γ vanishes, faster than $e^{-\gamma^{-1} c}$, for any possible c (e.g. $c = \mathcal{F}(\hat{m}) + 2\sigma$). If the number of spin flips is not greater than N, the largest change in magnetization in a coarse cell I is $(2\gamma/|I|)N = 2/(|I|(\ln \gamma^{-1})^b)$. It follows that

$$\sum_I |\int_I \sigma_{t+\Delta t} - \sigma_t| < \frac{4L}{|I|(\ln \gamma^{-1})^b} < \theta'$$

provided γ sufficiently small. With an abuse of notation we denoted with σ_t the following function: $\sum_{y \in \mathcal{S}} \sigma_t(y) 1\!\!1_{y \leq x < y + \gamma}$.

For any realization $\bar{\sigma}_t$ such that $\bar{\sigma}_0$ is distributed according to $\mu_{\gamma,W}$ and there is some $T > 0$ such that $\bar{\sigma}_T \in W_{r_-}(-m_\beta)$, we define the sequence $\sigma_i := \bar{\sigma}_{i\Delta t}$. With overwhelming probability (see above) $\rho(\sigma_{i+1}, \sigma_i) < \theta/2$; since the process starts in $(U^c \cap D_+)_\theta$ and after a while enters D_- and these two sets are disjoint, we can conclude that there is l such that $\sigma_l \in (U^c \cap D_+)_\theta$ but

$\sigma_{l+1} \notin (U^c \cap D_+)_\theta$. It follows that $\sigma_l \in \partial_\theta(U^c \cap D_+)$, because otherwise the distance between σ_l and σ_{l+1} would be larger than θ.

Define $M = [T^-/\Delta t]$, then

$$\begin{aligned} P_{\mu_{\gamma,W}}(\tau_\gamma < T^-) &\leq (M+1)\tfrac{(2Lc_M\gamma^{-1}\Delta t)^N}{N!} + \textstyle\sum_{i=1}^M P_{\mu_{\gamma,W}}(\sigma_i \in \partial_\theta(U^c \cap D_+)) \\ &\leq e^{\gamma^{-1}(\mathcal{F}(\hat{m})-\delta)}\gamma^{-a}\big(e^{-c\gamma^{-1}} + e^{-\gamma^{-1}(\mathcal{F}(\hat{m})-\epsilon-2\delta')}\big) \to 0 \end{aligned}$$

If we choose $\delta' < (\delta-\epsilon)/2$. This concludes the first step.

5.2 Step b)

We consider the following metric of the weak topology

$$\rho(m_1, m_2) := \sup_{k\in\mathbb{N}} \frac{\big|\int_{-L}^{L} dx(m_1(x) - m_2(x)) \cos\big(\frac{2\pi k}{4L}(x-L)\big)\big|}{1+k}$$

$$\rho_t(\sigma', \sigma'') := \sup_{0\leq\tau\leq t} \rho(\sigma'_\tau, \sigma''_\tau)$$

Using the reflection symmetry one sees that $\int_{-L}^{3L} m(x)dx = 2\int_{-L}^{L} m(x)dx$. The reason to introduce this metric is the following lemma

Lemma 1. *Let m_1 and m_2 be two profiles, then*

$$\rho(J * m_1, J * m_2) \leq \rho(m_1, m_2) \tag{3}$$

and, for sufficiently smooth J, there exists a constant $C_L > 0$ such that

$$|J * m(x)| \leq C_L \rho(m, 0).$$

The first inequality holds because of the following calculation

$$\begin{aligned} \left|2\int_{-L}^{L} dx J * m(x) \cos\big(\frac{2\pi k}{4L}(x-L)\big)\right| &= \left|\int_{-L}^{3L} dx J * m(x) \Re e^{i\frac{2\pi k}{4L}(x-L)}\right| \\ &= \left|\Re\Big(\int_{-L}^{3L} dx J(x) e^{i\frac{2\pi k}{4L}x} \int_{-L}^{3L} dx m(x) e^{i\frac{2\pi k}{4L}(x-L)}\Big)\right| \\ &\leq \left|\int_{-L}^{3L} dx J(x) e^{i\frac{2\pi k}{4L}x}\right| \left|\Re\int_{-L}^{3L} dx m(x) e^{i\frac{2\pi k}{4L}(x-L)}\right| \end{aligned}$$

where the last equality follows from the fact that by the symmetry $\Im\big(\int_{-L}^{3L} dx m(x) e^{i\frac{2\pi k}{4L}(x-L)}\big) = 0$. Of course we can assume $L > 1$. Observing that $\big|\int_{-L}^{3L} dx J(x) e^{i\frac{2\pi k}{4L}x}\big| \leq 1$ one obtains the inequality (3).

In order to prove the second inequality we express the convolution in terms of the Fourier transform

$$J*m(x) = \frac{1}{2}\left[\frac{1}{2L}\int_{-L}^{3L} dy J * m(y)\right] + \sum_{k=1}^{\infty} \cos\left(\frac{2\pi k}{4L}(x-L)\right)\left[\frac{1}{2L}\int_{-L}^{3L} J * m(y) \cos\left(\frac{2\pi k}{4L}(y-L)\right)\right]$$

Then

$$|J * m(x)| \leq \frac{1}{2L}\sum_{k=0}^{\infty}\left|\int_{-L}^{3L} dy J * m(y) \cos\left(\frac{2\pi k}{4L}(y-L)\right)\right|$$

and since

$$\left|\int J(x) e^{i\frac{2\pi k}{4L}x} k^3\right| \leq \left(\frac{4L}{2\pi}\right)^3 2||J'''||_\infty$$

we can conclude that

$$|J * m(x)| \le \rho(m,0)CL^2||J'''||_\infty$$

Of course this estimate is rough and can be improved in terms of L and of smoothness of J.

These two inequalities can be used to derive that in a sufficiently small ball around m_β the metric decays exponentially in time. The integral form of the mean field equation yields

$$\rho(m_t, m_\beta) \le \rho(m_0, m_\beta)e^{-t} + \int_0^t dse^{-(t-s)}\rho\big(\tanh(\beta J * m_s), \tanh(\beta J * m_\beta)\big)$$

According to Taylor expansion

$$\tanh(\beta J * m_s) - \tanh(\beta J * m_\beta) = \frac{\beta}{\cosh^2(\beta m_\beta)}(J * m_s - J * m_\beta) - \frac{\beta^2 \tanh(\xi)}{\cosh^2(\xi)}(J * m_s - J * m_\beta)^2$$

one gets for the distance

$$\rho\big(\tanh(\beta J * m_s), \tanh(\beta J * m_\beta)\big) \le \beta(1 - m_\beta^2)\rho(J * m_s, J * m_\beta) + \beta^2 2L \sup_x |J * m_s(x) - J * m_\beta|^2$$

Applying the first inequality to the first summand on the r.h.s. and the second one to the second summand, one obtains, where $C' > 0$, $C' \approx L^3$,

$$\rho(m_t, m_\beta) \le \rho(m_0, m_\beta)e^{-t} + \int_0^t dse^{-(t-s)}\Big(\beta(1 - m_\beta^2)\rho(m_s, m_\beta) + C'\rho(m_s, m_\beta)^2\Big)$$

As $\beta(1 - m_\beta^2) < 1$, we proved

Lemma 2. *There exists an $r_0, \delta_0 > 0$ such that if $\rho(m_0, m_\beta) \le r_0$ then*

$$\rho(m_t, m_\beta) \le e^{-\delta_0 t}\rho(m_0, m_\beta).$$

In this part we prove that if one starts from a small ball, $W_{r_1}(m_\beta)$, around m_β in the weak topology, the process will leave the ball $W_r(m_\beta)$, $r > 2r_1$, only after a very long time, i.e. the following proposition hold:

Proposition 1. *There exists $c > 0$ such that*

$$\sup_{\sigma \in W_{r_1}(m_\beta)} P_{\gamma,\sigma}\Big(\tau(W_r(m_\beta)^c) < e^{c\gamma^{-1}}\Big) \le e^{-c\gamma^{-1}}, \tag{4}$$

where $r_1 < \min\{r/2, r_0\}$.

Due to the uniform (exponential) decay of the metric ρ under the dynamics, there exists a time T such that any deterministic trajectory which starts from a point $\sigma \in W_{r_1}(m_\beta)$ will be in $W_{r_1/2}(m_\beta)$ at time T. Therefore, by the large deviation principle, for any $\sigma_0 \in W_{r_1}(m_\beta)$

$$P_{\gamma,\sigma_0}\big(\tau(W_r(m_\beta)^c) > T, \sigma_T \in W_{r_1}(m_\beta)\big) \ge P_{\gamma,\sigma}\big(\rho_T(\sigma_t, m_t^\sigma) < r_1/2\big) \ge 1 - e^{-c_1\gamma^{-1}},$$

where

$$c_1 := \inf\big\{\mathcal{I}_T(\tilde{m})\big|\sigma \in W_{r_1}(m_\beta), \tilde{m} \text{ such that } \rho_T(\tilde{m}, m^\sigma) \ge r_1/2 \text{ and } \tilde{m}_0 = \sigma\big\}$$

and m^σ denotes the deterministic path started at σ. In more generality, we used that for compact W, $\tilde{r} > 0$ by the large deviation principle

$$\sup_{\sigma \in W} P_{\gamma,\sigma}\big(\rho_T(\sigma_t, m_t^\sigma) \ge \tilde{r}\big) \le e^{-c_W\gamma^{-1}} \tag{5}$$

and that the corresponding constant

$$c_W := \inf\big\{\mathcal{I}_T(\tilde{m})\big|\sigma \in W, \tilde{m} \text{ such that } \rho_T(\tilde{m}, m^\sigma) \ge \tilde{r} \text{ and } \tilde{m}_0 = \sigma\big\}$$

is strictly greater than zero. As we want to prove this by contradiction we assume that $c_W = 0$. Then there exists a sequence of pairs $(\sigma^{(n)}, \tilde{m}^{(n)})_n$ such that $\tilde{m}_0^{(n)} = \sigma^{(n)}$, $\rho_T(\tilde{m}^{(n)}, m^{\sigma^{(n)}}) \geq \tilde{r}$ and $\lim_{n\to\infty} \mathcal{I}_T(\tilde{m}^{(n)}) = 0$. Due to the convergence of the rate function the sequence $(\tilde{m}^{(n)})_n$ lies in a level set and hence has a convergent subsequence. The corresponding subsequence of $(\sigma^{(n)})_n$ converges in the weak topology. Without loss of generality we hence assume that there exists a pair $(\sigma, \tilde{m})$ such that $\rho_T(\tilde{m}^{(n)}, \tilde{m})$ and $\rho(\sigma^{(n)}, \sigma)$ converge to zero. By lower semi-continuity also $\mathcal{I}_T(\tilde{m}) = 0$. But the deterministic flow is continuous w.r.t. the initial data, in other words $\rho_T(m^{\sigma^{(n)}}, m^\sigma)$ converges to zero, then also $\rho_T(\tilde{m}, m^\sigma) \geq \tilde{r}$ holds. Hence $\tilde{m}$ is not the deterministic path started at σ which contradicts $\mathcal{I}_T(\tilde{m}) = 0$.

Applying the result $e^{c_1/2\gamma^{-1}}$ times one obtains the result for $c < c_1/2$ by the Markov property.

5.2.1 Coupling and memory loosing

In the following we want to show that after a time of order $(\ln \gamma^{-1})^2$ the process loses is memory near the minimum. More precisely, we want to show that there exists a coupling Q_γ of $P_{\gamma,\sigma'}$ with $P_{\gamma,\sigma''}$ such that

$$Q_\gamma\left(\{\sigma'_t \neq \sigma''_t \text{ for } t \geq (\ln \gamma^{-1})^2\}\right) \to 0$$

when γ goes to zero. One coupling which fulfils this is the so-called *best coupling*, defined as the path measure of the process generated by the following generator

$$\begin{aligned} L^{(c)} f(\sigma', \sigma'') \quad := \quad & \sum_{x\in\mathcal{S}} \mathbb{1}_{\sigma'(x)\neq\sigma''(x)} \Big(c(x,\sigma')[f(\sigma'^x, \sigma'') - f(\sigma', \sigma'')] + c(x,\sigma'')[f(\sigma', \sigma''^x) - f(\sigma', \sigma'')] \Big) \\ & + \mathbb{1}_{\sigma'(x)=\sigma''(x)} \Big(c(x,\sigma') \wedge c(x,\sigma'')[f(\sigma'^x, \sigma''^x) - f(\sigma', \sigma'')] \\ & + (c(x,\sigma') - c(x,\sigma') \wedge c(x,\sigma''))[f(\sigma'^x, \sigma'') - f(\sigma', \sigma'')] \\ & + (c(x,\sigma'') - c(x,\sigma') \wedge c(x,\sigma''))[f(\sigma', \sigma''^x) - f(\sigma', \sigma'')] \Big) \end{aligned}$$

with initial condition (σ'_0, σ''_0).

The idea of the best coupling is to make the two process to coincide: if $\sigma'(x) \neq \sigma''(x)$ we look at two Poissonian clocks with parameters $c(x,\sigma')$ and $c(x,\sigma'')$ that can ring independently from each other. If $\sigma'(x) = \sigma''(x)$ then the clocks are different: the first clock has parameter $c(x,\sigma') \wedge c(x,\sigma'')$, the second $|c(x,\sigma') - c(x,\sigma'')|$. So there are $2(2L\gamma^{-1})$ clocks. When the first clock rings we check to which site it is associated; if σ' and σ'' differ there, then we flip the spin of the configuration whose clock rang; if σ' and σ'' are the same, then we flip the spins of both the configurations if the clock that rang is the one with parameter $c(x,\sigma') \wedge c(x,\sigma'')$, otherwise we flip only the spin of the configuration corresponding to the largest parameter. After that we restart all the clocks.

As $L^{(c)}$ is obviously a generator of a Markov process, Q is a probability measure. In order to show that the marginals of Q are given by P, we use that the generator $L^{(c)}$ has, for functions which depend only on one variable, the form

$$\begin{aligned} L^{(c)} f(\sigma') \quad = \quad & \sum_{x\in\mathcal{S}} \mathbb{1}_{\sigma'(x)\neq\sigma''(x)} \Big(c(x,\sigma')[f(\sigma'^x) - f(\sigma')] \\ & + \mathbb{1}_{\sigma'(x)=\sigma''(x)} \Big(c(x,\sigma') \wedge c(x,\sigma'')[f(\sigma'^x) - f(\sigma')] \\ & + (c(x,\sigma') - c(x,\sigma') \wedge c(x,\sigma''))[f(\sigma'^x) - f(\sigma')] \Big) \\ = \quad & \sum_{x\in\mathcal{S}} c(x,\sigma')[f(\sigma'^x) - f(\sigma')] = Lf(\sigma'). \end{aligned}$$

Hence, the marginal in the first variable is $P_{\gamma,\sigma'}$. The marginal in the second variable can be treated analogously. It remains to show that the coupling gives coalescence.

Consider the function $\alpha(\sigma',\sigma'') := \sum_y |\sigma'(y)-\sigma''(y)|$. First, we compute for any summand separately

$$\begin{aligned} L|\sigma'(y)-\sigma''(y)| &= \mathbb{1}_{\sigma'(y)\neq\sigma''(y)}\big(c(y,\sigma')+c(y,\sigma'')\big)[-2] + \\ & \quad \mathbb{1}_{\sigma'(y)=\sigma''(y)}\big|c(y,\sigma')-c(y,\sigma'')\big|2 \end{aligned} \tag{6}$$

Using the form of $c(x,\sigma) \equiv \frac{e^{-\sigma(x)h_\gamma(x,\sigma)}}{e^{h_\gamma(x,\sigma)}+e^{-h_\gamma(x,\sigma)}}$ where $h_\gamma(x,\sigma) \equiv \sum_{y\in\mathcal{S}\setminus\{x\}} J_\gamma(x,y)\sigma(y)$. Then one can rewrite for $\sigma'(x) = -\sigma''(x)$

$$\begin{aligned} c(x,\sigma')+c(x,\sigma'') &= \frac{e^{-\sigma'(x)h_\gamma(x,\sigma')}}{e^{h_\gamma(x,\sigma')}+e^{-h_\gamma(x,\sigma')}} + \frac{e^{\sigma'(x)h_\gamma(x,\sigma'')}}{e^{h_\gamma(x,\sigma'')}+e^{-h_\gamma(x,\sigma'')}} \\ &= \frac{e^{-\sigma'(x)h_\gamma(x,\sigma')}}{e^{h_\gamma(x,\sigma')}+e^{-h_\gamma(x,\sigma')}} - \frac{e^{-\sigma'(x)h_\gamma(x,\sigma'')}}{e^{h_\gamma(x,\sigma'')}+e^{-h_\gamma(x,\sigma'')}} + 1 \end{aligned}$$

Therefore, it remains to estimate for σ',σ'' such that $\sigma'(y)=\sigma''(y)$

$$\begin{aligned} 2|c(y,\sigma')-c(y,\sigma'')| &= \Big|\frac{e^{\pm h_\gamma(x,\sigma')}}{\cosh(h_\gamma(x,\sigma'))} - \frac{e^{\pm h_\gamma(x,\sigma'')}}{\cosh(h_\gamma(x,\sigma'')]}\Big| \\ &\leq \int_0^1 ds \frac{1}{\cosh^2(sh_\gamma(x,\sigma')+(1-s)h_\gamma(x,\sigma''))}|h_\gamma(x,\sigma')-h_\gamma(x,\sigma'')| \end{aligned}$$

where we use that the derivative

$$\frac{d}{dx}\frac{e^{\pm x}}{\cosh(x)} = \frac{\pm e^{\pm x}\cosh(x) - e^{\pm x}\sinh(x)}{\cosh^2(x)} = \frac{\pm 1}{\cosh^2(x)}$$

Working as before in proving lemma (2), this can be bounded by

$$\leq \quad \beta(1-m_\beta^2)\sum_{x\neq y} J_\gamma(x,y)\ |\sigma'(x)-\sigma''(x)| + \beta^2\Big|\sum_{x\neq y} J_\gamma(x,y)\ (\sigma'(x)-\sigma''(x))\Big|^2$$

where x and y are running in $\mathcal{S}$. Inserting this in (6) we get

$$L|\sigma'(y)-\sigma''(y)| \leq -2\mathbb{1}_{\sigma'(y)\neq\sigma''(y)} + \beta(1-m_\beta^2)\sum_{x\neq y} J_\gamma(x,y)\ |\sigma'(x)-\sigma''(x)| \tag{7}$$

$$+\beta^2\Big(\sum_{x\neq y} J_\gamma(x,y)\ |\sigma'(x)-\sigma''(x)|\Big)^2 \tag{8}$$

Then we use the bounds

$$\sum_y\sum_{x\neq y} J_\gamma(x,y)\ |\sigma'(x)-\sigma''(x)| = \sum_x\sum_{y\neq x} J_\gamma(x,y)\ |\sigma'(x)-\sigma''(x)| \leq \sum_x J_\gamma(x,0)\ \alpha(\sigma',\sigma'')$$

and

$$\begin{aligned} &\sum_y\Big(\sum_{x_1\neq y} J_\gamma(x_1,y)\ (\sigma'(x_1)-\sigma''(x_1))\Big)\Big(\sum_{x_2\neq y} J_\gamma(x_2,y)\ (\sigma'(x_2)-\sigma''(x_2))\Big) \\ &\leq \Big(\sup_{y'}\big|J*(\sigma'-\sigma'')(y')\big| + c\gamma\Big)\sum_y\sum_{x_1\neq y} J_\gamma(x_1,y)\ |\sigma'(x_1)-\sigma''(x_1)| \\ &\leq \Big(C_L\rho(\sigma',\sigma'')+c\gamma\Big)\sum_x J_\gamma(x,0)\alpha(\sigma',\sigma'') \end{aligned}$$

Finally, we note that, since J is a probability kernel, $|\sum_x J_\gamma(x,0)-1|\le c\gamma$. Then

$$\begin{aligned} L\alpha(\sigma',\sigma'') &\le -2\sum_y \mathbb{1}_{\sigma'(y)\neq\sigma''(y)} + \beta(1-m_\beta^2)\sum_y\sum_{x\neq y} J_\gamma(x,y)\ |\sigma'(x)-\sigma''(x)| && (9)\\ &\quad +\beta^2\sum_y\Big(\sum_{x\neq y} J_\gamma(x,y)\ |\sigma'(x)-\sigma''(x)|\Big)^2 && (10)\\ &\le \Big(-1+\beta(1-m_\beta^2)(1+c\gamma)+\beta^2(1+c\gamma)\big(C\rho(\sigma',\sigma'')+c\gamma\big)\Big)\alpha(\sigma',\sigma'') && (11)\end{aligned}$$

As $\beta(1-m_\beta^2)<1$, for γ small enough one can choose an $r>0$ independent of γ such that for all σ',σ'' with $\rho(\sigma',\sigma'')<2r$ the pre-factor is less than zero, i.e. less or equal to $-\omega$.

$$\begin{aligned} \mathbb{E}_Q\left[L\alpha_t\right] &\le -\omega\mathbb{E}_Q\left[\alpha_t\right]+C'\gamma^{-1}\ Q\big(\sigma_t'\notin W_r(m_\beta) \text{ or } \sigma_t''\notin W_r(m_\beta)\big)\\ &\le -\omega\mathbb{E}_Q\left[\alpha_t\right]+C'\gamma^{-1}\ \Big(P_{\sigma'}\big(\sigma_t'\notin W_r(m_\beta)\big)+P_{\sigma''}\big(\sigma_t''\notin W_r(m_\beta)\big)\Big),\end{aligned}$$

where $C'=16Lc_M$ and we use that the marginals of Q are $P_{\sigma'}$ and $P_{\sigma''}$. In order to estimate the last summand uniformly in times of order $e^{c\gamma^{-1}}$ we use stopping times and the bound (4)

$$\begin{aligned} \mathbb{E}_Q\left[L\alpha_t\right] &\le -\omega\mathbb{E}_Q\left[\alpha_t\right]+2C'\gamma^{-1}\sup_{\sigma\in W_{r_1}(m_\beta)} P_\sigma\big(\tau(W_r(m_\beta)^c)\le e^{c\gamma^{-1}}\big)\\ &\le -\omega\mathbb{E}_Q\left[\alpha_t\right]+2C'\gamma^{-1}e^{-c\gamma^{-1}}.\end{aligned}$$

Therefore, by iterating the estimate one obtains that for all $t\le e^{c\gamma^{-1}}$

$$\mathbb{E}_Q\left[\alpha_t(\sigma',\sigma'')\right]\le e^{-\omega t}\alpha_0(\sigma',\sigma'')+2C'\gamma^{-1}e^{-c\gamma^{-1}}\omega^{-1}$$

Hence one has, according to the dynamics defining the coupling, since if at some point $\sigma'=\sigma''$ then this remains true forever, that

$$Q\big(\sigma_t'\neq\sigma_t'' \text{ for some } t\ge \ln^2(\gamma^{-1})\big)\le Q\big(\sigma'_{\ln^2(\gamma^{-1})}\neq\sigma''_{\ln^2(\gamma^{-1})}\big)\le Q(\alpha_{\ln^2\gamma^{-1}}\ge 1)\le \mathbb{E}_Q\left[\alpha_{\ln^2(\gamma^{-1})}(\sigma',\sigma'')\right]$$

which goes to zero when γ goes to zero uniformly for $\sigma',\sigma''\in W_{r_1}(m_\beta)$.

After this preparation we are able to conclude from $\lim_{\gamma\to0}P_{\mu_{\gamma,B}}(\tau_\gamma<T^-)=0$ that also $\lim_{\gamma\to0}\sup_{\sigma\in W_{r_1}(m_\beta)}P_\sigma(\tau_\gamma<T^-)=0$ in the following way

$$\begin{aligned} &\big|P_{\sigma'}(\tau_\gamma<T^-)-P_{\sigma''}(\tau_\gamma<T^-)\big| = \Big|\int Q(d\sigma',d\sigma'')\big(\mathbb{1}_{\tau_\gamma<T^-}(\sigma')-\mathbb{1}_{\tau_\gamma<T^-}(\sigma'')\big)\Big|\\ &\le \int Q(d\sigma',d\sigma'')\mathbb{1}_{\tau(W_r(m_\beta)^c)>\ln^2(\gamma^{-1})}(\sigma')\mathbb{1}_{\tau(W_r(m_\beta)^c)>\ln^2(\gamma^{-1})}(\sigma')\big|\mathbb{1}_{\tau_\gamma<T^-}(\sigma')-\mathbb{1}_{\tau_\gamma<T^-}(\sigma'')\big|\\ &\quad +P_{\sigma'}\big(\tau(W_r(m_\beta)^c)<\ln^2(\gamma^{-1})\big)+P_{\sigma''}\big(\tau(W_r(m_\beta)^c)<\ln^2(\gamma^{-1})\big)\end{aligned}$$

The first summand is bounded by $2Q\big(\sigma_t'\neq\sigma_t''$ for some $t\ge\ln^2(\gamma^{-1})\big)$, because $\tau_\gamma>\tau(W_r(m_\beta)^c)$ and σ' and σ'' coalesce before $\tau(W_r(m_\beta)^c)$, hence $\tau_\gamma(\sigma')=\tau_\gamma(\sigma'')$; the second summand can be bounded by (4). Therefore, $\big|P_{\sigma'}(\tau_\gamma<T^-)-P_{\sigma''}(\tau_\gamma<T^-)\big|$ tends to zero uniformly in $\sigma',\sigma''\in W_{r_1}(m_\beta)$ when γ tends to zero. Finally, it remains to be observed that if we choose w.l.o.g. $r_1\le\theta/2$

$$\begin{aligned} \big|P_\sigma(\tau_\gamma<T^-)\big| &= \Big|\int \mu_{\gamma,W_{r_1}(m_\beta)}(d\sigma')P_\sigma(\tau_\gamma<T^-)\Big|\le \frac{\mu_\gamma(B)}{\mu_\gamma(W_{r_1}(m_\beta))}P_{\mu_{\gamma,B}}(\tau_\gamma<T^-)\\ &\quad+\sup_{\sigma'\in W_{r_1}(m_\beta)}\big|P_{\sigma'}(\tau_\gamma<T^-)-P_\sigma(\tau_\gamma<T^-)\big|\end{aligned}$$

this converges to zero. Note that $\mu_\gamma(B)/\mu_\gamma(W_{r_1}(m_\beta))=1+\mu_\gamma(B\setminus W_{r_1}(m_\beta))/\mu_\gamma(W_{r_1}(m_\beta))$ and the second summand goes to zero exponentially fast in γ^{-1}, because of the l.d.p. for μ_γ.

5.3 Step c)

For each $\sigma \in D^+$, by definition, the deterministic motion m^σ starting in σ enters in $W_{r_1/2}(m_\beta)$ in a finite time T_σ independent of γ. As the deterministic motion is continuous in the weak topology there exists a $\delta_\sigma \in (0, r_1/2)$ such that $\rho_{T_\sigma}(m_t^\sigma, D_+^c) \geq \delta_\sigma$ for all t. According to (5) there exists a c' such that

$$P_{\gamma,\sigma}\big(\rho_{T_\sigma}(m_t^\sigma, \sigma_t) \geq \delta_\sigma/2\big) \leq e^{-c'\gamma^{-1}}.$$

This estimate holds uniformly on compact sets. Then by Markov property we have that

$$\begin{aligned}
P_{\gamma,\sigma}\big(\tau_\gamma < T^-\big) &\leq P_{\gamma,\sigma}\big(\tau_\gamma < T^-,\ \rho_{T_\sigma}(m_t^\sigma, \sigma_t) < \delta_\sigma/2\big) + e^{-c'\gamma^{-1}} \\
&\leq P_{\gamma,\sigma}\big(T_\sigma \leq \tau_\gamma < T^-, \sigma_{T_\sigma} \in W_{r_1}(m_\beta)\big) + e^{-c'\gamma^{-1}} \\
&= \mathbb{E}_{\gamma,\sigma}\big[P_{\gamma,\sigma_{T_\sigma}}\big(\tau_\gamma < T^- - T_\sigma\big) 1\!\!1_{W_{r_1}(m_\beta)}(\sigma_{T_\sigma})\big] + e^{-c'\gamma^{-1}} \\
&\leq \sup_{\sigma \in W_{r_1}(m_\beta)} P_\sigma(\tau_\gamma < T^-) + e^{-c'\gamma^{-1}}
\end{aligned}$$

6 Upper bound

We are going to prove

Theorem 3. *For all $\delta > 0$ and all $\sigma \in D_+$*

$$\lim_{\gamma \to 0} P_{\gamma,\sigma}\big(\tau_\gamma > T^+\big) = 0$$

where T^+, τ_γ and r_- are as in theorem (1).

Repeating the arguments of the third step of the lower bound (section 5.3), one can reduce the problem to initial points in a small neighbourhood of m_β. The last part of the argument has to be changed in the following way

$$\begin{aligned}
P_{\gamma,\sigma}\big(\tau_\gamma > T^+\big) &\leq P_{\gamma,\sigma}\big(\tau_\gamma > T^+,\ \rho_{T_\sigma}(m_t^\sigma, \sigma_t) < \delta_\sigma/2\big) + e^{-c'\gamma^{-1}} \\
&\leq P_{\gamma,\sigma}\big(T_\sigma \leq \tau_\gamma, \tau_\gamma > T^+, \sigma_{T_\sigma} \in W_{r_1}(m_\beta)\big) + e^{-c'\gamma^{-1}} \\
&= \mathbb{E}_{\gamma,\sigma}\big[P_{\gamma,\sigma_{T_\sigma}}\big(\tau_\gamma > T^+ - T_\sigma\big) 1\!\!1_{W_{r_1}(m_\beta)}(\sigma_{T_\sigma})\big] + e^{-c'\gamma^{-1}} \\
&\leq \sup_{\sigma \in W_{r_1}(m_\beta)} P_\sigma(\tau_\gamma > T^+/2) + e^{-c'\gamma^{-1}}.
\end{aligned}$$

First, we want to construct the channel of exit from the minimum. This can be done in a finite time T and with a cost w.r.t. the rate functional $\mathcal{I}_T$ as near to the free energy of the saddle $\hat{m}$ as we want, say $\mathcal{I}_T - \mathcal{F}(\hat{m}) \leq \tilde{\delta}$, $0 < \tilde{\delta} < \delta$, uniformly in the initial point in G. For this we choose L^∞ balls $B(m_\beta)$, $B(\hat{m})$, $B(-m_\beta)$, $B(-m_\beta) \subset W_{r_-}(-m_\beta)$, around the three stationary points of G, so small that the cost of the interpolation between any two points picked up in one of them is less than $\tilde{\delta}/2$. This is possible because of a result in Comets (chap. 6, prop. VI.1 c) of [3]). Then it is known (see [2]) that there exists a manifold $\mathcal{W}$ linking the two pure phases and going through the instanton according to the solution of the mean field equation and to its time reversal. Thus we take in $B(\hat{m})$ two points m^- and m^+ respectively on the branches of $\mathcal{W}$ going to $-m_\beta$ and to m_β and we call m_β^- and m_β^+ two points in $B(-m_\beta)$ and $B(m_\beta)$ reached in a finite time by the deterministic evolution starting from m^- and by its time reversal starting from m^+. The cost for going from m_- to m_β^- is zero because we are following the deterministic flow. The cost for going from m_β^+ to m_+ is bounded by $\mathcal{F}(\hat{m})$ (see [3], prop VI.1 a,b)). Then the cost for the path going from m_β^+ to m^+, interpolating to m^- and arriving in m_β^- is less than $\mathcal{F}(\hat{m}) + \tilde{\delta}/2$. If the initial point belongs to $B(m_\beta)$ or to $B(\hat{m})$, then the result is trivial because we just interpolate to m_β^+ or to m^-, adding to the final cost a contribution less than $\tilde{\delta}/2$.

Now we want to show that there exists a finite time T' such that the deterministic flow starting from any point in G enters $B(\pm m_\beta)$ or $B(\hat{m})$ at least once within T'. We argue by contradiction: for any K there is a point m_K in G such that the flow starting from it stays away from the balls above for a time larger than K. Then there is a sequence of points m_K with diverging K. Since G is compact there is a converging subsequence; let us call $\bar{m} \in G$ the limit point. We know that the flow starting from $\bar{m}$ will converge to $\pm m_\beta$ or to $\hat{m}$ and then it will enter one of the balls B in a finite time; since the deterministic flow is continuous w.r.t. the initial data, we get the desired contradiction.

At this point we just follow the deterministic evolution until it reaches one of the balls and then we interpolate to the suitable point on $\mathcal{W}$. Denote the path resulting from this construction which starts in m by v^m.

There exists for $\delta' > 0$ and each $m \in G$ an r_m such that by Theorem IV.1 of [3]

$$\inf_{\sigma_0:\rho(m,\sigma_0)\le r_m} P_{\gamma,\sigma_0}\big(\rho_T(v^m,\sigma) < r_-/2\big) \ge e^{-\gamma^{-1}(\mathcal{I}_t(v^m)+\delta')} \ge e^{-\gamma^{-1}(\mathcal{F}(\hat{m})+\tilde{\delta}+\delta')}$$

where σ is the process started at σ_0. Note that $\sigma_T \in W_{r_-}(-m_\beta)$ for all σ as above. By compactness there exists a finite cover of G by balls of the from $(W_{r_{m_i}/2}(m_i))_{i=1}^I$. Choose $r' := \min_i r_{m_i}/2$ and one has

$$\inf_{\sigma_0\in G_{r'}} P_{\gamma,\sigma_0}\big(\min_i \rho_T(v^{m_i},\sigma) < r_-/2\big) \ge e^{-\gamma^{-1}(\mathcal{F}(\hat{m})+\tilde{\delta}+\delta')}$$

because if $\sigma_0 \in G_{r'}$, then there is some $m \in G$ such that $\rho(\sigma_0, m) < r'$. But m belongs to some ball $W_{r_{m_i}/2}(m_i)$, so $\rho(\sigma_0, m_i) < r' + r_{m_i}/2 \le r_{m_i}$ and we can apply the result above.

Hence divide the interval $[0, T^+]$ in intervals of length T

$$\begin{aligned}
P_{\gamma,\sigma}\big(\tau_\gamma > T^+, \tau(G^c_{r'}) > T^+\big) &\le P_{\gamma,\sigma}\big(\sigma_{kT} \notin W_{r_-}(-m_\beta) \text{ for all } k = 0,\dots,[T^+/T], \tau(G^c_{r'}) > T^+\big)\\
&\le \prod_{k=0}^{[T^+/T]} \sup_{\sigma\in G_{r'}} P_{\gamma,\sigma}\big(\sigma_{kT} \notin W_{r_-}(-m_\beta)\big) \le \big(1 - \inf_{\sigma\in G_{r'}} P_{\gamma,\sigma}\big(\min_i \rho_T(v^{m_i},\sigma) < \frac{r_-}{2}\big)\big)^{T^+/T}\\
&\le \big(1 - e^{-\gamma^{-1}(\mathcal{F}(\hat{m})+\tilde{\delta}+\delta')}\big)^{e^{\gamma^{-1}(\mathcal{F}(\hat{m})+\delta)}/T} \to 0
\end{aligned}$$

provided that $\delta' < \delta - \tilde{\delta}$.

Let $r < r'$, so that $W_r(m_\beta) \subset G_{r'}$. We have for $\hat{\delta} > 0$

$$\begin{aligned}
P_{\mu_{\gamma,W_r(m_\beta)}}(\sigma_T \in \partial_{r'}G) &\le \frac{P_{\mu_\gamma}(\sigma_T \in \partial_{r'}G)}{\mu_\gamma(W_r(m_\beta))}\\
&\le \frac{e^{-\gamma^{-1}(\mathcal{F}(\hat{m})+\sigma-\hat{\delta})}}{e^{-\gamma^{-1}(\inf_{m\in W_r(m_\beta)}\mathcal{F}(m)+\hat{\delta})}}
\end{aligned}$$

and we note that $\inf_{m\in W_r(m_\beta)} \mathcal{F}(m) = 0$. Here we have used the stationarity of the Gibbs measure w.r.t. the Glauber dynamics and the large deviation principle for μ_γ. Given $r'/2$ there is a partition of $[-L, L]$ in subintervals I and a number θ' such that if $\sum_I |\int_I m_1 - m_2| < \theta'$ then $\rho(m_1, m_2) < r'/2$. Let Δt be a small time interval, say $\Delta t = \gamma^a$, $a > 0$. Since the jump rates are bounded by a constant c_M, we can estimate the probability of having more than N jumps in a time interval Δt by

$$\frac{(2L\gamma^{-1}c_M\Delta t)N}{N!} \approx \exp\{-N \ln \frac{N}{2Lc_M\gamma^{-1}\Delta t} + \dots\}$$

If we choose $N = \gamma^{-1}/(\ln\gamma^{-1})^b$, $0 < b < 1$, then the above probability goes to zero, when γ vanishes, faster than $e^{-\gamma^{-1}c}$, for any possible c. If the number of spin flips is not greater than N, the largest change in magnetization in a coarse cell I is $(2\gamma/|I|)N = 2/(|I|(\ln\gamma^{-1})^b)$. It follows that

$$\sum_I |\int_I \sigma_{t+\Delta t} - \sigma_t| < \frac{4L}{|I|(\ln\gamma^{-1})^b} < \theta'$$

provided γ sufficiently small. For any realization $\bar{\sigma}_t$ such that $\bar{\sigma}_0$ is distributed according to $\mu_{\gamma,W_r(m_\beta)}$ and that there is some $T>0$ such that $\bar{\sigma}_T \notin G_{r'}$, we define the sequence $\sigma_i := \bar{\sigma}_{i\Delta t}$. With overwhelming probability $\rho(\sigma_{i+1},\sigma_i) < r'/2$; since the process starts in $W_r(m_\beta)$ and after a while enters $G^c_{r'}$ we can conclude that there is l such that $\sigma_l \in G_{r'}$ but $\sigma_{l+1} \notin G_{r'}$. It follows that $\sigma_l \in \partial_{r'}G$, because if σ_l belonged to G then the distance between σ_{l+1} and G would be smaller than $r'/2$ implying that $\sigma_{l+1} \in G_{r'}$.

Let $M = [T^+/\Delta t]$, then

$$P_{\mu_{\gamma,W_r(m_\beta)}}(\tau(G^c_{r'}) < T^+) \leq (M+1)\frac{(2Lc_M\gamma^{-1}\Delta t)^N}{N!} + \sum_{i=1}^{M} P_{\mu_{\gamma,W_r(m_\beta)}}(\sigma_i \in \partial_{r'}G)$$

which goes to zero, because without loss of generality we can assume that $\sigma > \delta$. Applying the second step as before (see section 5.2), we obtain that there exists $r_1 > 0$ such that

$$\sup_{\sigma\in W_{r_1}(m_\beta)} P_{\gamma,\sigma}(\tau(G^c_{r'}) < T^+) \to 0.$$

So we just derived that

$$\sup_{\sigma\in W_{r_1}(m_\beta)} P_{\gamma,\sigma}(\tau_\gamma > T+) \to 0.$$

7 Conclusions

We proved that the time needed to observe a tunneling between two equally stable states of a ferromagnetic system is given in terms of the free energy gap with the first excited state, namely the instanton.

Acknowledgements. The authors are grateful to E. Presutti for illuminating discussions and suggestions, especially about the coupling method used in the proof.

References

[1] G. Bellettini, A. De Masi, E. Presutti *Energy levels of a nonlocal functional*, J. Math. Phys. **(46)** (2005), no. 8, 083302, 31pp.

[2] P. Buttà, A. De Masi, E. Rosatelli *Slow motion and metastability for a nonlocal evolution equation*, J. Statist. Physics. **(112)** (2003), no. 3-4, 709–764.

[3] F. Comets, *Nucleation for a long range magnetic model*, Ann. Inst. H. Poincaré Probab. Statist. **(23)** (1987), no. 2, 135–178.

[4] G. Manzi, *The Wulff problem For Diffuse Interfaces*, to appear in Z. Anal. Anwend.

[5] E. Olivieri, M.E. Vares, *Large deviations and metastability*, Cambridge University Press, (2004).