J. Nonlinear Sci. Vol. 15: pp. 223–253 (2005)
DOI: 10.1007/s00332-005-0671-z

Journal of
Nonlinear
Science

Magnetic Elements at Finite Temperature and Large Deviation Theory

R. V. Kohn,[1] M. G. Reznikoff,[2] and E. Vanden-Eijnden[1]

[1] Courant Institute of Mathematical Sciences, New York University, USA
E-mail: {kohn,eve2}@cims.nyu.edu

[2] Institut für Angewandte Mathematik, Universität Bonn, Germany
E-mail: reznikoff@iam.uni-bonn.de

Received October 8, 2004; accepted March 3, 2005
Online publication June 29, 2005
Communicated by C. R. Doering

Summary. We investigate thermally activated phenomena in micromagnetics using large deviation theory and concepts from stochastic resonance. We give a natural mathematical definition of finite-temperature astroids, finite-temperature hysteresis loops, etc. Generically, these objects emerge when the (generalized) Arrhenius timescale governing the thermally activated barrier crossing event of magnetic switching matches the timescale at which the magnetic element is pulsed or ramped by an external field; in the special and physically relevant case of multiple-pulse experiments, on the other hand, short-time switching can lead to non-Arrhenius behavior. We show how large deviation theory can be used to explain some properties of the astroids, like their shrinking and sharpening as the number of applied pulses is increased. We also investigate the influence of the dynamics, in particular the relative importance of the gyromagnetic and the damping terms. Finally, we discuss some issues and open questions regarding spatially nonuniform magnetization.

Key words. micromagnetics, Landau-Lifshitz-Gilbert equation, stochastic perturbation, rare events, large deviation theory, stochastic resonance, action minimization

Contents

1. Introduction

Submicron-sized ferromagnetic elements are the main building blocks in magnetoelectronics, where they are widely used as information storage devices [24], [25], [42], [52]. As these elements get smaller, the effects of thermal noise increase, particularly the ability of the noise to change the magnetization and thereby limit the data-retention time of the memory element [45], [49]. For this reason, noise-induced magnetization reversal has received a lot of attention in the magnetics community, from experimental, analytical, and numerical points of view [5], [6], [11], [22], [32], [43], [44].

A few standard experiments are used to study the stability of ferromagnetic elements against magnetization reversal. In pulsed experiments [27], the element is set to one of its preferred orientations. An (H_1, H_2) field is then applied for a certain time (a pulse), where H_1 is parallel to the easy axis and H_2 is perpendicular to it. After the pulse, the resistance is measured to determine whether the magnet has switched. Sometimes, multiple pulses are applied, in which case the magnet is allowed to settle down after each pulse, then pulsed again, and the magnetization is measured at the end of the series of pulses. The outcome is plotted as a black (switch) or white (no switch) square in the (H_1, H_2) plane (see Figure 1). This plot is called an astroid (or pulsed astroid). In a variant of this experiment, a magnetic field of increasing intensity is applied in one direction until the magnetization switches. Repeating the experiment along each direction and plotting the outcome produces a "ramped" or "swept" astroid [50].

Under either pulsing or ramping, one may also fix an applied field direction and consider one-dimensional hysteresis loops [4]. A hysteresis loop is a plot of magnetization versus field intensity. To generate a pulsed hysteresis loop, for instance, one measures whether a magnet set to one of the energy minimizers and pulsed at a given field strength switches to the opposing minimum. The loop consists of the magnetizations which do not switch, i.e., which are stable under the pulse (Figure 2).

At zero-temperature, such experiments amount to a bifurcation analysis of the stable states of the system. The boundary of the zero-temperature astroid (Figure 1) or the vertical jumps in the zero-temperature hysteresis loop (Figure 2) mark the "critical fields" that are just large enough to make unstable the state to which the system is originally set. The real experiments, however, are conducted at finite temperature, and they still produce fairly sharp astroids and hysteresis loops. At first sight, this may seem strange.

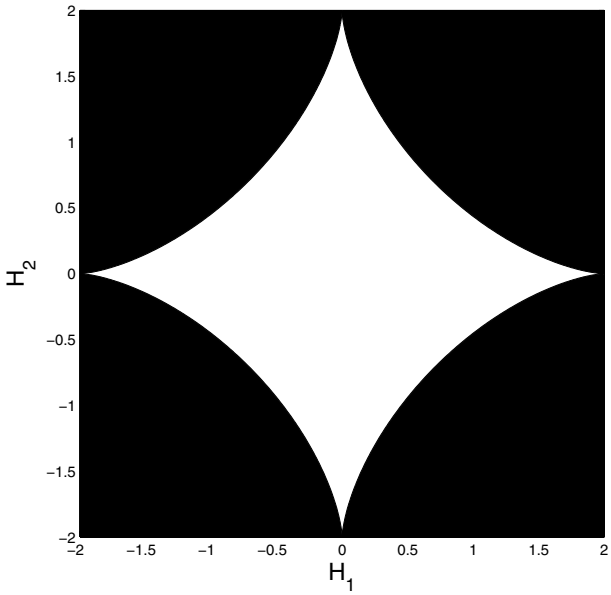

Fig. 1. The zero-temperature astroid associated with (3.2) for $\beta_2 = 1$. The boundary between the black (switch) and white (no-switch) regions is the parametric curve given in (3.3).

After all, the thermal noise eventually allows the magnetization to surmount any energy barrier, and thereby visit all possible configurations, no matter what the applied field is. As the outcome of a single trial of a thermally driven experiment, why should the astroids and hysteresis loops be so well-defined and sharp?

The key is the interplay between the timescale at which the experiments are conducted (for instance, the duration of the applied pulse) and the typical timescale on which the thermal fluctuations drive the system to overcome the energy barrier and change its magnetization. For a given pulse duration, the probability of switching may be either very small (if the applied field is too weak and the barrier high compared to the thermal energy available) or close to one (if the applied field is strong enough and the barrier small compared to the thermal energy). In a suitable limit, the boundary between these two regimes becomes sharp, thereby producing well-defined finite-temperature astroids and hysteresis loops. The main purpose of this paper is to quantify these statements using large deviation theory [15] and concepts similar to those used in stochastic resonance [14], [16]. As we will show, such an approach offers a comprehensive understanding of the experiments described above and sheds light on issues like the sharpening of the boundary of the astroid as the number of pulses is increased. The approach also reveals how short or ultrashort pulses can lead to non-Arrhenius switching behavior, a phenomenon which has been experimentally and numerically observed [22], [21], [25]. Finally, the approach by large deviation theory and stochastic resonance has the advantage that it can easily be generalized to more complicated situations, for instance when the magnetization in the sample is nonuniform and is described by a stochastic partial differential equation.

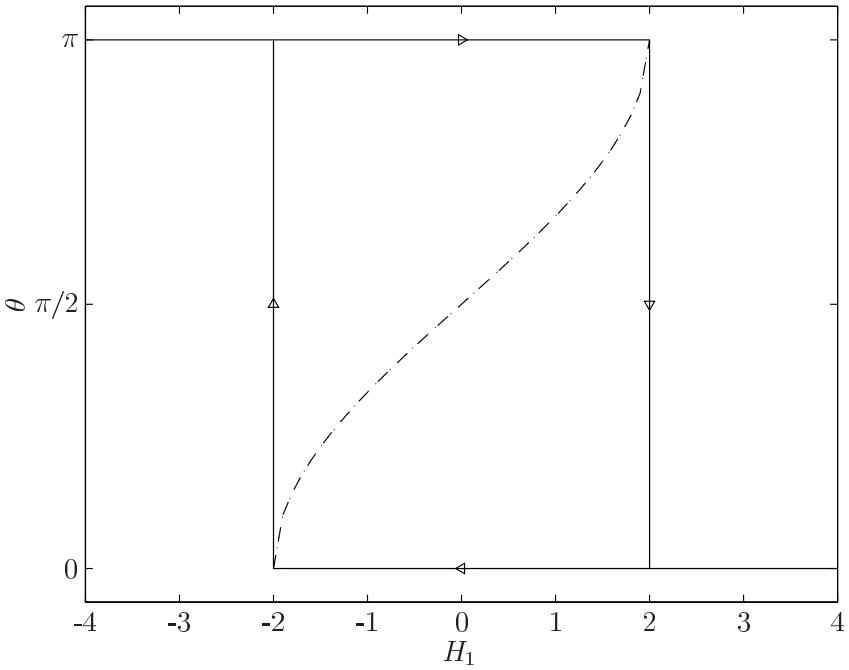

Fig. 2. The zero-temperature hysteresis loop for (3.2), with $H_2 = 0$ and $\beta_2 = 1$. The dotted line shows the location of the maximum, which collides with and annihilates the minimum at $\theta = \pi$ (or $\theta = 0$) when H_1 reaches 2 (resp. -2). The right-pointing arrows show the behavior of the magnetization as H_1 is increased (and the left-pointing arrows, as it is decreased), at zero temperature. The magnetization jumps only when the applied field actually annihilates the well in which it sits.

The remainder of this paper is organized as follows. In Section 2, we introduce the stochastic Landau-Lifshitz-Gilbert (LLG) equations for a uniform magnet. In Sections 3–7, we study the thin film limit of these equations. This is the simplest situation, but it retains the essence of the argument. In Section 3, we discuss single-pulse situations and determine the scaling relating the amplitude of the noise and the length of the applied pulse to obtain well-defined finite-temperature astroids. In Section 4, we generalize the argument to multiple-pulse astroids. We see that under the influence of a large number of pulses, exponentially unlikely events are important and the classical long-time switching path fails to capture the relevant physics. In Section 5, we discuss some properties of the finite-temperature astroids. In Section 6, we investigate finite-temperature ramped astroids, and in Section 7, finite-temperature hysteresis loops. In Section 8, we go back to the general situation and investigate thermal switching in the context of the full LLG equations. In this case, the switching path becomes nontrivial and depends on the duration of the pulse, the strength of the damping, and the anisotropy parameters. In Section 9, we look more closely at short-time switching and its connection to non-Arrhenius switching behavior. Deterministic effects arising for short pulses are discussed in Section 10. Section 11 lists some generalizations and open problems.

The focus is on nonuniform magnetization, especially the dependence of the optimal switching pathways on the applied field, and the identification of a low-dimensional reduced problem in a suitable asymptotic limit. For the reader's convenience, we also include two appendices. Appendix A discusses elementary properties of the stochastic LLG equations. Appendix B summarizes the necessary background on large deviation theory.

2. Modeling

We shall mostly focus on elements whose magnetization is uniform, which is relevant for a system of sufficiently small size [4], [52]. The zero-temperature dynamics of the system are described by the Landau-Lifshitz-Gilbert equation [1], [4], [7], [24], which after suitable nondimensionalization reads

$$\dot{m} = m \times h - \alpha m \times (m \times h). \tag{2.1}$$

In the uniform case, this is an *ordinary* differential equation, in which the unit vector $m = (m_1, m_2, m_3)$ describes the orientation of the magnetization in the ferromagnetic element. The effective field, h, is minus the gradient of the magnetic energy:

$$h = -\frac{\partial E}{\partial m}. \tag{2.2}$$

The uniform micromagnetic energy consists of the sum of anisotropy and applied field energies [4]. Both crystalline and shape anisotropy (the contribution of magnetostatic energy to the uniform model) play an important role. For simplicity, we consider an element that is a thin film in the m_1-m_2 plane, with uniaxial crystalline anisotropy and easy axis in the direction of m_1. Under a planar applied field (H_1, H_2), the energy is

$$E(m) = \beta_2 m_2^2 + \beta_3 m_3^2 - H_1 m_1 - H_2 m_2, \tag{2.3}$$

where $\beta_2, \beta_3 > 0$ are the anisotropy parameters. The parameter β_3 includes a shape anisotropy contribution proportional to the inverse aspect ratio, so for thin films, the relation $\beta_3 > \beta_2$ (or even $\beta_3 \gg \beta_2$) reflects the preference of the thin film magnetization to remain in the $m_3 = 0$ plane [51], [52]. The first term on the right-hand side of (2.1) is called the gyromagnetic term, and it is conservative. The second term on the right-hand side of (2.1), whose relative strength is controlled by the parameter $\alpha > 0$, can be written as $\alpha h^{\perp} = \alpha(h - (h \cdot m)m)$ and accounts for damping. Due to this term, the dynamics in (2.1) drive the system towards the closest local minimum of the energy (2.3).

Thermal effects in micromagnetics have been studied since the work of Brown in the 1960s [8]. They can be incorporated into the model by modifying the effective field in (2.1) to include a random term:

$$\dot{m} = m \times \left(h + \sqrt{\frac{2\alpha\varepsilon}{1+\alpha^2}}\dot{W} \right) - \alpha m \times \left(m \times \left(h + \sqrt{\frac{2\alpha\varepsilon}{1+\alpha^2}}\dot{W} \right) \right), \tag{2.4}$$

where $\dot{W}$ is a three-dimensional standard white-noise, and ε is the (dimensionless) temperature. Equation (2.4) is to be interpreted in the Stratonovich sense (which guarantees

that this equation preserves the norm constraint, $|m| = 1$), and it can be shown (see Appendix A) that the equilibrium distribution associated with (2.4) is the Gibbs distribution,

$$\rho(m) = Z^{-1} e^{-E(m)/\varepsilon}, \tag{2.5}$$

where $Z = \int_{|m|=1} e^{-E(m)/\varepsilon} d\sigma(m)$ is a normalization factor. We note that although there are differences of opinion as to whether to modify h and include the random term in both the gyromagnetic and the damping terms, as in (2.4), or in the gyromagnetic term alone, it does not matter as far as one-point statistics are concerned, because the two equations lead to magnetization fields which are identical in distribution (see Appendix A and [18], [44]). Furthermore, (2.4) can be written as the Itô stochastic differential equation,

$$\begin{cases} \dot\theta = -\dfrac{\partial E}{\partial z} - \dfrac{\alpha}{1-z^2}\dfrac{\partial E}{\partial \theta} + \dfrac{\sqrt{2\varepsilon\alpha}}{\sqrt{1-z^2}}\dot W_1, \\ \dot z = \dfrac{\partial E}{\partial \theta} - \alpha(1-z^2)\dfrac{\partial E}{\partial z} - 2\varepsilon\alpha z + \sqrt{2\varepsilon\alpha}\sqrt{1-z^2}\,\dot W_2, \end{cases} \tag{2.6}$$

using the representation

$$m = (\cos\theta\sqrt{1-z^2}, \sin\theta\sqrt{1-z^2}, z), \tag{2.7}$$

which automatically accounts for the constraint $|m| = 1$ (see Appendix A). The energy in (2.6) is (2.3) expressed in the (θ, z) variables:

$$E(\theta, z) = \beta_2(1-z^2)\sin^2\theta + \beta_3 z^2 - \sqrt{1-z^2}(H_1\cos\theta + H_2\sin\theta). \tag{2.8}$$

3. Single-Pulse Astroid in the Thin Film Limit

In the thin film limit of (2.3) or (2.8), the vanishing aspect ratio forces the shape anisotropy to infinity, i.e., $\beta_3 \to \infty$. Physically, this corresponds to the fact that an out-of-plane magnetization in a thin film element is strongly energetically penalized, and the magnetization tends to remain in-plane. In the limit, a finite energy requirement implies that $z = m_3 = 0$. The only remaining degree of freedom is the in-plane angle, θ, of the planar magnetization, $m = (m_1, m_2, 0)$. The equations (2.6) reduce to a single equation for θ,

$$\dot\theta = -\alpha E_R'(\theta) + \sqrt{2\alpha\varepsilon}\,\dot W_1, \tag{3.1}$$

and the "reduced energy" (i.e., the energy of a planar magnetization with angle θ) is

$$E_R(\theta) := \beta_2\sin^2(\theta) - H_1\cos\theta - H_2\sin\theta. \tag{3.2}$$

Note that (3.1) represents pure steepest descent dynamics, with no conservative term. The deterministic evolution is driven entirely by a gradient term that acts to reduce the energy, in contrast to equations (2.1), in which the evolution is driven by a conservative as well as a damping term. In this section and the sections immediately following (Sections 4–7), we focus on (3.1) because it simplifies the discussion while retaining the essence of the argument. We return to the general case in Section 8.

Consider the reduced energy, (3.2). At zero applied field, E_R is a standard double-well potential with two local minima, $\theta_A = 0$ and $\theta_B = \pi$. Both have zero energy. For a small but nonzero applied field, there are still two local minima, θ_A and θ_B, but the location and energy of the minima are shifted. As the intensity of the applied field in any fixed direction $(\cos\theta', \sin\theta')$ is increased, a "critical field" is encountered at which a bifurcation occurs and one minimum disappears. The field along any ray is called subcritical (resp. supercritical) if its magnitude is smaller (resp. larger) than the critical field in that direction. The family of critical fields can be parameterized as

$$(H_1^c, H_2^c) = (-2\beta_2 \cos^3(\theta'), 2\beta_2 \sin^3(\theta')), \quad \theta' \in [-\pi, \pi). \tag{3.3}$$

Recall that a field belongs to the interior of a pulsed astroid if applying the field (for a pulse of length T) causes one of the preferred magnetizations to switch to the opposing minimum after the pulse. At zero temperature, a subcritical field will never cause such a switching event, while a supercritical field always will. Therefore, the boundary of the zero-temperature astroid (Figure 1) is precisely (3.3). Next, we consider the situation at finite temperature.

Suppose without loss that the element is initially set to state $\theta = \pi$, and a subcritical field $H = (H_1, H_2)$ is applied for a pulse of length T. If the temperature ε is finite but small, with high probability the magnetization will first quickly relax to the vicinity of θ_B. Subsequently switching to θ_A has a *nonzero* probability, whose value depends on the pulse length T, the value of H, and the temperature ε. The large deviation estimate for this probability is (see Appendix B)

$$\text{Prob (no switch from } \theta_B \text{ to } \theta_A \text{ in time } T) \sim e^{-T/\tau(H)}. \tag{3.4}$$

Here and below $\sim$ denotes asymptotic equivalence ($f_\varepsilon \sim g_\varepsilon$ if $f_\varepsilon/g_\varepsilon \to 1$ as $\varepsilon \to 0$). $\tau(H)$ is the mean time to switch from θ_B to θ_A under the applied field H, which is related to the energy barrier $\Delta E_R(H)$ to escape θ_B as

$$\tau(H) \sim \nu^{-1} e^{\Delta E_R(H)/\varepsilon}. \tag{3.5}$$

The prefactor ν is given by Kramers' estimate (see Appendix B):

$$\nu = \frac{\alpha\sqrt{E_R''(m)E_R''(s)}}{2\pi}. \tag{3.6}$$

(We are assuming the generic condition that the minimal energy saddle, s, is unique. See Appendix B for more detail.) Viewed as a function of H at fixed T and $\varepsilon \ll 1$, (3.4) is a function that is mostly near 0 or 1, with a sharp transition between these two states (see Figure 3). Plotting (3.4) in the H-plane defines a "probabilistic astroid" that is consistent with the well-defined finite-temperature astroids observed in experiments. The transition region in the H-plane between extremely unlikely switching and almost certain switching is the region where the mean switching time $\tau(H)$ is logarithmically equivalent to the pulse length T, which we denote

$$\tau(H) \asymp T. \tag{3.7}$$

(Two functions f_ε and g_ε are called logarithmically equivalent, $f_\varepsilon \asymp g_\varepsilon$, if $\log f_\varepsilon / \log g_\varepsilon \to 1$ as $\varepsilon \to 0$.) See Figure 4 for an illustration. This selection by matching of timescales

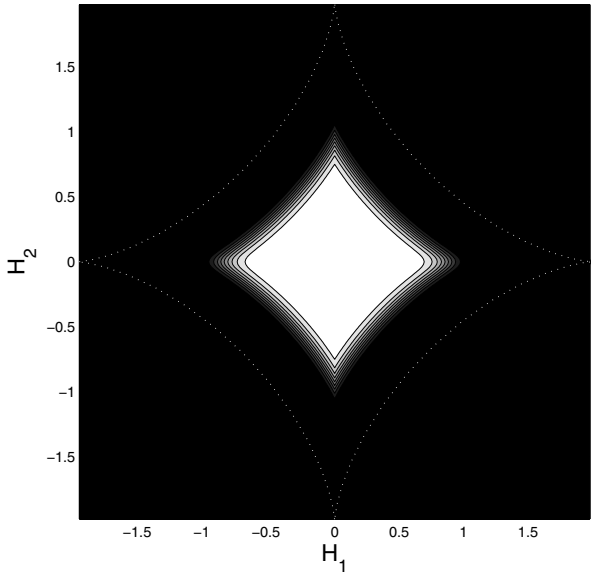

Fig. 3. We plot the switching probability as a function of H, using the approximation in (3.4). The grey scaling ranges from white for zero to black for one. The parameters are $T = 1000$, $\varepsilon = 0.05$, $\beta_2 = 1$. At this temperature, (3.4) is a function varying rapidly from zero to one, and the grey region is of small extent. This is consistent with the emergence of a reduced but sharp finite-temperature astroid, as observed in the experiments. The finite-temperature astroid is defined precisely in Proposition 1. See also Figure 5. The boundary of the zero-temperature astroid is also shown (white dotted line).

is similar to the one observed in stochastic resonance. (See [14], [16], and also Section 6.) Notice that (3.7) does not involve the prefactor ν, but it implies that T must be exponentially large in ε^{-1} in order that the finite-temperature astroid be different from its zero-temperature counterpart—in which case it is also necessarily smaller. In fact, the precise statement is:

Proposition 1 (Finite-Temperature, Single-Pulse Astroids). *Consider the sequence of pulses $T_\varepsilon := e^{A/\varepsilon}(B + O(\varepsilon))$, where $A, B > 0$. Let*

$$\Omega(A) := \{H : \Delta E_R(H) > A\}$$

be the region in the H-plane where $\Delta E_R(H)$, the energy barrier out of θ_A, is larger than A. Then

$$\lim_{\varepsilon\to 0} \textit{Prob}(\textit{no switch for } t \le T_\varepsilon) = \begin{cases} 1, & H \in \Omega, \\ 0, & H \in \mathbb{R}^2/\overline{\Omega}. \end{cases}$$

Proof. For any H, the mean switching time satisfies the logarithmic equivalence

$$\tau(H) \asymp e^{\Delta E_R(H)/\varepsilon}.$$

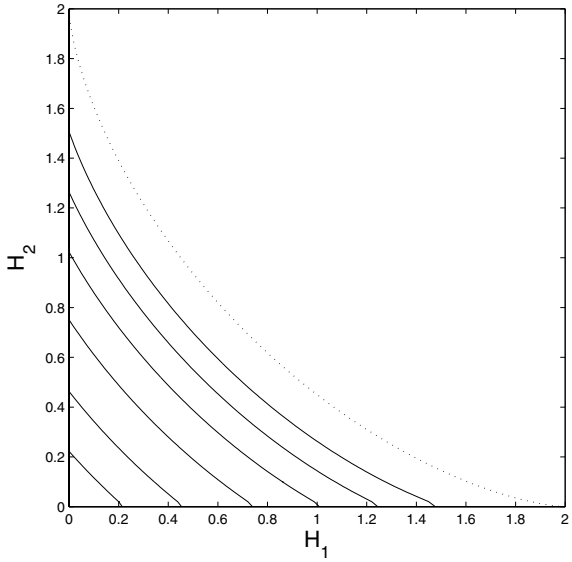

Fig. 4. Level sets (isolines) of $e^{-\Delta E_R(H)/\varepsilon} \asymp \tau(H)$. For a given T, the matching condition (3.5) is satisfied on one of these level sets, which therefore approximate the boundary of the finite-temperature astroid. The dotted line is the boundary of the zero-temperature astroid given by (3.3).

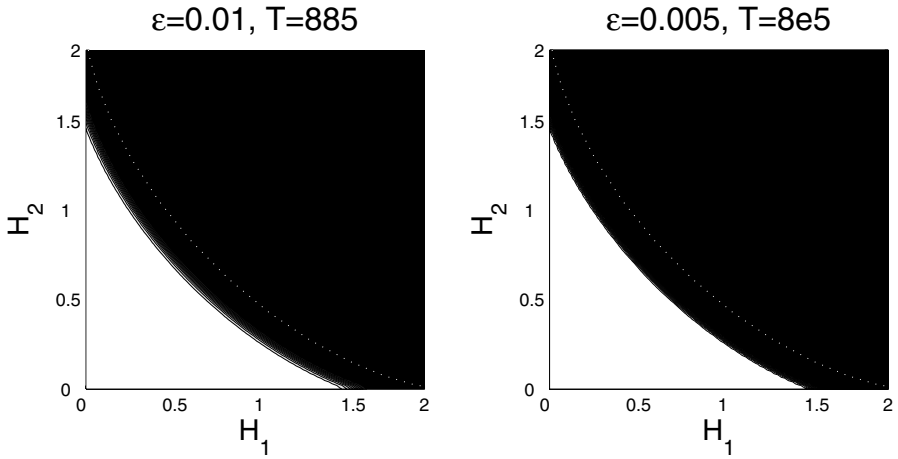

Fig. 5. When the pulse length increases exponentially in ε^{-1} as the noise decreases, the probability of not switching, (3.4), converges to a piecewise constant function taking the value 0 or 1. The limiting region constitutes the well-defined, finite-temperature astroid (which is different from the zero-temperature astroid, whose boundary is shown as the white dotted-line). This figure is an illustration of Proposition 1. (Here, $\beta_2 = 1$.)

We now use the exponential distribution (3.4) (cf. Appendix B) in three cases.

Case (1)*,* $\Delta E_R(H) = A$*, so that* $T_\varepsilon \asymp \tau(H)$. The exponential distribution applies, i.e.,

$$\lim_{\varepsilon\to 0} \text{Prob (no switch for } t \le T_\varepsilon) \in (0, 1).$$

Case (2)*,* $\Delta E_R(H) < A$. Here, the probability of not switching goes to zero. To see this, choose any $M > 0$.

$$\begin{aligned}\lim_{\varepsilon\to 0} \text{Prob (no switch for } t \le T_\varepsilon) &\le \lim_{\varepsilon\to 0} \text{Prob (no switch for } t \le M\tau(H)) \\ &\le e^{-M}.\end{aligned}$$

Let $M \to \infty$.

Case (3)*,* $\Delta E_R(H) > A$. Here, the probability of not switching goes to one. This time, choose any $m > 0$,

$$\begin{aligned}\lim_{\varepsilon\to 0} \text{Prob (no switch for } t \le T_\varepsilon) &\ge \lim_{\varepsilon\to 0} \text{Prob (no switch for } t \le m\tau(H)) \\ &\ge e^{-m}.\end{aligned}$$

Let $m \to 0$. □

Proposition 1 leads to the following natural definition:

Definition 1 (Finite-temperature, single-pulse astroid). *For given temperature and pulse time, the finite-temperature, single-pulse astroid is defined to be the following region in the* H*-plane:*

$$\Omega(\varepsilon, T) := \{H : \Delta E_R(H) > \varepsilon \ln(\nu T)\}.$$

4. Multiple-Pulse Astroids and Action Minimization

As shown in Section 3, when a single pulse is applied, the length T must be exponentially large in the inverse temperature, ε^{-1}, in order to observe a finite-temperature astroid that is both well-defined and nontrivial (i.e., smaller than its zero-temperature counterpart). The situation changes if one considers multiple-pulse experiments and allows the number of pulses, N, to be very large. This is physically relevant since magnetic memory devices need to be able to withstand many (on the order of 10^{17}) subcritical pulses without accidentally switching [26]. In this case, T can be $O(1)$, provided that N is exponentially large in ε^{-1} (and, of course, intermediate situations are possible as well). This is interesting because the switching event occurs within a time that is *very short* compared to the mean switching time! The exponential distribution breaks down, and action minimization enters the picture, as explained next in Sections 4.1 and 4.2.

We emphasize that multiple-pulse astroids are a real-world example in which we can observe exponentially unlikely events. In the long-time limit, the overwhelming

probability is that switching is achieved by flowing uphill to the minimal saddle. In multiple-pulse astroids, however, the extremely unlikely event of short-time switching dominates.

4.1. Action Minimization

Large deviation theory asserts that the probability of switching within a fixed time, T, is estimated via an action minimization problem. Consider an initial magnetization $\theta = \pi$, and let C_A denote the basin of attraction of the energy minimizer, θ_A, under the deterministic dynamics,

$$\dot{\theta} = -E_R'(\theta).$$

(Generically, $\theta_A \neq \pi$ for $|H| \neq 0$.) The probability of switching from π to θ_A in time T is estimated:

$$\text{Prob}(\text{switch for } t \leq T) \asymp e^{-S_T/\varepsilon}, \tag{4.1}$$

with the action, S_T, defined as

$$S_T := \inf_{\substack{\phi(0)=\pi \\ \phi(T)\in C_A}} \frac{1}{4}\int_0^T |\dot{\phi}(s) + E_R'(\phi(s))|^2\, ds. \tag{4.2}$$

The pathway $\phi_\star : [0, T] \to \mathbb{R}$ that minimizes the action functional is called the optimal switching path, and it is the most likely switching path: With probability 1 as $\varepsilon \to 0$, the actual switching paths remain within an arbitrarily small neighborhood of $\phi_\star$. The graph of $\phi_\star$ is trivial in one dimension (it connects π to the optimal saddle point, θ_s), but not in two dimensions; see Section 8.

For finite $T \in (0, \infty)$, the action has to be found via numerical minimization, and an accurate estimate is important if we are interested in taking exponentially many pulses. In the long-time limit ($T \to \infty$), the action converges to the energy barrier,

$$S_T \to \Delta E_R(H), \tag{4.3}$$

and the optimal path converges to the heteroclinic orbit connecting θ_B and θ_s (see Appendix B). The rate of convergence in (4.3) depends only on the characteristic timescale of the deterministic dynamics, $\dot{\theta} = -E_R(\theta)$, which we will denote by t_d. (In the present context working with (3.2), we can for instance take $t_d = \nu^{-1}$, the inverse of the prefactor (3.6).) The role of the deterministic timescale, in combination with the role of the mean switching time in Proposition 1, suggests the classification:

Definition 2 (Classification of Pulse Length).

1. *Normal pulses last for a time T that is logarithmically equivalent to the mean switching time (cf. Proposition 1).*
2. *Short pulses last for a time T that is short compared to the mean switching time, but long compared to the deterministic timescale:*

$$t_d \ll T \ll \tau(H).$$

3. *Ultrashort pulses last for a time that is comparable to, or even shorter than, the deterministic timescale t_d.*

Single-pulse experiments lead to finite-temperature astroids that are both well-defined and nontrivial (i.e., different from their zero-temperature counterparts), and that are controlled by switching under normal pulses (cf. Definition 1). The generalization to ultrashort and short pulses is given next.

4.2. Multiple-Pulse Astroids

From (4.1), we see that for $N := e^{A/\varepsilon}$ and fixed T, the switching probability converges to zero or one, according to whether the action is greater or less than A, which we now state as a proposition.

Proposition 2 (Finite-Temperature, Ultrashort, Multiple-Pulse Astroid). *Fix T and define $N_\varepsilon := e^{A/\varepsilon}(B + O(\varepsilon))$, where $A, B > 0$. Then for*

$$\Omega(A) := \{H;\, S_T(H) > A\},$$

we have

$$\lim_{\varepsilon\to 0} \mathit{Prob}(\textit{no switch for } N_\varepsilon \textit{ pulses of length } T) = \begin{cases} 1, & H \in \Omega, \\ 0, & H \in \mathbb{R}^2/\overline{\Omega}. \end{cases}$$

Proof. The result follows immediately from the fact that

$$\text{Prob(no switch after } N \text{ pulses)} = (1 - C_\varepsilon \exp(-S_T(H)/\varepsilon))^{N_\varepsilon}, \tag{4.4}$$

where C_ε is the subexponential prefactor. □

A proposition similar to Proposition 2 applies for short pulses if one uses $\Omega(A) = \{H : \Delta E_R(H) > A\}$. It suggests the generalization of Definition 2 for both single and multiple pulses of arbitrary length as follows:

Definition 3 (Single- and Multiple-Pulse Finite-Temperature Astroid). *The finite-temperature astroid is defined as the following region in the H-plane:*

$$\Omega(\varepsilon, T, N) := \{H : S_T(H) > \varepsilon \ln(\nu T N)\}.$$

5. Astroid Properties

Experiments observe *shrinking and sharpening* of the astroid under multiple pulses. We look at necessary and sufficient conditions for these properties.

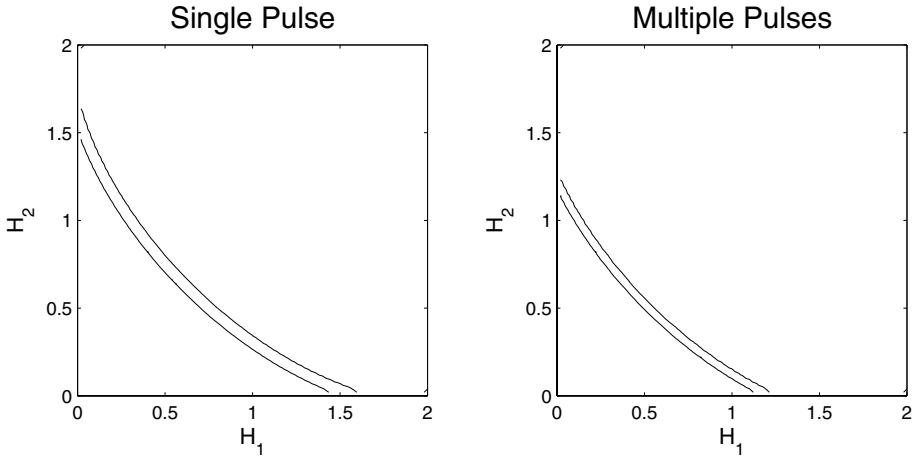

Fig. 6. The δ-transition region of (Prob(no switch for $t \leq T$))^N, for $\varepsilon = 0.01$, $T = 1000$, $\delta = 0.07$. In the first plot, $N = 1$, while in the second, $N = 65,000$. The thinning of the transition region as the number of pulses increases is consistent with the experimental discovery of sharper multiple-pulse astroids. Also consistent is the fact that the multiple-pulse astroid is smaller than the single-pulse astroid.

Shrinking. First, we observe that the astroid shrinks under multiple pulses if the switching probability is monotonic in $|H|$, which is physically reasonable. A necessary and sufficient condition for this monotonicity is for the action to be monotonic in $|H|$. The monotonicity certainly holds for fields within the exponential regime (by the monotonicity of τ).

Sharpening and Log-Convexity. It is observed that astroids become sharper under multiple pulses, in the sense that there are fewer "freckles," white spots in a field of grey or grey spots in a field of white [26]. We find an explanation for this sharpening by looking at the region of probability transition. The numerics for our simple model demonstrate the thinning of the transition region under multiple pulses, as shown in Figure 6.

Analytically, we show that the transition region shrinks under multiple pulses if and only if the mean switching time for applied field $H = h(\cos\phi, \sin\phi)$ is strictly log-convex as a function of h. (Strict log-convexity means that the log of the function is strictly convex.) First, we define the notions of the δ-transition region and the band-narrowing property.

Definition 4 (δ**-Transition Region).** *For $f : \mathbb{R}^2 \to [0, 1]$ and any $\delta > 0$, we call the region of $\mathbb{R}^2$ for which*

$$f(x) \in [\delta, 1 - \delta]$$

the δ-transition region of f.

Definition 5 (Band-Narrowing Property). *The positive, decreasing function $f(x)$ is band-narrowing if whenever $0 < A < B$ and $T_1 < T_2$, then*

$$f^{-1}(T_2 A) - f^{-1}(T_2 B) < f^{-1}(T_1 A) - f^{-1}(T_1 B).$$

A simple example of a band-narrowing function is $1/x$.

Propositions 3 and 4 below show that within the exponential regime, a necessary and sufficient condition for the narrowing of the δ-transition region of the probability under multiple pulses is that the mean switching time be strictly log-convex. First of all, the δ-transition region of the probability shrinks if and only if the mean switching time is band-narrowing.

Proposition 3. *Consider any ray $H = h(\cos\phi, \sin\phi)$ and suppose*

$$\textit{Prob}\,(\textit{no switch for } t \le T) = \exp(-T/\tau^{\phi}(h)),$$

for some positive, decreasing function τ^{ϕ}. Then the δ-transition region of the probability is decreasing in T for every $\delta \in (0, 0.5)$ if and only if τ is band-narrowing on that ray.

Proof. This is a direct calculation. □

Next, the mean switching time is band-narrowing if and only if it is strictly log-convex.

Proposition 4. *The positive, decreasing, continuously differentiable function f is band-narrowing if and only if $f'(x)/f(x)$ is increasing or, equivalently, if and only if f is strictly log-convex.*

Proof. Let $g(t) := f^{-1}(At) - f^{-1}(Bt)$. f is band-narrowing if and only if g is strictly decreasing. Let $f^{-1}(tA) =: X_2$ and $f^{-1}(tB) =: X_1$, and consider the derivative.

$$\begin{aligned}
g'(t) < 0 &\Leftrightarrow A(f^{-1})'(tA) < B(f^{-1})'(tB) \\
&\Leftrightarrow \frac{A}{f'(f^{-1}(tA))} < \frac{B}{f'(f^{-1}(tB))} \\
&\Leftrightarrow \frac{A}{f'(X_2)} < \frac{B}{f'(X_1)} \\
&\Leftrightarrow \frac{f(X_2)}{tf'(X_2)} < \frac{f(X_1)}{tf'(X_1)} \\
&\Leftrightarrow \frac{f(X_2)}{f'(X_2)} < \frac{f(X_1)}{f'(X_1)}.
\end{aligned}$$

□

6. Ramped Astroids

A different but related physical experiment consists of applying an external field by "ramping," which refers to steadily increasing the applied field until switching is

achieved. Ramped astroids are also analyzed (with a different emphasis) in [50]. As in the pulsed case, we find a 0–1 distribution in the limit of vanishing noise strength, where now the phenomenon is stochastic resonance [14], [16] in the traditional sense: The increasing applied field corresponds to a tilting potential, the noise sets the timescale for switching (with nearly deterministic precision), and as soon as surmounting the barrier is *possible*, it is also *probable*, with probability converging to one in the limit. For ramped astroids, the *ramping rate is critical* in defining the time and field at which switching occurs.

Consider the question of when switching will occur. Let the applied field be of the form $H(t) = \gamma t(\cos\phi, \sin\phi)$ where ϕ is a fixed angle, t is time, and $\gamma \ll 1$ is a parameter that sets the timescale for ramping. At zero temperature, switching occurs when γt reaches the boundary of the zero-temperature astroid. At finite-temperature, switching occurs when

$$e^{\Delta E_R(H(t))/\varepsilon} \asymp \gamma^{-1}.$$

We make this precise in the following proposition.

Proposition 5 (Ramped Astroid). *Let the ramped field be $H(t) = \gamma t(\cos\phi, \sin\phi)$ with rate $\gamma := Be^{-A/\varepsilon}$, $A > 0$. Then $\Omega := \{H;\, \Delta E_R(H) > A\}$ is such that*

$$\lim_{\varepsilon\to 0} \textit{Prob}(\textit{not to have switched when field is } H) = \begin{cases} 1, & H \in \Omega, \\ 0, & H \in \bar{\Omega}^c. \end{cases}$$

Proof. Fix any ray through the origin and consider ramping along the ray. Find the value $\hat{H}$ along the ray such that $\Delta E_R(\hat{H}) = A$. Consider the probability of switching while the ramping field is within a tolerance $0 < \delta \ll 1$ of $\hat{H}$. Without loss of generality, rotate the frame of reference and consider $\hat{H} \in \mathbb{R}$.

The time interval such that $H(t) \in [\hat{H}, \hat{H} + \delta]$ is $I := [\hat{H}/\gamma, (\hat{H} + \delta)/\gamma]$, of length δ/γ. Furthermore, for this time range, $\Delta E_R(H(t)) \le \Delta E_R(\hat{H}) = A$. For simplicity, assume that the prefactor in the mean switching time is unity, since it is subexponential and does not affect the result. Then, using the exponential distribution,

$$\begin{aligned} \text{Prob}\left(\text{no switch while } H \in [\hat{H}, \hat{H} + \delta]\right) &= \text{Prob}(\text{no switch for } t \in I) \\ &\le \exp\left(-\frac{\delta B^{-1} e^{A/\varepsilon}}{e^{\Delta E_R(\hat{H})/\varepsilon}}\right) \\ &= \exp\left(-\delta B^{-1}\right) \\ &\in (0, 1). \end{aligned}$$

Next, by the method of Proposition 1 in Section 3, one shows that the probability of switching for $H(t) \in [0, \hat{H} - \delta]$ goes to zero, and the probability of having switched by the time $H(t) = \hat{H} + \delta$ goes to one, for any $\delta > 0$. □

7. Hysteresis Loops

As discussed in the introduction, thermally reduced hysteresis loops are another example in which physical observations can be understood via large deviation theory and a concept similar to stochastic resonance. The hysteresis loop tracks the location of the critical points of the energy as a function of the applied field magnitude, h, in a fixed direction, i.e., $H = h(\cos\phi, \sin\phi)$. (Recall, for example, the zero-temperature hysteresis loop from Figure 2.) Its most important feature is that it identifies the applied field (positive and negative) at which a metastable state loses stability and disappears. The obvious extension of the ideas of the preceding sections is that under the influence of noise, the zero-temperature hysteresis loop should be replaced with a plot of the *probability of switching out of the metastable basin* under an applied pulse. Plotting this probability versus applied field magnitude, there is a sharp transition from probability nearly zero to probability nearly one. The transition occurs in a small neighborhood of the applied field whose mean switching time is equal to the given pulse time.

To illustrate the idea, let us consider a specific example. We choose to look at $H_2 = 0$, since in that case we can solve explicitly for τ as a function of H_1. (The energy barrier in this case is $\Delta E_R = (H_1/2 - 1)^\beta$, with $\beta = 2$. Wernsdorfer points out in his review article [52] that $\beta = 2$ is nongeneric, and specific to the case when the applied field and the uniaxial anisotropy are aligned.) Letting $\beta_2 = 1$ and solving for critical points of the energy for $H_1 \in (0, 2)$ yields

$$(m_1, s_1, m_2, s_2) = \left(0, \cos^{-1}\left(\frac{-H_1}{2}\right), \pi, 2\pi - \cos^{-1}\left(\frac{-H_1}{2}\right)\right).$$

Using the Kramers' approximation for the switching time (cf. Appendix B), the equation $\tau = T$ becomes

$$\frac{\pi\sqrt{2}}{\alpha\sqrt{(2 - H_1)(4 - H_1^2)}} \exp\left(\frac{(H_1/2 - 1)^2}{\varepsilon}\right) = T.$$

We calculate H_1 iteratively. The reduced loop is shown in Figure 7 for $T = 1000$, $\alpha = 1$, and $\varepsilon = .01$.

8. The General Case

We have seen how to extract an explanation of finite-temperature effects for the reduced model, (3.2). Of course, this simple model can be analyzed by other means as well, for instance by solving the Fokker-Planck equation for the mean first passage time [10], [19], [41]. The power of the large deviation perspective is that it can be generalized to higher dimensional situations when the Fokker-Planck approach is not practical anymore. In particular, large deviation theory and prefactor estimates offer a general framework by which to understand thermal effects for the full energy, (2.3), or even the case of spatially nonuniform magnetization, which we discuss in Section 11. Considering now (2.3), we look at the implications of large deviation theory for the question of thermally activated switching between the basins of attraction of the modified energy landscape, i.e. the

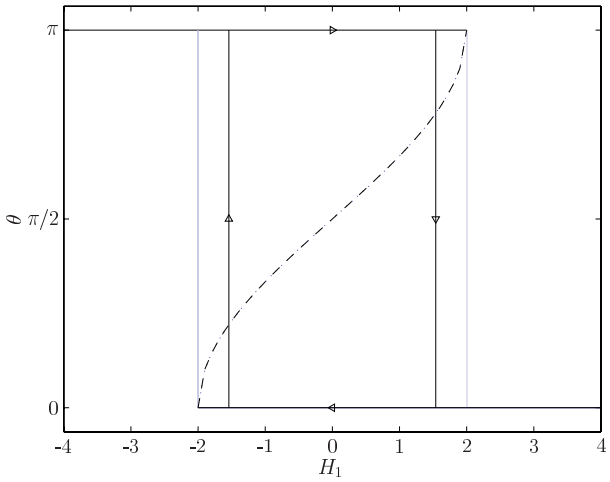

Fig. 7. Finite-temperature hysteresis loop for (2.3), with $H_2 = 0$, $\beta_2 = 1$, $\varepsilon = 0.01$, and $T = 1000$. Compare to the zero-temperature loop, Figure 2. The zero-temperature loop shrinks to the finite-temperature counterpart: The jump in magnetization occurs at the H_1 value for which $\tau_{LD} = T$, *before* annihilation of the energy barrier, and thermally activated switching is observed at subcritical fields.

basins of the energy landscape under the applied field. (This can be different from the question of *having switched between the basins of the zero-field system* after the pulse has terminated, e.g. for short pulses and low damping. We discuss that complication in Section 10.)

8.1. Normal Pulses

We saw that for normal pulses (or, equivalently, for single-pulse astroids) the switching probability for the reduced model was determined by the energy barrier from the minimum to the minimal-energy saddle point. We now show that the same is true for the full system, because of the property that the gyroscopic term in the LLG equations is perpendicular to the damping term. We will also see a new feature, namely that the optimal switching path, which is trivial in one dimension, can be an interesting object for the full model. Our system is of the form

$$\dot{X} = -(\gamma + r + \varepsilon c) + \sqrt{2\varepsilon}\sigma\dot{W}, \tag{8.1}$$

with $X = (\theta, z)$ and

$$\gamma := \begin{pmatrix} \dfrac{\alpha}{1-z^2}E_\theta \\ \alpha(1-z^2)E_z \end{pmatrix}, \qquad r := \begin{pmatrix} E_z \\ -E_\theta \end{pmatrix},$$

$$c = \begin{pmatrix} 0 \\ 2\alpha z \end{pmatrix}, \qquad \sigma = \begin{pmatrix} \dfrac{\sqrt{\alpha}}{\sqrt{1-z^2}} & 0 \\ 0 & \sqrt{\alpha}\sqrt{1-z^2} \end{pmatrix}.$$

The Wentzell-Freidlin action functional [15] is

$$S_{[0,T]}(\phi) = \frac{1}{4}\int_0^T |\sigma^{-1}(\phi_s)(\dot{\phi}_s + \gamma(\phi_s) + r(\phi_s))|^2\, ds, \tag{8.2}$$

for all ϕ such that the integral is defined, and ∞ otherwise. Because of the relation $(\sigma^{-2}\gamma, r) = (\nabla E, r) = 0$, we see that we can manipulate the action:

$$\begin{aligned} |\sigma^{-1}(\dot{\phi} + \gamma + r)|^2 &= |\sigma^{-1}(\dot{\phi} - \gamma + r + 2\gamma)|^2 \\ &= |\sigma^{-1}(\dot{\phi} - \gamma + r)|^2 + 4|\sigma^{-1}\gamma|^2 + 4(\sigma^{-2}\dot{\phi}, \gamma) - 4|\sigma^{-1}\gamma|^2 \\ &= |\sigma^{-1}(\dot{\phi} - \gamma + r)|^2 + 4\dot{E}. \end{aligned} \tag{8.3}$$

From (8.2) and (8.3), we recognize that, as before,

$$S_{[0,T]} \geq \Delta E,$$

i.e., the action is bounded below by the energy barrier. This leads to the conclusion (cf. [15]) that the energy barrier controls the mean switching time. Furthermore, Day shows that the exponentiality of the limiting distribution of $t_B/\bar{t}_B$ also holds for the full dynamics [9]. Thus, Definition 1 extends *precisely* to the general case. Furthermore, it is not hard to see that as long as β_3 is sufficiently large compared to β_2, the minimal-energy saddle point lies along $z = 0$, and the reduced model *completely characterizes* the single-pulse astroid for the full model.

8.2. Optimal Switching Path

To achieve the sharp bound in (8.3), we see that the optimal trajectory should solve

$$\dot{\phi} = \gamma - r$$

to travel from $(\theta_B, 0)$ to the minimal saddle, and solve

$$\dot{\phi} = -\gamma - r$$

to relax from the saddle to $(\theta_A, 0)$. This is the optimal path in the long-time limit, and is close to the optimal path as long as T is large compared to the deterministic timescale of the system. This long-time optimal path is a function of the applied field, the damping, and the anisotropy. For the sake of illustration and to emphasize the dependence on the system parameters, we consider two examples. The first, shown in Figure 8, corresponds to a situation where $\beta_2 = 1$, $\beta_3 = 5$, $\alpha = 01$, $H_1 = 0.5$, $H_2 = 0$. Observe that the "uphill path" is not simply time-reversed flow along the heteroclinic orbit, as in the reduced model; the sign is reversed on the damping term, but preserved for the gyroscopic term. Furthermore, even if the minimal saddle lies on $z = 0$, the path is two-dimensional.

The second example, shown in Figure 9, is for the case $\beta_2 = 1$, $\beta_3 = 0.5$, $\alpha = 0.1$, $H_1 = 0.5$, $H_2 = 0$. With the out-of-plane penalization relaxed, even the saddle points have $z \neq 0$.

The optimal path in the *spatially nonuniform* problem is yet more interesting and complicated. See Section 11 for a discussion.

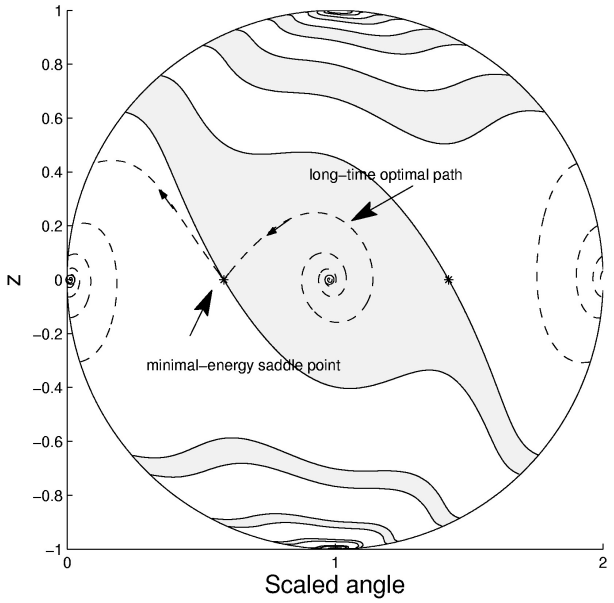

Fig. 8. In the general case, the optimal path is distinct from (but related to) the heteroclinic orbit. Here, $H_1 = 0.5$, $H_2 = 0.1$, $\alpha = 0.1$, $\beta_2 = 1$, and $\beta_3 = 5$. Here and in Section 10, we plot the two-dimensional projection in terms of the height z vs. the scaled angle $(\theta/\pi - 1)\sqrt{1 - z^2} + 1$ (preserving length along lines of constant latitude). The minima are at $(\theta, z) = (0.040, 0)$ and $(3.075, 0)$. The minimal-energy saddle point is at $(1.838, 0)$, and a secondary saddle is at $(4.472, 0)$. Although they lie on $z = 0$, the optimal path is **not** confined to the circle.

8.3. Short and Ultrashort Pulses

As always, short-time switching forces us to confront the full action minimization problem, which is given by (4.2) with the action functional defined by (8.2). As it did for the reduced model, the estimate (4.4) assures that the probability of multiple-pulse switching, as a function of H, will tend to either 0 or 1, depending on the action associated with the given applied field. Thus, Definition 3 also extends to the general case. (See also Section 9.) Again, these ideas generalize and become yet more interesting in the spatially nonuniform problem.

9. Non-Arrhenius Behavior

The so-called Arrhenius behavior refers to the probability distribution of switching time being exponential in T. If this is not the case, one says that the behavior is non-Arrhenius.

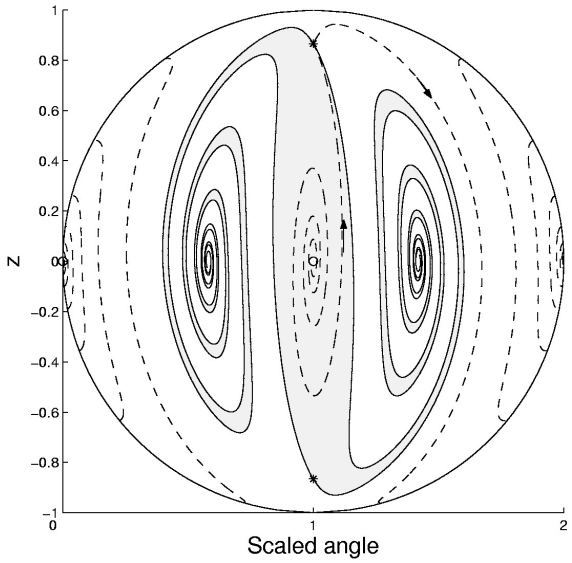

Fig. 9. Here, $H_1 = 0.5$, $H_2 = 0$, $\alpha = 0.1$, $\beta_2 = 1$, and $\beta_3 = 0.5$. The minima are at $(\theta, z) = (0, 0)$ and $(\pi, 0)$. The minimal-energy saddle points, marked with a $*$, are at $(\pi, \pm\sqrt{3}/2)$. They do **not** lie on $z = 0$.

Non-Arrhenius behavior has been observed in physical and numerical experiments [21], [22], [25].[3]

One source of non-Arrhenius behavior, which enters even for normal pulses, is the existence of intermediate states (which are not present in the simple, uniaxial model). When there are intermediate states in the switching process, the system is described by a multistate (rather than two-state) Markov chain, much in the spirit of the ladder model introduced in [22]. The reduction of magnetic switching to the dynamics of a multistate Markov chain is explored in [11], [22], [25]. Intermediate states are found in the coherent rotation model with cubic anisotropy, for example. Intermediate states are also a *signature feature of noncoherent rotation*. Intermediate states in the nonuniform micromagnetic energy with zero applied field have been investigated numerically in [11].

A second source of non-Arrhenius behavior is short-time switching. This is clear from the non-Arrhenius distribution in (4.4). When short or ultrashort switching is involved, the optimal switching path and corresponding action can be computed numerically by solving the action minimization problem (as functions of the pulse time and applied field). It is important to emphasize that a path which is optimal at one value of T (or H) may be *far from optimal* at another value of T (or H). Identifying the correct pathway is necessary in order to estimate accurately the switching probability.

[3] In other settings, it is natural to study the exponentiality of the mean switching time with respect to the temperature rather than the pulse length. Non-Arrhenius effects with respect to temperature are studied, for instance, in [35].

To quantify the effect of the nonexponentiality induced by short-time switching, consider the difference between the single-pulse astroid under pulse-length $N_\varepsilon T$ and the multiple-pulse astroid built by taking sequences in which N_ε pulses of length T are applied. In the former case, the finite-temperature astroid has a boundary of thermally critical fields for which $S_{N_\varepsilon T} = \varepsilon \ln(N_\varepsilon T)$. In the latter case, the thermally critical fields are those for which $S_T = \varepsilon \ln N_\varepsilon$. For T fixed and $\varepsilon \ll 1$, this basically amounts to comparing $S_T(H)$ with $\Delta E(H)$, and the difference can be large or small, depending on T.

10. Ultrashort Switching, Low Damping, and Deterministic Effects

We mentioned in Section 8 that we should consider whether there is a difference between having switched in the pulsed landscape and having switched in the zero-field landscape after termination of the pulse. As we discuss below, for high damping, the questions are roughly equivalent; for low damping and ultrashort pulses, however, the situation is different, and the effect of the deterministic, gyroscopic motion can dominate.

When the pulse is turned on, the magnetization finds itself perturbed away from the nearby minimum of the pulsed landscape (as long as $H_1 \neq 0$). Under the deterministic dynamics, it relaxes to the minimum. Furthermore, if the damping is large enough (Figure 11), or, regardless, if the pulse is long enough for relaxation, then the only way for

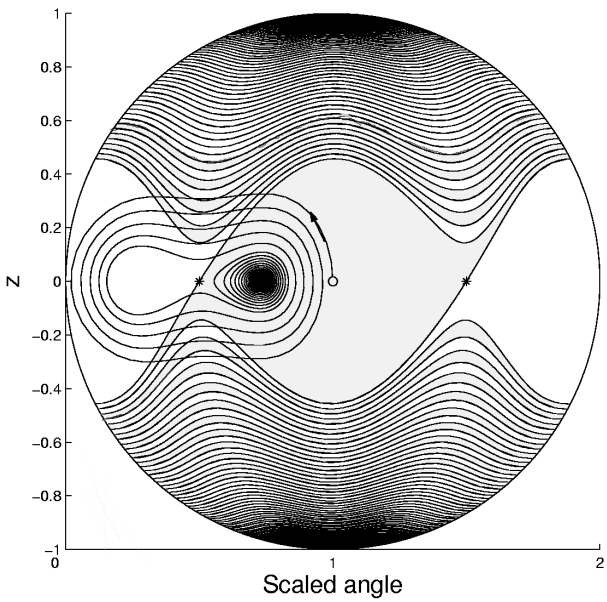

Fig. 10. For $\alpha = 0.01$, $H_1 = 0$, $H_2 = 1.5$, the magnetization is initially perturbed away from the nearby minimum of the pulsed landscape, which can send it into a complicated deterministic orbit. The trajectory flies around the sphere, and ultrashort pulses can "catch" the magnetization in either basin, overriding thermal effects.

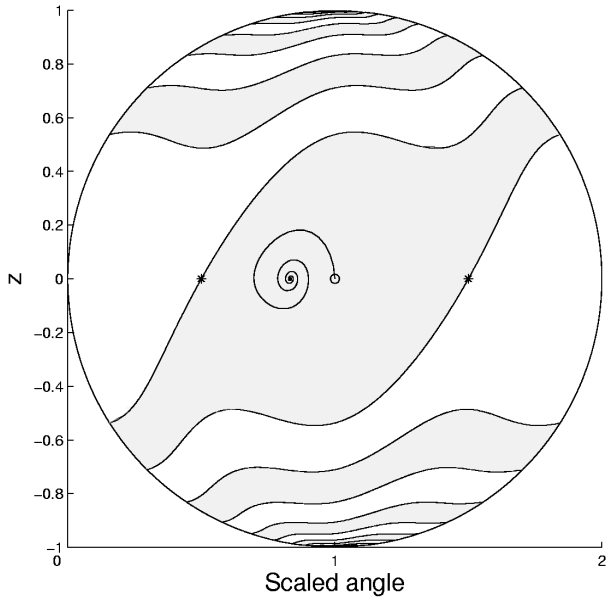

Fig. 11. For $\alpha = 0.1$, $H_1 = 0$, $H_2 = 1.5$, the damping is high enough that, even for ultrashort pulses, the deterministic orbit does not cause switching.

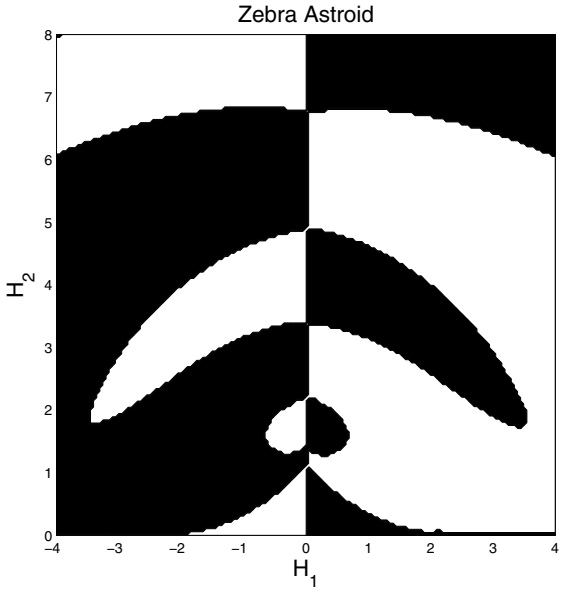

Fig. 12. Short-pulse, low-damping "zebra astroid." Here, $\alpha = 0.01$ and $T = 28$.

switching to happen is by thermal driving. When damping is low, however, the relaxing trajectory can wind around the sphere, as illustrated in Figure 10. If the pulse is short enough to "catch" the magnetization along the trajectory (i.e. if it is ultrashort), then switching can be achieved precisely by turning the pulse off while the magnetization is travelling through the basin of attraction of θ_A in the zero-field landscape. The deterministic effect of pulse length overrides thermal effects in this case. This mechanism has been captured in recent physical experiments [47].

One is motivated to reconsider the astroid for low damping and ultrashort pulses. A more complicated, "zebra astroid" emerges (for example, Figure 12), in which the pulse time plays a critical role, and switching is no longer monotonic in applied field magnitude, nor confined to the zero-temperature astroid. Switching depends sensitively on the pulse time, which controls the position of the magnetization along its orbit. Relationships among pulse time, anisotropy, applied field, and damping have been studied in [23], [34], for instance.

11. Outlook and Generalization

Switching probabilities for the one-dimensional equation (3.1) can be computed by brute force via the Fokker Planck equation [10], [19], [41]. The advantage of the perspective developed here, applying large deviation theory and ideas from stochastic resonance, is that it generalizes naturally to more complicated equations and even to spatially nonuniform systems. Relevant examples occur in many application areas. For instance, large deviation theory could be applied to spin-transfer-induced magnetic switching [40], although numerical study may be necessary to find the minimum action even for long-time switching. In biology, the dependence of unbinding force in dynamic force spectroscopy on the pulling force [2], [13] is analogous to the dependence of ramped astroids on ramping rate. The case of spatially distributed systems seems particularly rich, both because it raises deep issues of analysis and probability, and because it has strong links to many application areas, so we briefly comment upon it here.

Removing the assumption of uniform magnetization means reinstating the full Landau-Lifshitz-Gilbert (LLG) equations, which are difficult because of the nonlocal term in the effective field and, to a lesser degree, the norm constraint. Below, we discuss some issues related to nonuniform magnetization. We conclude with two subtopics that are largely open.

In extremely small samples, the exchange energy enforces approximately uniform magnetization; in larger samples, however, spatial variation plays an increasingly important role in magnetic switching pathways. A similar connection between sample size and spatial complexity of optimal pathways exists for the Allen-Cahn problem,

$$\dot{u} = \Delta u - V'(u) + \sqrt{2\varepsilon}\eta. \tag{11.1}$$

This stochastic gradient flow is a model that, while simpler than the stochastic LLG, captures many relevant phenomena [12], [28], [43], [44], including a crossover between spatially uniform and nonuniform optimal switching pathways as the sample size in-

creases [28], [33]. A crossover from spatial uniformity to nonuniformity of the minimal saddle has also recently been observed for a reduced micromagnetic energy, for a two-dimensional magnetic ring [35].

The weak volume dependence of saddle points in the spatially nonuniform landscape is important in LLG. Whereas the uniform model suggests a saddle-point energy that scales linearly with volume, this is far from true for typical samples. Under applied field pulses, the micromagnetic energy is biased, which tends to lead to saddle points whose energies depend only weakly on volume once the sample is larger than a "critical nucleus." The analogous situation exists for (11.1) when V is an *asymmetric* potential. The existence of the critical nucleus is well known in the physics community (see, for instance, [5]). In the PDE literature, the spherical symmetry of ground states and existence of a critical radius for (11.1) were studied, for instance, in [3], [20].

Pursuing the generalization of the uniform case to the nonuniform: The stochastic dynamics on nonuniform landscapes may be reduced to a Markov chain in the small-noise limit, as discussed in Section 9 for the uniform magnetic problem. In this case, the saddle points are spatially complex intermediate states, interesting in their own right. Formally at least, the switching probabilities and optimal switching pathways are studied via the action minimization problem, where now the minimization is over pathways through function space. Quantitative, numerical investigations of saddle points and the switching rates for LLG at zero applied field appear in [11], [43]. The optimal pathways connecting the states of the chain generate a host of questions, including the ones posed below. Both of the following subtopics find analogy in the Allen-Cahn problem, and both contain questions that, for LLG, are largely open.

Dependence of Optimal Pathways on Applied Field. Because the spatially extended landscape is so much richer than the uniform landscape, there are *many* saddle points, and many switching pathways. One question is *how the optimal pathway depends on the applied field*. Entirely different pathways may be selected as the applied field ranges over different values. It is pointless to obtain good action estimates for the wrong pathways! A different but related problem studied in [12], [28], [29], [44] is: How does the optimal pathway depend on the switching time and small parameter in the Allen-Cahn phase transformation problem?

Reduced Problem: Sharp Interfaces and Vortex Motion. Asymptotic limits of the scalar Ginzburg Landau energy are a familiar topic in the calculus of variations community. In [12], [28], [29], [44], related limits in the action minimization problem lead to a reduced, sharp-interface problem.

In certain limiting regimes of micromagnetics, metastable states are characterized by the locations of domain walls or boundary vortices. These limiting regimes have been studied in the physics community for a long time [4], [24], and have recently received a lot of attention in the mathematics community (see, for instance, [31], [38], [39]). It would be appealing to investigate an asymptotic reduction of the magnetic switching problem in which action-minimizing pathways are characterized by the motion of domain walls or boundary vortices.

Appendix A. Stochastic Landau-Lifshitz-Gilbert Dynamics

We claimed in Section 2 that it does not matter whether we add a random noise in both the gyromagnetic and the damping terms, as in (2.4), or in the gyromagnetic term alone, because the one-point statistics are the same. This is shown in [18], and later in [44]. In [44], the proof proceeds by changing to (θ, z) variables in both (2.4) and the Stratonovich equation

$$\dot{m} = m \times (-h + \sqrt{2\alpha\varepsilon}\, \dot{W}) - \alpha m \times (m \times h), \tag{A.1}$$

converting to Itô form, and showing that the resultant Fokker-Planck equation is the same as for the Itô equation (2.6).

Fokker-Planck Equation and Invariant Measure. We verify that the noisy LLG (2.6) are consistent with the Gibbs distribution, by showing that the Fokker-Planck equation for (2.6) may be written

$$p_t = \nabla \cdot (K \nabla E p + \varepsilon \sigma \sigma^T \nabla p), \tag{A.2}$$

and that a consequence is the stationarity of

$$p_s := Z^{-1} \exp(-E/\varepsilon).$$

(Z^{-1} is the normalization factor.) To proceed, we observe that (2.6) is a two-dimensional system for $X = (\theta, z)$, of the form

$$\dot{X} = -K \nabla E - \varepsilon c + \sqrt{2\varepsilon}\, \sigma \dot{W},$$

where

$$K = \begin{pmatrix} \dfrac{\alpha}{1-z^2} & 1 \\ -1 & \alpha(1-z^2) \end{pmatrix}, \qquad c = \begin{pmatrix} 0 \\ 2\alpha z \end{pmatrix}, \qquad \sigma = \begin{pmatrix} \dfrac{\sqrt{\alpha}}{\sqrt{1-z^2}} & 0 \\ 0 & \sqrt{\alpha}\sqrt{1-z^2} \end{pmatrix},$$

and $\dot{W}$ is a two-dimensional white noise with independent components. Decompose K into its symmetric and antisymmetric parts, $K = K^S + K^A$. Two distinguishing features of the system are

- $\sigma \sigma^T = K^S$.
- K^A is constant.

With this structure, being able to write the Fokker-Planck equation in the divergence form (A.2) suffices to confirm the stationarity of p_s, since

$$K^S \nabla E p_s + \varepsilon \sigma \sigma^T \nabla p_s = 0$$

identically, and

$$\nabla \cdot (K^A \nabla E p_s) = K^A : \nabla \nabla E p_s - \varepsilon^{-1} (\nabla E, K^A \nabla E) p_s = 0$$

by antisymmetry. Thus, we need only show (A.2). By direct calculation,

$$\begin{aligned} p_t &= \nabla\cdot(K\nabla Ep) + \partial^2_{\theta\theta}\left(\frac{\varepsilon\alpha}{1-z^2}p\right) + \partial^2_{zz}\left(\varepsilon\alpha(1-z^2)p\right) + \partial_z(2\varepsilon\alpha zp) \\ &= \nabla\cdot(K\nabla Ep) + \partial_\theta\left(\frac{\varepsilon\alpha}{1-z^2}\partial_\theta p\right) + \partial_z\left(\varepsilon\alpha(1-z^2)\partial_z p\right) \\ &= \nabla\cdot(K\nabla Ep + \varepsilon\sigma\sigma^T\nabla p). \end{aligned}$$

Appendix B. Large Deviation Theory

It is well-known that thermally activated transitions occur via saddle points, with an Arrhenius law that depends on the energy barrier. Part of the contribution of large deviation theory is to make these assertions rigorous, in the limit $\varepsilon \to 0$. Large deviation theory also gives more detailed information about rare events and the pathways by which they occur. While the exponential distribution applies for timescales of the order of the mean switching time, large deviation theory also estimates, for instance, the probability of an event to occur within a given time, T. This estimate, which relies on action minimization, was important for studying multiple-pulse astroids in Subsection 4.2.

To make sharp statements in the limit of vanishing noise strength, we have used results for the mean switching time, exponential distribution, and switching probability from large deviation theory. The results come from Freidlin and Wentzell [15] and Martinelli, Olivieri, and Scoppola [36]. In this appendix, we state the large deviation theorems for the process governed by (3.1). Analogous theorems hold for (2.6).

As usual, let $\theta_A, \theta_B \in [0, 2\pi)$ represent the two minima of the energy E_R with applied field H. We are interested in the probability of switching from the basin of attraction of the less-favored metastable state, θ_B, to either (a) the boundary of the basin of attraction of this state (which we call hitting), or (b) a small neighborhood of the energetically preferred metastable state, θ_A (which we call switching).

Large deviation theory [15] estimates the hitting and switching probabilities in terms of the Wentzell-Freidlin action. We make the following definitions.

$C_A :=$ the basin of attraction of θ_A $\qquad$ $C_B :=$ the basin of attraction of θ_B
$t_B(x) := \inf\{t \geq 0; \theta^\varepsilon(0) = x, \theta^\varepsilon(t) C_B\}$ $\qquad$ $\bar{t}_B(x) := \mathbb{E}[t_B(x)]$
$P_{hit}(x,t) := \text{Prob}(t_B(x) \leq t)$.

The Wentzell-Freidlin action functional for (3.1) is

$$S_{[0,T]}[\phi] = \frac{1}{4}\int_0^T |\dot{\phi}(s) + E_R'(\phi(s))|^2 ds.$$

The gradient structure simplifies the action minimization problem in the long-time limit (cf. Example B.1).

Mean Hitting Time. In the zero-temperature limit, the mean hitting time is independent of x, and is controlled by the energy barrier (cf. [15], Sections 4.3 and 4.4):

$$\lim_{\varepsilon\to 0} \varepsilon \ln \bar{t}_B = E_R(s) - E_R(\theta_B) =: \Delta E_R.$$

In other words, $\bar{t}_B \asymp e^{\Delta E_R/\varepsilon}$.

Hitting Probability. The probability of hitting the boundary of the basin of attraction within a given time T is found via an action minimization problem (cf. [15], Section 4.1). Specifically, for all $x \in C_B$,

$$\lim_{\varepsilon \to 0} -\varepsilon \ln P_{hit}(x, T) = \inf_{\phi \in F^x} S_{[0,T]}[\phi] =: S_T, \tag{B.1}$$

the *action* for the hitting event. Here, $F^x := \{\phi \in C[0, T]; \phi(0) = x, \phi(s) \notin C_B \text{ for some } s \leq T\}$. The pathway $\phi_\star$ that minimizes the action functional is the most likely pathway by which the stochastic process will switch, in the sense that the switching pathway stays within an arbitrarily small neighborhood of the action-minimizing path, with probability one as $\varepsilon \to 0$ (cf. [15], Section 4.2).

The more common, exponential estimate of hitting probabilities (see below) looks on a timescale which is of the order of the mean hitting time, so that $T \asymp \exp(\Delta E_R/\varepsilon)$. For ε-independent times, however, it is the action that gives the correct exponential factor, requiring action minimization, i.e., solving the deterministic variational problem in (B.1).

Exponential Distribution. In the limit as $\varepsilon \to 0$, appropriately rescaling t_B yields an exponential distribution [36]. To be precise,

$$\lim_{\varepsilon \to 0} \text{Prob}(t_B > c\,\bar{t}_B) = \exp(-c).$$

For finite epsilon, if we rescale time by the mean hitting time, t_B is approximately exponentially distributed. In other words, given $T = O(\bar{t}_B)$, we have the estimate

$$P_{hit}(T) \approx 1 - e^{-T/\bar{t}_B}.$$

We call the region of validity of this approximation in H_1-H_2 space the *exponential regime*. (On this timescale, the hitting probability is x-independent in the limit as $\varepsilon \to 0$.)

Mean Switching Time and Switching Probability. What about the event of *switching* rather than just *hitting*? The mean time to switch, τ, satisfies

$$\tau \asymp \exp(\Delta E_R/\varepsilon)$$

(logarithmic equivalence), which can be seen from a standard argument along the lines of [15]. Alternatively, this estimate is a corollary of the estimates above and the fact (cf. [44]) that within the exponential regime,

$$\tau = 2\bar{t}_B, \tag{B.2}$$

and thus the two are logarithmically equivalent. Also,

$$\text{Prob}(\text{switch to } C_A \text{ for } t \leq T) \asymp \exp\left(- \inf_{\substack{\phi(0)=x \\ \phi(T)\in C_A}} \frac{1}{4\varepsilon} \int_0^T |\dot{\phi}(s) + E_R'(\phi(s))|^2 ds \right).$$

Illustrative Examples. When the form of the deterministic equation is gradient or "gradient-plus-orthogonal," there are simplifications that allow us to determine the mean switching time and the most likely switching path (given an arbitrarily long time).

Example B.1 (Simplification for the gradient case). *Consider the noisy gradient flow,*

$$dX = -\nabla V(X)dt + \sqrt{2\varepsilon}dW.$$

Suppose that x_0 lies in the basin of attraction of m, an isolated minimum of V. Call the basin D_m. Suppose that ∂D_m is smooth and, furthermore, that there is a unique point $s \in \partial D_m$ that minimizes V. Consider the action to switch from $x(T_1) = m$ to $x(T_2) \in \partial D_m$. One manipulates the action functional:

$$\begin{aligned} S_{[T_1,T_2]}[x] &= \frac{1}{4}\int_{T_1}^{T_2} |\dot{x} + \nabla V(x)|^2 dt \\ &= \frac{1}{4}\int_{T_1}^{T_2} |\dot{x} - \nabla V|^2 + 4\langle \dot{x}, \nabla V\rangle dt \\ &= \frac{1}{4}\int_{T_1}^{T_2} |\dot{x} - \nabla V|^2 dt + V(x(T_2)) - V(m) \\ &\geq V(s) - V(m), \end{aligned}$$

the inequality following by dropping a positive term. From here, one identifies the long-time limit of the action as the energy barrier between the minimum and the minimal energy saddle, and the optimal long-time switching path as precisely that which satisfies

$$\dot{x} = \nabla V(x), \qquad x(-\infty) = m, \qquad x(\infty) = s,$$

i.e., the heteroclinic orbit of the deterministic gradient flow connecting m and s (in backwards time).

Example B.2 (Gradient-plus-orthogonal). *For the generalization*

$$\dot{X} = -\nabla V(X) + \ell(X) + \sqrt{2\varepsilon}\dot{W}, \qquad \text{with } \langle \nabla V, \ell\rangle = 0,$$

where the potential V has only isolated critical points, the story remains almost the same. Reversing the sign on ∇V as before and using the orthogonality condition, we discover that the system exits via the saddle at a rate with exponential factor $V(s) - V(m)$. The optimal path in this case satisfies

$$\dot{x} = \nabla V(x) + \ell(x), \qquad x(-\infty) = m, \qquad x(\infty) = s.$$

The structure of the equations assures the existence of such a path.

Prefactors and Kramers' Analysis. The prefactor for the switching event is derived by using Kramers' estimate for the hitting rates [37], [43], [48]. According to this analysis, the switching rate k_{m,s_i} from minimum m via the saddle s_i $(i = 1, 2)$ is given by

$$k_{m,s_i} = e^{-\varepsilon^{-1}\Delta E_{m,s_i}} \left((2\pi)^{-1}|\mu_{s_i}||\det(J_m J_{s_i}^{-1})|^{1/2} + O(\varepsilon)\right), \tag{B.3}$$

where J is the Hessian matrix of E, the subscript m denotes evaluation at the minimum, the subscript s_i denotes evaluation at the saddle point, μ_{s_i} is the unique negative eigenvalue of $K^T J_{s_i}$, and $\Delta E_{m,s_i}$ is the energy difference between the saddle and the minimum. Note that although the formula above was derived for the case with constant K, it holds for our case as well because of the convenient fact that the derivatives of K vanish at the saddle and therefore

$$\left.\frac{\partial(K^T \nabla E)}{\partial(\theta, z)}\right|_{s_i} = K^T J(s_i).$$

Using this identity and denoting the unit tangent vector to the least action path by $\hat{t}$, we obtain $\mu_{s_i}\hat{t}_{s_i} = K_{s_i}^T J_{s_i}\hat{t}_{s_i}$. This fact is enough to generalize constant K analysis from [43], for instance, to the magnetic case. From the switching rates, we calculate the mean switching time τ:

$$\tau = (k_{m,s_1} + k_{m,s_2})^{-1}. \tag{B.4}$$

Generically, the rates are widely separated and

$$\tau \approx (\max\{k_{m,s_1}, k_{m,s_2}\})^{-1}.$$

Note that for the reduced energy (3.2), the rate formula (B.3) reduces to the particularly simple estimate

$$k_{m,s_i} = e^{-\varepsilon^{-1}\Delta E_{m,s_i}} \left((2\pi)^{-1}\sqrt{E_R''(m)E_R''(s_i)}\right).$$

Acknowledgments

We would like to thank Weinan E, Geoff Grinstein, and Roger Koch for stimulating conversations that contributed to the development of this work. This article derives from work done as part of the Ph. D. thesis of M. G. Reznikoff, while partially supported by NSF grant DMS01-01439. R. V. Kohn was partially supported by NSF grants DMS01-01439 and DMS03-13744. E. Vanden-Eijnden was partially supported by NSF grants DMS01-01439, DMS02-09959, and DMS02-39625, and ONR grant N00014-04-1-0565.

References

[1] A. Aharoni, *Introduction to the Theory of Magnetism*, Clarendon Press, Oxford, 1998.
[2] D. Bartolo, I. Derényi, and A. Ajdari, *Dynamic response of adhesion complexes: Beyond the single-path picture*, Phys. Rev. E **65** (2002), 051910.

[3] H. Berestycki and P.-L. Lions, *Some applications of the method of super and subsolutions*, Bifurcation and Nonlinear Eigenvalue Problems, Lecture Notes in Math. **782** (1980), 16–41.
[4] G. Bertotti, *Hysteresis in Magnetism*, Academic Press, San Diego, 1998.
[5] H. B. Braun, *Fluctuations and instabilities of ferromagnetic domain-wall pairs in an external magnetic field*, Phys. Rev. B. **50** (1994), 16485–16500.
[6] G. Brown, M. A. Novotny, and P. A. Rikvold, *Thermal magnetization reversal in arrays of nanoparticles*, J. Appl. Phys. **89** (2001), 7588–7590.
[7] W. F. Brown, Jr., *Micromagnetics*, Robert E. Krieger Publishing Company, Huntington, 1978.
[8] W. F. Brown, Jr., *Thermal fluctuations of a single-domain particle*, Phys. Rev. **130** (1963), 1677–1686.
[9] M. V. Day, *On the exponential exit law in the small parameter exit problem*, Stochastics **8** (1983), 297–323.
[10] R. Durrett, *Stochastic Calculus: A Practical Introduction*, CRC Press, Boca Raton, 1996.
[11] W. E, W. Ren, and E. Vanden-Eijnden, *Energy landscape and thermally activated switching of submicron-size ferromagnetic elements*, J. Appl. Phys. **93** (2003), 2275–2282.
[12] W. E, W. Ren, and E. Vanden-Eijnden, *Minimum action method for the study of rare events*, Commun. Pure Appl. Math. **57** (2004), 637–656.
[13] E. Evans, *Probing the relation between force–lifetime– and chemistry in single molecular bonds.* Ann. Rev. Biophys. Biomol. Struct. **30** (2001), 105–127.
[14] M. I. Freidlin, *Quasi-deterministic approximation, metastability and stochastic resonance*, Physica D **137** (2000), 333–352.
[15] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer-Verlag, 2nd edition, New York, 1998.
[16] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, *Stochastic resonance*, Rev. Mod. Phys. **70** (1998), 223–242.
[17] C. J. García-Cervera and W. E, *Effective dynamics for ferromagnetic thin films*, J. Appl. Phys. **90** (2001), 370–374.
[18] J. L. García-Palacios and F. J. Lázaro, *Langevin-dynamics study of the dynamical properties of small magnetic particles*, Phys. Rev. B **58** (1998), 14937–14958.
[19] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, Springer-Verlag, Berlin, 1996.
[20] B. Gidas, W. M. Ni, L. Nirenberg, *Symmetry and related properties via the maximum principle*, Commun. Math. Phys. **68** (1979), 209–243.
[21] J. M. González, R. Ramírez, R. Smirnov-Rueda, and J. González, *Non-Arrhenius relaxation in micromagnetic models of systems with many degrees of freedom*, Phys. Rev. B **53** (1995), 16034–16040.
[22] G. Grinstein and R. H. Koch, *Switching probabilities for single-domain magnetic particles*, to appear.
[23] L. He and W. D. Doyle, *A theoretical description of magnetic switching experiments in picosecond field pulses*, J. Appl. Phys. **79** (1996), 6489–6491.
[24] A. Hubert and R. Schäfer, *Magnetic Domains: The Analysis of Magnetic Microstructures*, Springer-Verlag, New York, 1998.
[25] R. H. Koch, G. Grinstein, G. A. Keefe, Y. Lu, P. L. Trouilloud, W. J. Gallagher, and S. S. P. Parkin, *Thermally assisted magnetization reversal in submicron-sized magnetic thin films*, Phys. Rev. Lett. **84** (2000), 5419–5422.
[26] R. Koch, private communication.
[27] R. Koch's web site: www.research.ibm.com/people/c/coke/index.html.
[28] R. V. Kohn, F. Otto, M. G. Reznikoff, and E. Vanden-Eijnden, *Action minimization and sharp interface limits for the noisy Allen-Cahn equation*, submitted.
[29] R. V. Kohn, M. G. Reznikoff, and Y. Tonegawa, *The sharp interface limit of the Allen Cahn action in one space dimension*, submitted.
[30] R. V. Kohn and V. V. Slastikov, *Effective dynamics for ferromagnetic thin films: A rigorous justification*, Proc. Roy. Soc. London Ser. A **461** (2005), 143–154.
[31] M. Kurzke, *Analysis of Boundary Vortices in Thin Magnetic Films*, Ph. D. thesis, Universität Leipzig, 2003.

[32] D. Liu, *Topics in the Analysis and Computation of Stochastic Differential Equations*, Ph. D. thesis, Princeton University, 2003.
[33] R. S. Maier and D. L. Stein, *Droplet nucleation and domain wall motion in a bounded interval*, Phys. Rev. Lett. **87** (2001), 270601-1–4.
[34] J. C. Mallinson, *Damped gyromagnetic switching*, IEEE Trans. Magnetics **36** (2000), 1976–1981.
[35] K. Martens, D. L. Stein, and A. D. Kent, *Magnetic reversal in mesoscopic ferromagnetic rings*, submitted. Available via cond-mat archive cond-mat/0410561.
[36] F. Martinelli, E. Olivieri, and E. Scoppola, *Small random perturbations of finite- and infinite-dimensional dynamical systems: Unpredictability of exit times*, J. Stat. Phys. **55** (1989), 477–504.
[37] B. J. Matkowsky and Z. Schuss, *The exit problem for randomly perturbed dynamical systems*, SIAM J. Appl. Math. **33** (1977), 365–382.
[38] R. Moser, *Boundary vortices for thin ferromagnetic films*, Arch. Ratl. Mech. Anal. **174** (2004), 267–300.
[39] R. Moser, *Ginzburg-Landau vortices for thin ferromagnetic films,* Appl. Math. Res. Ex. **1** (2003), 1–32.
[40] E. B. Myers, F. J. Albert, J. C. Sankey, E. Bonet, R. A. Buhrman, and D. C. Ralph, *Thermally activated magnetic reversal induced by a spin-polarized current*, Phys. Rev. Lett. **89** (2002), 196801-1–4.
[41] B. Øksendal, *Stochastic Differential Equations*, Springer-Verlag, Berlin, 2000.
[42] G. A. Prinz, *Magnetoelectronics*, Science **282** (1998), 1660–1663.
[43] W. Ren, *Numerical Methods for the Study of Energy Landscapes and Rare Events*, Ph. D. thesis, New York University, 2002.
[44] M. G. Reznikoff, *Rare Events in Finite and Infinite Dimensions*, Ph. D. thesis, New York University, 2004.
[45] P. A. Rikvold, G. Brown, S. J. Mitchell, and M. A. Novotny, *Dynamics of magnetization reversal in models of magnetic nanoparticles and ultrathin films*, in Nanostructured Magnetic Materials and Their Applications, edited by D. Shi, B. Aktas, L. Pust, and F. Mikailov, Springer Lecture Notes in Physics **593** (2002), 164-182.
[46] S. E. Russek, S. Kaka, and M. J. Donahue, *High-speed dynamics, damping, and relaxation times in submicrometer spin-valve devices*, J. Appl. Phys. **87** (2000), 7070–7072.
[47] H. W. Schumacher, C. Chappert, P. Crozat, R. C. Sousa, P. P. Freitas, J. Miltat, J. Fassbender, and B. Hillebrands, *Phase coherent precessional magnetization reversal in microscopic spin valve elements*, Phys. Rev. Lett. **90** (2003), 017201-1–4.
[48] Z. Schuss, *Singular perturbation methods in stochastic differential equations of mathematical physics*, SIAM Review **22** (1980), 119–155.
[49] Special Issue on Magnetoelectronics, Phys. Today **48** (1995).
[50] J. Z. Sun, J. C. Slonczewski, P. L. Trouilloud, D. Abraham, I. Bacchus, W. J. Gallagher, J. Hummel, Y. Lu, G. Right, S. S. P. Parkin, and R. H. Koch, *Thermal activation-induced sweep-rate dependence of magnetic switching astroid*, Appl. Phys. Lett. **78** (2001), 4004–4006.
[51] A. Thiaville, *Coherent rotation of magnetization in three dimensions: A geometric approach*, Phys. Rev. B **61** (2000), 12221–12232.
[52] W. Wernsdorfer, *Classical and quantum magnetization reversal studied in nanometer-sized particles and clusters*, Adv. Chem. Phys., Wiley, 2nd edition, 2000.