Ann. Henri Poincaré 13 (2012), 1291–1303

1424-0637/12/051291-13
published online December 1, 2011
DOI 10.1007/s00023-011-0148-6

Annales Henri Poincaré

Decision Making Times in Mean-Field Dynamic Ising Model

Yuri Bakhtin

Abstract. We consider a dynamic mean-field ferromagnetic model in the low-temperature regime in the neighborhood of the zero magnetization state. We study the random time it takes for the system to make a decision, i.e., to exit the neighborhood of the unstable equilibrium and approach one of the two stable equilibrium points. We prove a limit theorem for the distribution of this random time in the thermodynamic limit.

1. Introduction

The asymptotic properties of exit from a small neighborhood of an unstable equilibrium of a dynamical system under small white noise perturbation were first studied rigorously in [14]. It was shown that as the noise intensity ε goes to zero, the exit time τ_ε behaves roughly as $a^{-1}\ln\varepsilon^{-1}$, where a is the local expansion rate (Lyapunov exponent) at the equilibrium point. After that, several authors worked on ramifications of these results, see [2–5, 8, 11].

It was shown in [8] and rediscovered in [2, 3] that $\tau_\varepsilon - a^{-1}\ln\varepsilon^{-1}$ converges to a limiting distribution that is a dilation and translation of $\ln|G|^{-1}$, where G is a standard Gaussian random variable.

Understanding distributional asymptotics for the exit time was pivotal in describing the vanishing noise asymptotics for noisy heteroclinic networks, see [3, 4]. These systems occur naturally in the context of neural dynamics and sequential decision making, see, e.g., [17] and references therein. Exit times for diffusion models have been used in psychology to describe reaction times in decision tasks, see [18] and references therein, and it is natural to ask if the limiting behavior of exit times described above is reproduced in statistical mechanics models of neural computation.

In fact, the process of escape from an unstable equilibrium in statistical mechanics models, especially in the context of phase separation, has been studied extensively and well understood, see, e.g., a recent book [16], papers

[6,9,10], and references therein. The literature contains subtle results on the geometry of clusters formed in the course of phase separation, and cases where the instability is quadratic or quartic have been considered. However, the infinite volume distributional limit of the correction to the main logarithmic term has not been addressed in the literature. It cannot be derived directly from previous results (to the best of our knowledge), and our goal is to close this gap.

In this paper we choose to study one of the simplest statistical mechanics models, the dynamic mean field ferromagnetic model, also known as the Curie–Weiss model, in the low temperature regime with two minima of free energy. We start the evolution of the system at the completely disordered state with zero magnetization, where the number of plus spins equals the number of minus spins. We stop the dynamics as soon as magnetization enters a neighborhood of one of the stable equilibrium values and interpret that event as a decision made by the system between the two choices. We show that as the number N of spin variables (representing individual neurons in the neural computation context) goes to infinity, the exit time behaves as $\ln N$ and the correction to the main term converges to an affine transformation of $\ln |G|^{-1}$, thus reproducing the above result for the diffusion in the neighborhood of an unstable equilibrium.

We have to make some remarks concerning the proof. It is well known that systems like the one under consideration can be approximated by diffusion processes with diffusion coefficient of order of $N^{-1/2}$. This suggests a natural idea of a seemingly easy proof. However, if we try to use this approximation directly, we would have to prove that two sequences of probability distributions are asymptotic to each other (since the basic model and the diffusion both depend on N). This convergence, which is more general than weak convergence of a sequence of measures to a fixed measure (e.g., none of the two sequences has to be tight), is known in the literature as merging or proximity of two sequences of measures, see, e.g., [7] and references therein. Unfortunately, currently there is no contentful theory of merging in trajectory spaces.

Of course, we can modify this approach and try to rescale the solution of our system appropriately, obtain a functional limit theorem for it, and derive the asymptotics for the exit time from the exit asymptotics for the limiting diffusion (the endpoints of the domain must be appropriately rescaled, too). However, there is still an obstacle. The standard set of tools of functional limit theorems for diffusions is essentially useful for proving weak convergence of processes within a finite time horizon, whereas the exit times we are dealing with require the time scale that grows at least logarithmically in N. That is why the standard diffusion approximation results do not imply the desired asymptotics of exit times straightforwardly.

So, to prove the result we do use the technique of diffusion approximation and mimic the derivation of the asymptotics for the exit distribution in the diffusion case, but it takes more effort to finish the proof due to the limitations of the existing weak convergence techniques.

2. The Model and the Main Result

First let us recall the mean-field ferromagnetic equilibrium Ising model also known as Curie–Weiss model, see [12, Sect. IV.4]. Let us fix a large number N and consider N spin variables. Each variable $X_k, k = 1, \ldots, N$ takes values ± 1, and the energy assigned to a configuration $(x_k)_{k=1}^N$ is given by

$$E(x) = -\frac{1}{2N} \sum_{i,j=1}^{N} x_i x_j.$$

We then can fix an inverse temperature value $\beta > 0$ and consider the Boltzmann–Gibbs distribution defined by E and β:

$$\mathsf{P}_N\{X_k = x_k, \quad k = 1, \ldots N\} = \frac{e^{-\beta E(x)}}{Z_N},$$

where

$$Z_N = \sum_{x \in \{-1,1\}^N} e^{-\beta E(x)}$$

is the partition function.

Since there is no geometry involved in this mean-field model and the strength of interactions between two spins is the same for all pairs of spins, one can describe the macroscopic behavior of the system by a single variable called magnetization,

$$M(x) = \frac{1}{N} \sum_i x_i \in [-1, 1].$$

Notice that

$$E(x) = -NM^2(x)/2$$

Therefore,

$$\mathsf{P}_N\left\{M(X) = \frac{n}{N}\right\} = \frac{1}{Z_N} \binom{N}{(N+n)/2} e^{N \frac{\beta}{2} \cdot \left(\frac{n}{N}\right)^2}$$

if $(N + n)/2$ is integer.

Recall that (see, e.g., [12, Lemma I.3.2]) uniformly in $k = 0, \ldots, n$,

$$\frac{1}{N} \ln \binom{N}{k} = h\left(\frac{k}{N}\right) + O\left(\frac{\ln N}{N}\right),$$

where

$$h(x) = -x \ln x - (1 - x) \ln(1 - x), \quad x \in [0, 1],$$

is the entropy of the Bernoulli distribution with probabilities x and $1 - x$. Therefore,

$$\mathsf{P}_N\left\{M(X) = \frac{n}{N}\right\} = \frac{1}{Z_N} e^{-NF(n/N) + O(\ln N)},$$

where the free energy per spin F is defined by

$$F(m) = -\frac{\beta}{2}m^2 - h\left(\frac{1}{2} + \frac{m}{2}\right).$$

It is easy to show that the sequence of distributions $\mathsf{P}_N\{M(X) \in \cdot\}$ satisfies a large deviation principle on $[0,1]$ with rate function J given by

$$J(m) = F(m) - \min_{[0,1]} F.$$

Differentiating F, we see that the minimizers of F satisfy

$$\beta m = \frac{1}{2}\ln\frac{1+m}{1-m}, \tag{2.1}$$

and, as elementary analysis shows, (i) for $\beta < 1$, a unique minimizer of F is $m = 0$ (corresponding to completely disordered case), and $F''(0) > 0$ so that F is approximately quadratic in the neighborhood of the minimizer; (ii) for $\beta > 1$, there are two minimizers $m = \pm m_*$, for some $m^* > 0$; (iii) if $\beta = 1$ then 0 is still a unique minimizer, but contrary to the first case, $F''(0) = 0$, and the leading term in the Taylor expansion of F at 0 is order 4.

In this paper, we are concerned with the low-temperature case (ii). In that situation, point 0 is also a solution of (2.1), but it is an unstable equilibrium of the system being the local maximum of the free energy F. We are going to consider Glauber dynamics, a stochastic process of spin flips compatible with Curie–Weiss model, and study it in the neighborhood of the unstable equilibrium in the case $\beta > 1$.

We must study a $\{-1,+1\}^N$-valued Markov process with intensities of spin flips $c_i(x), i \in \{1,\dots,N\}, x \in \{-1,+1\}^N$ defined by

$$\mathsf{P}\{X_i(t+\Delta t) \neq X_i(t) | X(t) = x\} = c_i(x)\Delta t + o(\Delta t), \quad \Delta t \downarrow 0.$$

If we want the process X to be reversible w.r.t. the Gibbs distribution P_N, it is required that

$$c_i(x)\exp\{-\beta E(x)\}$$

does not depend on x_i, see [15, Sect. IV.2]. Equivalently, it is required that

$$c_i(x)\exp\{\beta x_i M(x)\}$$

does not depend on x_i. There are many choices for rates c_i, and there is no physical reason to prefer one of them to others. In this paper we will work with

$$c_i(x) = \exp\{-\beta x_i M(x)\}, \tag{2.2}$$

although our results should hold for a variety of other choices of $c_i(x)$.

Notice that if the spin x_i is aligned with magnetization $M(x)$, then the resulting flipping rate of i-th spin is lower than that in the opposite situation where x_i is misaligned with $M(x)$. This is the result of the ferromagnetic nature of the model which favors configurations with most spins aligned with each other.

Suppose now that we observe only the magnetization, or, equivalently, the number of +1-spins. Flipping a −1 spin means then a transition from the current magnetization m to $m + 2/N$. Since the number of -1 spins equals $N(1-m)/2$, we see that the total transition rate $m \mapsto m + 2/N$ is given by $\lambda_+(m, N)$, where

$$\lambda_+(m, N) = N\frac{1-m}{2}\exp\{\beta m\}.$$

Flipping a $+1$ spin means a transition from the current magnetization m to $m - 2/N$. Since the number of $+1$ spins equals $N(1+m)/2$, we see that the total transition rate $m \mapsto m - 2/N$ is given by $\lambda_-(m, N)$, where

$$\lambda_-(m, N) = N\frac{1+m}{2}\exp\{-\beta m\}.$$

Let us consider the Markov process M_N describing the evolution of magnetization in the above model and set $M_N(0) = 0$ (this means that N has to be even, but this is not a really important restriction).

It is clear that M_N will spend some time in the neighborhood of 0 and then it will escape that neighborhood and head towards one of the minima of free energy, $\pm m_*$. We can interpret the exit in each of these directions as the decision made by the system. We set a threshold level $R \in (0, m_*)$ and as soon as M_N exceeds R in absolute value, we claim that the system has made the decision. The choice of one of the two alternatives is encoded by the sign of M_N at that time. Our main result describes the asymptotics of the random time it takes to reach the threshold R starting from the completely disordered state with zero magnetization. According to the interpretation above, this time can be viewed as the decision making time for the situation where the initial state is a completely unbiased indecisive state.

More formally, for any $R \in (0, m_*)$ we introduce

$$\tau_N(R) = \inf\{t : |M_N(t)| \geq R\}.$$

Our main result describes the joint asymptotic behavior of random variables $\tau_N(R)$ and $\operatorname{sgn} M_N(\tau_N)$. To state it, we need more notation. For $m \in [-1, 1]$, we denote

$$b(m) = \frac{2}{N}(\lambda_+(m, N) - \lambda_-(m, N)) = (1-m)e^{\beta m} - (1+m)e^{-\beta m}, \quad (2.3)$$

introduce $a = b'(0) = 2\beta - 2 > 0$ and $Q(x) = b(x) - ax, x \in \mathbb{R}$, and define

$$D(R) = K(R) + \frac{\ln R}{a} + \frac{\ln(a/2)}{2a}, \quad R \in \mathbb{R},$$

where

$$K(R) = -\int_0^R \frac{Q(x)}{axb(x)}\mathrm{d}x \in \mathbb{R}, \quad R \in \mathbb{R}. \quad (2.4)$$

Theorem 2.1. *For any $R \in (0, m^*)$, as $N \to \infty$*

$$\left(\operatorname{sgn} M_N(\tau_N(R)),\ \tau_N(R) - \frac{1}{2a}\ln N\right) \xrightarrow{distr} \left(\operatorname{sgn} G, -\frac{1}{a}\ln|G| + D(R)\right),$$

where G is a standard Gaussian random variable.

3. Proof

The proof is based on the theory of Markov processes, martingales and their convergence. We refer to [13] as an excellent source on the relevant background.

The form of rates λ_+, λ_- implies that the generator L_N of the magnetization process M_N is given by

$$L_N f(m) = \frac{N}{2}\left[(1-m)e^{\beta m}\left(f(m+\frac{2}{N}) - f(m)\right)\right. \\ \left. +(1+m)e^{-\beta m}\left(f(m-\frac{2}{N}) - f(m)\right)\right].$$

In particular,

$$f(M_N(t)) - \int_0^t L_N f(M_N(s))\mathrm{d}s$$

has to be a martingale for any f, see Proposition 1.7 in [13, Chapter 4]. Choosing $f(x) \equiv x$ on $[-1,1]$, we obtain that

$$Z_N(t) = M_N(t) - \int_0^t b(M_N(s))\mathrm{d}s$$

is a bounded variation cadlag martingale, where $b(\cdot)$ is defined in (2.3), so that it plays the role of drift coefficient that drives the deterministic component of the process. Notice that zeros of $b(m)$ coincide with solutions of equation (2.1), so that for any point $x \in (0, m_*)$, $\lim_{t\to+\infty} S^t x = m_*$ and $\lim_{t\to-\infty} S^t x = 0$, and for any point $x \in (-m_*, 0)$, $\lim_{t\to+\infty} S^t x = -m_*$ and $\lim_{t\to-\infty} S^t x = 0$, where S is the flow generated by b: $S^0 x = x$ and $\frac{\mathrm{d}}{\mathrm{d}t} S^t x = b(S^t x)$.

Using the representation

$$b(m) = am + Q(m), \quad m \in [-1,1], \tag{3.1}$$

where $a = b'(0) = 2\beta - 2$ and $|Q(m)| \le Km^2$ for $m \in [-1,1]$, we can write

$$M_N(t) = a\int_0^t M_N(s)\mathrm{d}s + \int_0^t Q(M_N(s))\mathrm{d}s + \int_0^t \mathrm{d}Z_N(s),$$

We can now use variation of constants to write

$$M_N(t) = e^{at}\int_0^t e^{-as}\mathrm{d}Z_N + e^{at}\int_0^t e^{-as}Q(M_N(s))\mathrm{d}s, \tag{3.2}$$

where the integral w.r.t. Z_N is understood as Lebesgue–Stieltjes integral.

Lemma 3.1. *Suppose there is a sequence of stopping times θ_N satisfying*

$$\theta_N \xrightarrow{\mathsf{P}} \infty, \quad N \to \infty, \tag{3.3}$$

and

$$\sup_{t\le\theta_N} |M_N(t)| \le N^{-\gamma}, \tag{3.4}$$

for some $\gamma > 0$ *and all* N*. Then, as* $N \to \infty, I_N = N^{1/2} \int_0^{\theta_N} e^{-as} \mathrm{d}Z_N(s)$ *converges in distribution to* $\mathcal{N}(0, 2a^{-1})$*.*

Proof. Let us introduce a process

$$V_N(t) = \int\limits_0^{t\wedge\theta_N} e^{-as} \mathrm{d}Z_N(s). \tag{3.5}$$

It is a martingale with

$$[V_N]_t = \int\limits_0^{t\wedge\theta_N} e^{-2as} \mathrm{d}[Z_N]_s,$$

where square brackets denote the quadratic variation process. Let us define $U_N(s) = V_N(f(s))$, where

$$f(s) = -\frac{\ln(1 - as/2)}{2a}, \quad s \in [0, 2a^{-1}].$$

We need the following statement which is a specific case of Theorem 1.4 in [13, Chapter 7] (see also bibliographical notes therein for the history of this theorem and related results):

Theorem 3.1. *For each* $N \in \mathbb{N}$*, let* U_N *be a martingale w.r.t. some filtration, with cadlag paths and* $U_N(0) = 0$*. Suppose for all* $t \in [0, 2a^{-1}], A_N(t) = [U_N]_t$ *satisfies*

$$N A_N(t) \xrightarrow{\mathsf{P}} t, \quad N \to \infty,$$

and

$$\lim_{N\to\infty} \mathsf{E}\left[N^{1/2} \sup_{0\le t\le 2a^{-1}} |U_N(t) - U_N(t-)|\right] = 0. \tag{3.6}$$

Then, as $N \to \infty, N^{1/2} U_N$ *converges in distribution in the Skorokhod topology to the standard Wiener process on* $[0, 2a^{-1}]$*.*

In our case, condition (3.6) is fulfilled automatically since all the jumps of U_N are bounded by $2N^{-1}$ in absolute value. All the jumps of Z_N are equal to $2N^{-1}$ in absolute value, so that

$$[Z_N](t) = \frac{4}{N^2} B_N(t), \quad t \ge 0,$$

where $B_N(t)$ denotes the number of jumps the process Z_N makes up to time t. Next,

$$A_N(t) = \frac{4}{N^2} \sum_{\substack{s: s\le f(t)\wedge\theta_N \\ M_N(s)\ne M_N(s-)}} e^{-2as},$$

and

$$NA_N(t) \xrightarrow{\mathsf{P}} 4 \int_0^{f(t)} e^{-2as} \mathrm{d}s = t. \tag{3.7}$$

This is a consequence of the following result:

Lemma 3.2. *For any non-increasing function $g : \mathbb{R}_+ \to \mathbb{R}_+$ and any $t \geq 0$,*

$$\frac{1}{N} \sum_{\substack{s: s \leq t \wedge \theta_N \\ M_N(s) \neq M_N(s-)}} g(s) \xrightarrow{\mathsf{P}} \int_0^t g(s)\mathrm{d}s.$$

Proof. It is sufficient to assume that $g(s) = \mathbf{1}_{s \in [0,h]}$ for some $h > 0$ since one can use linear combinations of functions of this form to approximate any non-increasing function. We have to show that for all t,

$$\frac{1}{N} B_N(t \wedge h \wedge \theta_N) \xrightarrow{\mathsf{P}} t \wedge h.$$

Due to (3.3), we can restrict ourselves to the high probability event $\{t < \theta_N\}$. Let $(s_{N,i})$ be the increasing sequence of times of jumps for each N, i.e., $M_N(s_{N,i}) \neq M_N(s_{N,i}-)$. Clearly, conditioned on $M_N(s_{N,i})$, the spacing random variable $s_{N,i+1} - s_{N,i}$ has exponential distribution with parameter

$$\lambda_+(M_N(s_{N,i}), N) + \lambda_N(M_N(s_{N,i}), N) = \lambda(M_N(s_{N,i}), N),$$

where

$$\lambda(m, N) = \lambda_-(m, N) + \lambda_+(m, N).$$

Therefore, using (3.4) we see that the number of points s_i not exceeding $t \wedge h$ is between two Poisson processes with intensities $\lambda(0, N) = N$ and $\lambda(N^{-\beta}, N)$ evaluated at time $t \wedge h$. Our claim follows since $\lambda(N^{-\beta}, N)/N \to 1$. □

Now all the conditions of Theorem 3.1 have been verified and we conclude that $N^{1/2} U_N$ converges in distribution in Skorokhod topology to the Wiener process on $[0, 2a^{-1}]$. Therefore, $I_N = N^{1/2} U_N(2a^{-1})$ converges in distribution to $\mathcal{N}(0, 2a^{-1})$, and the proof of Lemma 3.1 is complete. □

Let us now check that the conditions of Lemma 3.1 hold for $\theta_N = \inf\{t : |M_N(t)| \geq N^{-\gamma}\}$. Notice that condition (3.4) is satisfied automatically.

Lemma 3.3. *If $\frac{1}{4} < \gamma < \frac{1}{2}$, then (3.3) holds.*

Proof. Let us take any $T > 0$. On $\{\theta_N < T\}$

$$\begin{aligned} N^{-\gamma} &\leq |M_N(\theta_N)| \\ &= e^{a\theta_N} \left| V_N(\theta_N) + \int_0^{\theta_N} e^{-as} Q(M_N(s))\mathrm{d}s \right| \\ &\leq e^{aT}(|V_N(\theta_N)| + a^{-1} K N^{-2\gamma}), \end{aligned} \tag{3.8}$$

where V has been introduced in (3.5). The quadratic variation of the martingale V_N was computed in the proof of Lemma 3.1, and we can conclude that

$$\mathsf{E}[V_N]_{\theta_N} \le 2a^{-1}N^{-1}.$$

Therefore, by the Chebyshev inequality and the fact that $V_N^2 - [V_N]$ is a martingale (see Proposition 6.1 in [13, Chapter 2]),

$$\mathsf{P}\{|V_N(\theta_N)| > N^{-\gamma/2-1/4}\} \le \frac{\mathsf{E}V_N^2(\theta_N)}{N^{-\gamma-1/2}} \le \frac{\mathsf{E}[V_N]_{\theta_N}}{N^{-\gamma-1/2}} \le \frac{2a^{-1}N^{-1}}{N^{-\gamma-1/2}} \to 0.$$

Now, on $\{\theta_N < T\} \cap \{|V_N(\theta_N)| \le N^{-\gamma/2-1/4}\}$, we have

$$N^{-\gamma} \le e^{aT}(N^{-\gamma/2-1/4} + a^{-1}KN^{-2\gamma})$$

which is impossible for large N under our assumptions. We conclude that $P\{\tau_N < T\} \to 0$, and the lemma follows. □

We are ready to describe the asymptotics of the exit from $[-N^{-\gamma}, N^{-\gamma}]$.

Lemma 3.4. *If* $1/4 < \gamma < 1/2$, *then*

$$\left(\operatorname{sgn} M_N(\theta_N),\ \theta_N - \frac{1/2-\gamma}{a}\ln N\right) \xrightarrow{distr} \left(\operatorname{sgn} H, -\frac{1}{a}\ln|H|\right),$$

where $\mathrm{Law}(H) = \mathcal{N}(0, 2a^{-1})$.

Proof. Considering the process M_N at time θ_N and using (3.2), we obtain

$$\begin{aligned}\theta_N &= \frac{1}{a}\ln\frac{\left\lceil N^{-\gamma}\cdot\frac{N}{2}\right\rceil\cdot\frac{2}{N}}{\left|V_N(\theta_N) + \int_0^{\theta_N} e^{-as}Q(M_N(s))\mathrm{d}s\right|}\\ &= \frac{1/2-\gamma}{a}\ln N - \frac{1}{a}\ln\left|I_N + N^{1/2}\int_0^{\theta_N} e^{-as}Q(M_N(s))\mathrm{d}s\right| + o(1).\end{aligned}$$

Also,

$$\operatorname{sgn} M_N(\theta_N) = \operatorname{sgn}\left(I_N + N^{1/2}\int_0^{\theta_N} e^{-as}Q(M_N(s))\mathrm{d}s\right).$$

Since

$$N^{1/2}\int_0^{\theta_N} e^{-as}Q(M_N(s))\mathrm{d}s \le \frac{N^{1/2}K}{a}N^{-2\gamma} \to 0,$$

the desired statement follows from Lemma 3.1. □

Let us now study the exit of M_N from an interval $[-r, r]$ where r is a small number that does not depend on N. We define $Y_N(t) = M_N(t+\theta_N)$ and for any $r \in (0, R)$ we define

$$\nu_N(r) = \inf\{t \ge 0:\ |Y_N(t)| = r\}.$$

We are going to compare the evolution of the magnetization process to the deterministic trajectory of the flow S^t generated by the drift b.

For any $\delta > 0$ we introduce $t(\delta, r)$ as the only solution t of $S^t\delta = r$, i.e., it is the time it takes for the solution of ODE $\dot{x} = b(x)$ to travel from δ to r.

Lemma 3.5. *For any $r > 0$,*

$$\lim_{\delta \to 0} \left(t(\delta, r) - \frac{1}{a} \ln \frac{r}{\delta} \right) = K(r),$$

where $K(\cdot)$ was defined in (2.4).

Proof. By the basic formula for solutions of autonomous ODE's (see, e.g., [1, Sect. 1.2]):

$$\begin{aligned} t(\delta, r) &= \int_\delta^r \frac{\mathrm{d}x}{b(x)} = \int_\delta^r \left(\frac{1}{b(x)} - \frac{1}{ax} \right) \mathrm{d}x + \int_\delta^r \frac{\mathrm{d}x}{ax} \\ &= \int_\delta^r \frac{1}{b(x)} \frac{ax - b(x)}{ax} \mathrm{d}x + \frac{1}{a}(\ln r - \ln \delta), \end{aligned}$$

and the lemma follows. □

Lemma 3.6. *There is $r_0 > 0$ such that*

$$\sup_{0 \le t \le t(|Y_N(0)|, r_0)} |Y_N(t) - S^t(Y_N(0))| \xrightarrow{\mathsf{P}} 0, \quad N \to \infty.$$

Proof. Denote $\Delta_N(t) = Y_N(t) - S^t Y_N(0)$. Since

$$Y_N(t) = Y_N(0) + \int_0^t b(Y_N(s)) \mathrm{d}s + Z'_N(t),$$

where $Z'_N(t) = Z_N(t + \theta_N) - Z_N(\theta_N)$ is a martingale, and

$$S^t Y_N(0) = Y_N(0) + \int_0^t b(S^s Y_N(0)) \mathrm{d}s,$$

we see that for any $r \in (0, m_*)$ and any $t \in (0, t(|Y_N(0)|, r))$,

$$|\Delta_N(t \wedge \nu_N(r))| \le L(r) \int_0^{t \wedge \nu_N(r)} |\Delta_N(s)| \mathrm{d}s + \sup_{s \le t \wedge \nu_N(r)} |Z'_N(s)|,$$

where $L(r)$ is the Lipschitz constant of b on $[-r, r]$.

Since $[Z'_N]_t = 4N^{-2}(B_N(\theta_N + t) - B_N(\theta_N))$, and the number of jumps between θ_N and $\theta_N + t$ is stochastically dominated by the increment of the Poisson process with intensity N, we have

$$\mathsf{P} \left\{ \sup_{s \le t(|Y_N(0)|, r) \wedge \nu_N(r)} |Z'_N(s)| > N^{-\delta} \right\} \to 0$$

if $\delta < 1/2$. On the complementary event, applying Gronwall's inequality, we obtain for some constant $C > 0$,

$$|\Delta_N(t \wedge \nu_n(r))| \leq e^{L(r)t(|Y_N(0)|,r)} N^{-\delta} \leq C N^{\gamma L(r)/a} N^{-\delta}. \tag{3.9}$$

We can choose r to be so small that $L(r)$ is close to a enough to ensure that $\gamma L(r)/a < 1/2$. Consequently, we can choose $\delta < 1/2$ such that $\rho = \delta - \gamma L(r)/a > 0$, and the r.h.s. of (3.9) converges to 0. For any $r_0 \in (0, r)$ we conclude then that $\mathsf{P}\{\nu_N(r) < t(|Y_N(0)|, r_0)\} \to 0$, and

$$\mathsf{P}\left\{\nu_N(r) \geq t(|Y_N(0)|, r_0); \sup_{s \leq t(|Y_N(0)|, r_0)} |\Delta_N(s)| > N^{-\rho}\right\} \to 0.$$

which completes the proof of the lemma. $\square$

We can now combine the results of Lemmas 3.4 and 3.6.

Lemma 3.7. *For any $r \in (0, r_0)$,*

$$\left(\operatorname{sgn} M_N(\tau_N(r)), \tau_N(r) - \frac{1}{2a} \ln N\right) \stackrel{distr}{\longrightarrow} \left(\operatorname{sgn} H, -\frac{1}{a} \ln |H| + \frac{\ln r}{a} + K(r)\right),$$

Proof. Obviously,

$$\tau_N(r) = \theta_N + \nu_N(r). \tag{3.10}$$

Lemma 3.6 implies that

$$\nu_N(r) - t(|Y_N(0)|, r) \stackrel{\mathsf{P}}{\to} 0.$$

This together with Lemma 3.5 implies

$$\nu_N(r) - \frac{1}{a} \ln \frac{r}{N^{-\gamma}} - K(r) \stackrel{\mathsf{P}}{\to} 0. \tag{3.11}$$

The lemma follows now from (3.10),(3.11), and Lemma 3.4. $\square$

The next result follows from the same considerations as Lemma 3.6, except that it is easier since we consider a finite time horizon.

Lemma 3.8. *Let r be as in the last lemma. Let $Y_N(t) = M_N(\tau_N(r) + t)$. Then, for any $T > 0$,*

$$\sup_{0 \leq t \leq T} |Y_N(t) - S^t(Y_N(0))| \stackrel{\mathsf{P}}{\to} 0, \quad N \to \infty.$$

This lemma means that after $\tau_N(r)$ the process essentially follows the deterministic trajectory. Since $H = 2^{1/2} a^{-1/2} G$, where G is standard Gaussian, our main result is a direct consequence of Lemmas 3.7 and 3.8. In fact, it extends Lemma 3.7 since the latter is valid only for sufficiently small values of threshold, whereas our main result applies to any $R \in (0, m_*)$.

Acknowledgements

The author is grateful to Mark Meckes for pointing to some literature on merging of probability measures. The support from National Science Foundation through CAREER grant DMS-0742424 is gratefully acknowledged.

References

[1] Arnold, V.I.: Ordinary differential equations (Translated from the Russian by Roger Cooke, Second printing of the 1992 edition). Universitext. Springer, Berlin (2006)

[2] Bakhtin, Y.: Exit asymptotics for small diffusion about an unstable equilibrium. Stoch. Process. Appl. **118**(5), 839–851 (2008)

[3] Bakhtin, Y.: Noisy heteroclinic networks. Probab. Theory Relat. Fields, in print (2010)

[4] Bakhtin, Y.: Small noise limit for diffusions near heteroclinic networks. Dyn. Syst., in print (2010)

[5] Buterakos, L.A.: The exit time distribution for small random perturbations of dynamical systems with a repulsive type stationary point. PhD thesis, Virginia Polytechnic Institute and State University (2003)

[6] Caderoni, P., Pellegrinotti, A., Presutti, E., Vares, M.E.: Transient bimodality in interacting particle systems. J. Stat. Phys. **55**(3–4), 523–577 (1989)

[7] Davydov, Y., Rotar, V.: On asymptotic proximity of distributions. J. Theor. Probab. **22**(1), 82–98 (2009)

[8] Day, M.V.: On the exit law from saddle points. Stoch. Process. Appl. **60**(2), 287–311 (1995)

[9] De Masi, A., Orlandi, E., Presutti, E., Triolo, L.: Glauber evolution with Kac potentials. III. Spinodal decomposition. Nonlinearity **9**(1), 53–114 (1996)

[10] De Masi, A., Pellegrinotti, A., Presutti, E., Vares, M.E.: Spatial patterns when phases separate in an interacting particle system. Ann. Probab. **22**(1), 334–371 (1994)

[11] Eizenberg, A.: The exit distributions for small random perturbations of dynamical systems with a repulsive type stationary point. Stochastics **12**(3–4), 251–275 (1984)

[12] Ellis, R.S.: Entropy, large deviations, and statistical mechanics. In: Classics in Mathematics. Springer, Berlin (Reprint of the 1985 original, 2006)

[13] Ethier, S.N., Kurtz, Thomas G.: Markov processes. Characterization and convergence. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. Wiley, New York (1986)

[14] Kifer, Y.: The exit problem for small random perturbations of dynamical systems with a hyperbolic fixed point. Israel J. Math. **40**(1), 74–96 (1981)

[15] Liggett, T.M.: Interacting particle systems. In: Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 276. Springer, New York (1985)

[16] Presutti, E.: Scaling limits in statistical mechanics and microstructures in continuum mechanics. In: Theoretical and Mathematical Physics. Springer, Berlin (2009)

[17] Rabinovich, M.I., Huerta, R., Varona, P., Afraimovich, V.S.: Transient cognitive dynamics, metastability, and decision making. PLoS Comput. Biol. **4**(5), e1000072, 9 (2008)

[18] Ratcliff, R., McKoon, G.: The diffusion decision model: theory and data for two-choice decision tasks (PMID: 18085991). Neural Comput. **20**(4), 873–922 (2008)

Yuri Bakhtin
School of Mathematics
Georgia Tech
Atlanta, GA 30332-0160, USA
e-mail: bakhtin@math.gatech.edu

Communicated by Jean Bellissard.
Received: September 1, 2011.
Accepted: October 5, 2011.