

Rare Events, Action Minimization, and Sharp Interface Limits

Draft of proceedings for CRM workshop *Singularities in PDE and the Calculus of Variations*

Maria G. Westdickenberg
Georgia Institute of Technology

January 14, 2008

Abstract

Estimating the probability of and mechanism for rare events is a physically relevant and mathematically interesting question. Focussing on issues related to “constrained switching events,” we describe differences from the classical Arrhenius picture and how action minimization can be used as a tool. As a specific example, we turn to the stochastic Allen-Cahn equation and explain recent results for the scaling of the action problem and the sharp interface limit.

1 Motivation

Small random forces are unexpectedly powerful: In so-called rare events, a small perturbation drives a system to violate deterministic behavior in a “large” way. We set the stage with a few examples.

Three examples

As our canonical example, consider the deterministic ordinary differential equation (ODE)

$$\dot{x}(t) = x(t) - x(t)^3, \quad x(0) = x_0, \quad (1)$$

which is the gradient flow for the double-well potential

$$V(x) = \frac{1}{4}(1 - x^2)^2. \quad (2)$$

Notice that the energy has two absolute minimizers, $x_- = -1$ and $x_+ = +1$. Equation (1) is tame: Choosing $x_0 = x_-$ leads to $x(t) \equiv x_-$, and choosing x_0 near x_- leads to the relaxation $x(t) \rightarrow x_-$ as $t \rightarrow \infty$. Even an arbitrarily small stochastic perturbation spices things up, however, destroying the stationarity and stability of x_- . Consider the stochastic ODE

$$dX(t) = (X(t) - X(t)^3) dt + \sqrt{2\delta} dB(t), \quad X(0) = x_0,$$

where $B(t)$ is a Brownian motion and $\delta > 0$. Because of the perturbation, the solution for $x_0 = x_-$ is not stationary, but instead makes small, random hops. The more dramatic change, however, is that with probability one, $X(t)$ will eventually visit a *small neighborhood* of x_+ . In other words, an order $\sqrt{\delta}$ perturbation will force an order 1 increase in energy,

driving the process against the gradient flow and over the energy barrier. The expected value of the time for this switching event to occur, which we denote by τ_δ , is exponentially large in δ^{-1} (hence the descriptor “rare”). Even for much shorter timescales T , however, the probability of the switching event occurring for $t \leq T$ is positive. We will be interested in estimating switching probabilities.

New structure emerges in the two-dimensional analogue. Consider the energy with level curves as in Figure 1. We again denote the two absolute minimizers by x_- and x_+ . The stochastically perturbed gradient flow is

$$dX(t) = -\nabla V(X(t)) dt + \sqrt{2\delta} dB(t), \quad X(0) = x_0, \quad (3)$$

where $X(t), x_0 \in \mathbb{R}^2$ and $B(t)$ is a vector of two independent Brownian motions. As in the first example, we can ask about the probability to switch from x_- to a neighborhood of x_+ in time $t \leq T$. A new question that becomes relevant in this example is *how* the process switches. It turns out that there is a distinguished pathway, the so-called most likely switching pathway, such that when the process does transition from one minimizer to the other, it almost certainly stays close to this pathway. We will be interested in identifying switching pathways.

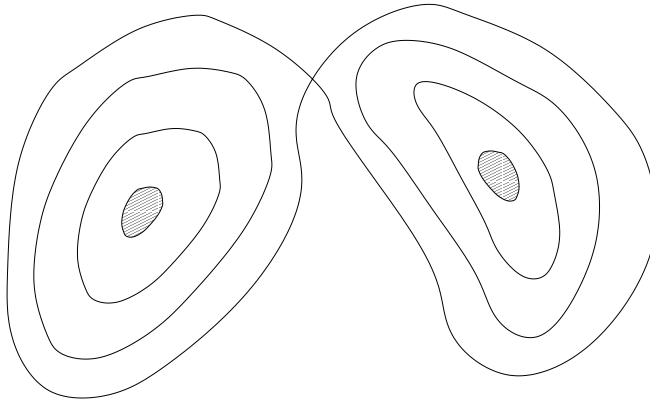


Figure 1: The level curves of a potential $V : \mathbb{R}^2 \rightarrow \mathbb{R}$. The shaded regions represent small neighborhoods of the absolute energy minimizers.

For the third example, we turn to the physical problem of developing reliable, nanoscale sized magnetic memory devices. In small devices, thermal noise is significant: Finite temperature effects can cause magnetic switching on physically relevant timescales. Micromagnetics is a continuum theory for magnetic behavior in terms of the magnetization vector $m : \Omega \rightarrow \mathbb{S}^2$, where Ω represents the magnetic sample and \mathbb{S}^2 represents the unit sphere. Associated to a given magnetization is the micromagnetic energy, which includes an exchange term that penalizes spatial variation, an anisotropy term that prefers certain orientations, and a nonlocal stray field term that favors divergence-free configurations. This rich energy functional has a host of local energy minimizers; an example of one minimizer in a thin-film sample is shown in Figure 2. The associated dynamic equations are the stochastically perturbed Landau-Lifshitz-Gilbert equations. We emphasize two features of the stochastic switching problem: (i) critical points of the energy have a nontrivial spatial structure, and (ii) the most likely switching pathway depends on a rich parameter regime including, for instance, the size of the sample, the time of the experiment, and the coefficient of the

exchange term. We will be interested in how spatial structure and competition between system parameters affect the switching problem.

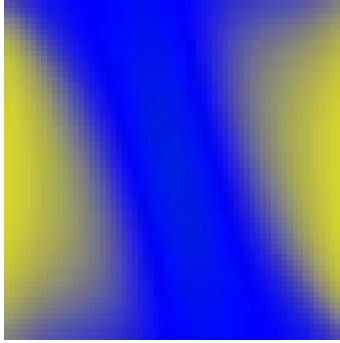


Figure 2: A local minimizer of the micromagnetic energy in a $200 \text{ nm} \times 200 \text{ nm} \times 10 \text{ nm}$ thin film sample. Here blue represents $(1, 0, 0)$, and yellow represents $(0, 1, 0)$. This figure has been calculated numerically by Weiqing Ren, and is reproduced with his permission. cf. <http://math.nyu.edu/~weiqing/MM/MM.html>

Timescales and switching probabilities

There is a classical analysis of the switching problem, well-known in the physics community and later analyzed in the mathematics literature in the 1980s, that we now describe. To fix ideas, consider (3). Let

$$\mathcal{A}^\epsilon = \{x : \mathbb{R}^+ \rightarrow \mathbb{R}^2 \mid x(0) = x_-, x(t) \in B_\epsilon(x_+) \text{ for some } t \leq T\},$$

where $\epsilon > 0$ and $B_\epsilon(x_+)$ denotes a ball of radius ϵ around x_+ . In this case we have

$$\mathbb{E}(\inf\{T \text{ such that } X \in \mathcal{A}^\epsilon\}) =: \tau_\delta \asymp \exp(\delta^{-1}\Delta E)$$

where the energy barrier ΔE is the difference in energy of x_- and the saddle point with smallest energy, and $f_\delta \asymp g_\delta$ denotes logarithmic equivalence, i.e., $\ln f_\delta \sim \ln g_\delta$ as $\delta \rightarrow 0$. The classical analysis says that the process is governed by the Arrhenius law

$$\text{Prob}(X \notin \mathcal{A}^\epsilon) \approx \exp(-T/\tau_\delta). \quad (4)$$

This statement was put on mathematically rigorous ground in [GOV]: Roughly speaking, (4) is proved *in the case* $T \asymp \tau_\delta$. However, and this will be important for us, the Arrhenius approximation (4) gives no information in the case $\ln T \ll \ln \tau_\delta$.

At first glance, considering switching on the timescale $T \asymp \tau_\delta$ may seem like the *only* reasonable question: Switching for $\ln T \ll \ln \tau_\delta$ is far less likely than switching for $T \asymp \tau_\delta$. Indeed, if one starts the dynamics and waits for a rare event, the timescale τ_δ is the relevant one. On the other hand, one can also fix $T > 0$ and decide a priori to wait *only* that long to see whether there is a rare event. Why do that? One motivation comes from the magnetic problem: Nanoscale magnetic devices are subjected to short, applied-field pulses that are intended to switch the magnetization in certain elements, but not others. The “non-targeted elements” must reliably withstand the pulses without switching. The pulse time T will be short compared to their expected switching time, so the probability of switching is exceedingly small, but there will be many of these pulses in their lifetime, so even a small

probability matters. What one seeks in this case is a quantitative estimate of the very small switching probability. See [KRV] for further discussion.

The Freidlin-Wentzell theory of large deviations [FW] analyzes switching probabilities for fixed time T in terms of the so-called large deviation action functional. For (3), the statement is

$$\begin{aligned} \text{Prob}(X \in \mathcal{A}^\epsilon) &\asymp \exp\left(-\delta^{-1} \inf_{\mathcal{A}^\epsilon} S(x)\right), \\ \text{where } S(x) &= \frac{1}{4} \int_0^T |\dot{x}(t) + \nabla V(x(t))|^2 dt. \end{aligned} \tag{5}$$

Freidlin-Wentzell theory is not the only tool by which the switching problem for (3) can be analyzed; indeed, (5) only estimates the exponential factor in the switching probability, and more refined estimates are possible by other methods. See for instance [BEGK]. What is remarkable, however, and the source of our interest in the tools of large deviation theory, is how naturally statements such as (5) generalize to the case of stochastic *partial* differential equations [FJ, F, DZ].

Timescales and switching pathways

The timescale of the experiment is also critical in determining the most likely switching pathway. It is almost folkloric that the optimal pathway for switching “should follow the time-reversed gradient flow connecting x_- and the minimal energy saddle point.” For T sufficiently large, this fact can be deduced from large deviation theory, as we now explain. Large deviation theory says that the action minimizer is the most likely switching pathway. In other words, if $X \in \mathcal{A}^\epsilon$, then with high probability, X stays within a small neighborhood of the pathway in \mathcal{A}^ϵ that minimizes $S(\cdot)$; cf. [FW, Chapter 3, Theorem 3.4]. On the other hand, for any $s \leq T$, we have

$$\begin{aligned} S(x) &= \frac{1}{4} \int_0^T |\dot{x}(t) + \nabla V(x(t))|^2 dt \\ &\geq \frac{1}{4} \int_0^s |\dot{x}(t) + \nabla V(x(t))|^2 dt \end{aligned} \tag{6}$$

$$\begin{aligned} &= \frac{1}{4} \int_0^s |\dot{x}(t) - \nabla V(x(t))|^2 dt + \int_0^s \dot{x}(t) \cdot \nabla V(x(t)) dt \\ &\geq \int_0^s \dot{x}(t) \cdot \nabla V(x(t)) dt = V(x(s)) - V(x(0)). \end{aligned} \tag{7}$$

Hence, (7) implies that the action cost is at least as large as the difference in energy between x_- and the minimal energy saddle point. Moreover, this bound can be achieved if (i) the inequality in (7) can be made tight by satisfying the time-reversed gradient flow:

$$\dot{x}(t) = +\nabla V(x(t)) \quad \text{for } t \leq s \tag{8}$$

and (ii) the bound in (6) can be made tight by satisfying

$$\dot{x}(t) = -\nabla V(x(t)) \quad \text{for } s < t \leq T. \tag{9}$$

Using this construction actually requires $T = \infty$, but for T sufficiently large compared to the timescale of the deterministic gradient flow, an approximation of this construction (with patching near the minima and saddle point) gives a good bound.

The classical picture changes when one considers shorter timescales. Indeed, it is easy to see via a scaling argument that for $T \downarrow 0$,

$$S(x) \sim \frac{1}{T} \int_0^T |\dot{x}(t)|^2 dt,$$

so that the action minimizer in \mathcal{A}^ϵ is a straight line connecting x_- and the closest point in $B_\epsilon(x_+)$. Similarly, in the case of magnetic switching, it has been observed that switching probabilities in short-time experiments and long-time experiments are different (see [KRV] and the references therein). In particular, the *pathway* in short-time experiments seems to be different. The optimal short-time switching pathway involves new spatial structure that seems to be related to the time of the experiment.

We will be interested in using action minimization to predict most likely switching pathways. We will be particularly interested in quantifying the way that the spatial structure of the switching pathway depends on the parameter regime.

2 Simplest interesting PDE example

Because we are interested in spatial structure, we want to consider a stochastic *partial* differential equation. Because the magnetic problem is complicated (vector-valued order parameter, nonconvex constraint $|m| = 1$, nonlocal stray field term), we want to study a simpler problem. We consider the simplest interesting PDE example, namely the stochastic Allen-Cahn equation.

2.1 Introduction

The deterministic Allen-Cahn equation

$$\dot{u} = \Delta u - V'(u) \tag{10}$$

is the L^2 -gradient flow for the energy

$$E(u) = \int_{\Omega} \left(\frac{|\nabla u|^2}{2} + V(u) \right) dx. \tag{11}$$

The energy consists of two terms, a gradient term that penalizes spatial variation and a potential term in which the minima of V represent the preferred states of the order parameter. For simplicity, let V be given by (2) and let $\Omega = [0, L]^d$. The stochastic PDE is

$$\dot{u} = \Delta u - V'(u) + \sqrt{2\delta} \eta. \tag{12}$$

Here δ is a measure of the temperature of the system and η is a space-time noise. Choosing a space-time *white* noise is natural in the sense that formally, the associated invariant measure is the Gibbs distribution with the energy (11) and temperature δ . However a solution to (12) with white noise exists only in the case $d = 1$.

Consider the functional

$$S(u) = \frac{1}{4} \int_0^T \int_{[0, L]^d} (\dot{u} - \Delta u + V'(u))^2 dx dt. \tag{13}$$

Notice that (13) depends on two parameters, L and T . That (13) is the large deviation action functional for (12) with white space-time noise in dimension $d = 1$ is proved in [FJ].

In higher space dimensions, some regularization of the noise in space is necessary. We are interested in (13) in $d > 1$ because we expect it to capture the leading order behavior of (12) in an appropriate regime in the space of L , T , δ , and the regularization parameter of the noise.

With the stochastic problem as motivation, we study the following leading order problem. We replace \mathcal{A}^ϵ by

$$\mathcal{A} := \{u \in C^\infty([0, L]^d \times [0, T] \mid u(\cdot, 0) \equiv -1, u(\cdot, T) \equiv +1)\}. \quad (14)$$

To fix ideas, we equip \mathcal{A} with Neumann boundary conditions; periodic or Dirichlet boundary conditions are also possible. We are interested in the value and minimizers of the action:

$$\mathcal{S} := \inf_{\mathcal{A}} S(u). \quad (15)$$

2.2 Scaling regimes and heuristics

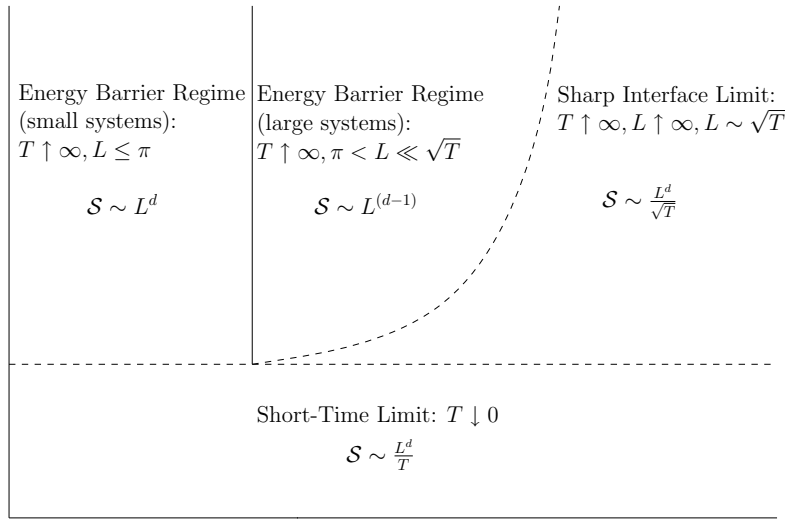


Figure 3: The scaling of the action in the four parameter regimes.

Four scaling regimes for the problem (15) were analyzed in [KORV]. See Figure 3 for an overview. The method in each case is to develop an upper bound and a lower bound that agree on the level of scaling. The first ingredient is an explicit construction; the value of $S(\cdot)$ for this construction gives an upper bound for the infimum (15). One tries to produce a bound that is close to optimal by using heuristics to develop a good construction. The second ingredient is to develop a lower bound for the infimum that says “no function can do better than the construction.” We summarize the results below; see [KORV] for details.

Short-time limit: $T \downarrow 0$

In this limit, the action scales as

$$\mathcal{S} \sim \frac{L^d}{T}.$$

The heuristic, as discussed above, is that for $T \downarrow 0$, the $(\dot{u})^2$ term dominates the action. The construction in this case is constant in space and linear in time. The lower bound uses the spatial mean and Jensen's inequality.

Energy barrier regime for small systems: $T \uparrow \infty$, $L \leq \pi$

In this limit, the action scales as

$$\mathcal{S} \sim L^d.$$

Because T is large, we expect to use the long-time most likely switching pathway, cf. Section 1. Because the system is small, the minimal energy saddle point is $u \equiv 0$. Since the path switches between spatially uniform critical points *via* a spatially uniform critical point, the heuristic is that developing spatial structure is unnecessary. Hence, the construction is constant in space, as in the short-time limit, but the time dynamic is completely different: $u(t)$ uses the gradient flow dynamics (1) in backward and then forward time, according to the heuristics discussed near (8) and (9) above.

Energy barrier regime for large systems: $T \uparrow \infty$, $\pi < L \ll T^{1/2}$

In this limit, the action scales as

$$\mathcal{S} \sim L^{d-1}.$$

For $L > \pi$, the minimal energy saddle point is spatially nonuniform and its energy scales like perimeter: L^{d-1} . This fact provides an easy lower bound in this region, as in (7). Because the saddle is nonuniform, any nearly optimal construction is also expected to be nonuniform. Extending the ideas used for small systems, one would like to use the time-reversed gradient flow connecting $u \equiv -1$ to the saddle. These ideas work fine when L is fixed or bounded. We have to be careful when considering limits in which $L \uparrow \infty$, however.

The ‘‘catch’’ for $L \uparrow \infty$ is that the deterministic timescale of this pathway is extremely slow; this is due to the slow motion of transition layers in the one-dimensional deterministic Allen Cahn equation [CP, FH, BK]. (Even in $d > 1$, the minimal energy saddle point seems to have one-dimensional structure [KORV].) Hence, we would expect only to be able to use such a pathway for L small enough, which means $L \sim \log T$ in this case. Indeed, for $L \gg \log T$, we are forced to use constructions that artificially accelerate the motion of the deterministic gradient flow. When L grows sufficiently quickly with respect to T , accelerating these transition layers will lead to an additional cost in the action functional. However it turns out that as long as $L \ll \sqrt{T}$, this additional propagation cost is negligible; see [KORV] for details.

Sharp interface limit: $T \uparrow \infty$, $L \uparrow \infty$, $L \sim T^{1/2}$

In this limit, the action scales as

$$\mathcal{S} \sim \frac{L^d}{T^{1/2}}. \tag{16}$$

Heuristically, the sharp interface limit consists of a limit in which the system size grows so quickly with respect to T that the optimal action includes a propagation cost. Because the system is so large, the optimal path is expected to have the following character: A small region of $+1$ is formed somewhere in the system, connected to the surrounding -1 phase by transition layers. The transition layers then spread across the system, increasing the support of the $+1$ phase (cf. Figures 4 and 6 below). The propagation cost is the cost in the action functional that is associated to forcing the transition layers to cover the large system in the relatively short time T . The nucleation cost, on the other hand, is the action cost to form the transition layers. The optimal configuration balances the competing costs

of nucleation and propagation, forming additional transition layers in larger systems so that each does not have to cover as much distance. For larger and larger systems there is a “cascade” of optimal configurations with more and more spatial structure. These ideas were first explored in [ERV], using a combination of numerical investigation and formal arguments. See also [FHS] for the one dimensional case.

In [KORV], the scaling (16) is proved by the method of upper and lower bounds. The construction used to develop the upper bound has an arbitrary number N of transition layers, each of which has the shape of an *energy minimizer* (i.e., a hyperbolic tangent), and the layers are propagated across the system with a constant velocity. Elementary calculations produce a lower bound that matches in terms of scaling. To develop a *sharp* lower bound, however, requires a more careful analysis. We describe this work below.

3 The sharp interface limit

The sharp interface limit has its name because on the scale $L \gg 1$, the order-one transition layers look like sharp interfaces connecting ± 1 . It will be convenient to rescale so that the interface length goes to zero while the global length and time scales of the system remain finite; let

$$\varepsilon := L^{-1} \quad \text{and} \quad x \rightarrow \varepsilon x, \quad t \rightarrow \varepsilon^2 t.$$

Under this rescaling, the action functional (13) becomes $\varepsilon^{-(d-1)} S_\varepsilon(u)$, where

$$S_\varepsilon(u) = \frac{1}{4} \int_0^T \int_{[0,L]^d} \left(\varepsilon^{1/2} \dot{u} - \varepsilon^{-1/2} (\varepsilon \Delta u - \varepsilon^{-1} V'(u)) \right)^2 dx dt. \quad (17)$$

According to (16), the action value

$$\mathcal{S} \sim \varepsilon^{-(d-1)},$$

so that in particular, the infimum

$$\inf_{u \in \mathcal{A}} S_\varepsilon(u) \quad (18)$$

is expected to have a well-defined limit as $\varepsilon \downarrow 0$. We are interested in the minimum value and the action minimizers. More broadly, we would like to identify the Γ -limit of $S_\varepsilon(\cdot)$.

3.1 Connections with other sharp interface problems

This sharp interface limit of the action functional is related to a family of previously studied sharp interface problems, namely, the sharp interface limit of the energy functional

$$E_\varepsilon(u) = \int_{\Omega_L} \left(\frac{\varepsilon}{2} |\nabla u|^2 + \varepsilon^{-1} V(u) \right) dx \quad (19)$$

and the sharp interface limit of the associated gradient flow

$$\dot{u} = \varepsilon \Delta u - \varepsilon^{-1} V'(u). \quad (20)$$

The Γ -convergence of (19) to the perimeter functional is studied in [MM, M]. Roughly speaking, Γ -convergence is a notion of convergence for functionals that implies in particular that minimizers of the finite ε problem converge to minimizers of the limit problem as $\varepsilon \downarrow 0$.

(See [B] for a nice introduction.) In order for the second term in the energy to remain finite as $\varepsilon \downarrow 0$, u must converge almost everywhere to ± 1 . The Γ -convergence of $E_\varepsilon(\cdot)$ to the perimeter functional is a way of making precise the idea that the leading order contribution to the energy in the limit $\varepsilon \downarrow 0$ is the length of the interface separating the two phases.

Given the convergence of $E_\varepsilon(\cdot)$, a natural question is whether the gradient flow (20) converges to the gradient flow of the perimeter functional, i.e., flow by mean curvature: interfaces with normal velocity v_n and mean curvature κ evolving according to

$$v_n = -\kappa. \tag{21}$$

The answer is yes, although only after rescaling time by ε : The dynamics (20) leads to the algebraically slow mean curvature flow $v_n = -\varepsilon\kappa$. See [RKS, DS, ESS, I].

The case of one space dimension is special, since the ‘‘perimeter of the interface’’ reduces to the number of jump discontinuities of u . The limiting energy is degenerate in the sense that the leading order energy of two N -layer configurations is the same, regardless of the location of the layers. Because of this degeneracy, the associated dynamics (20) in $d = 1$ features transition layers that move *exponentially* slowly with respect to ε^{-1} and the distance between layers [CP, FH, BK, OR].

With this background in mind, we would like to ask whether the sharp interface limit of (17) is, in some sense, the action functional corresponding to the sharp interface dynamics (21). More generally, we would like to understand connections among the sharp interface limits of (19), (20), and (17). One expects the situation in $d = 1$ to be special because of the degenerate energy and exponentially slow dynamics.

3.2 Rigorous results in $d = 1$

The sharp interface limit of the upper bound constructions from [KORV] is sketched in Figure 4. In [KRT], a sharp lower bound on the action value is proved. Assuming Neumann boundary conditions, it is shown that for any sequence of smooth functions u_ε , one has

$$\liminf_{\varepsilon \rightarrow 0} S_\varepsilon(u_\varepsilon) \geq c_0 \min_{N \in \mathbb{Z}^+} \left(N + \frac{L^2}{4NT} \right),$$

where

$$c_0 = \int_{-1}^{+1} \sqrt{2V(u)} du = \frac{2\sqrt{2}}{3}. \tag{22}$$

Since this lower bound matches the upper bound from [KORV], this proves that the constructions with sharp interface limit sketched in Figure 4 do achieve the optimal value.

Are the constructions used for the upper bound in fact *the only* configurations that achieve the optimal action? The answer is yes, which follows from the Γ -convergence of $S_\varepsilon(\cdot)$ proved in [TW]. Letting $S_0(\cdot)$ and \mathcal{A}_0 denote an appropriately defined limit functional and space, respectively, the upper and lower bounds of Γ -convergence are proved, namely:

Theorem 1 (Upper bound). *For every $u_0 \in \mathcal{A}_0$, there exists a sequence $u_\varepsilon \in \mathcal{A}$ with $u_\varepsilon \rightarrow u_0$ in $L^\infty(L^2)$ such that*

$$\limsup_{\varepsilon \rightarrow 0} S_\varepsilon(u_\varepsilon) \leq S_0(u_0).$$

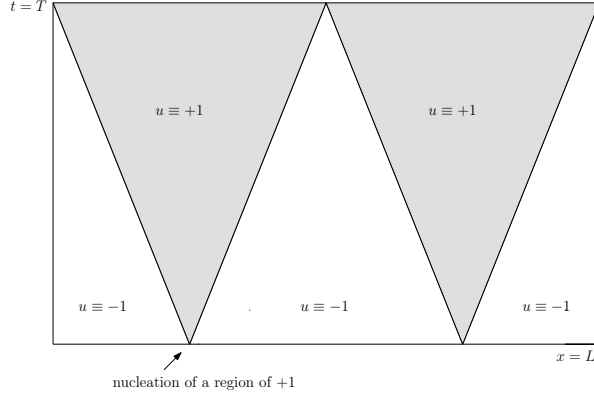


Figure 4: The sharp interface limit of the upper bound constructions. The N interfaces are formed at $t = 0$ and move linearly across the system in time T .

Theorem 2 (Lower bound). *Given any sequence $u_\varepsilon \in \mathcal{A}$ with*

$$\limsup_{\varepsilon \rightarrow 0} S_\varepsilon(u_\varepsilon) < \infty \quad \text{and} \quad u_\varepsilon \rightarrow u_0 \text{ in } L^\infty(L^2),$$

we have that $u_0 \in \mathcal{A}_0$ and moreover,

$$\liminf_{\varepsilon \rightarrow 0} S_\varepsilon(u_\varepsilon) \geq S_0(u_0). \quad (23)$$

In order to explain S_0 and \mathcal{A}_0 , it is convenient to begin by explaining a bit about the method of proof in [KRT] and [TW]; see Subsection 3.3 below.

3.3 Details and method

It is an elementary observation that for any sequence of functions u_ε with uniformly bounded action, the energy $E_\varepsilon(u_\varepsilon(\cdot, t))$ is uniformly bounded for every $t \in [0, T]$. Therefore it makes sense to consider the energy measures

$$d\mu_\varepsilon^t := \left(\frac{\varepsilon}{2}(u_x)^2 + \varepsilon^{-1}V(u) \right) dx$$

together with the space-time energy measure

$$d\mu_\varepsilon := \left(\frac{\varepsilon}{2}(u_x)^2 + \varepsilon^{-1}V(u) \right) dx dt$$

and action measure

$$d\nu_\varepsilon := \frac{1}{4} \left(\varepsilon^{1/2} \dot{u} - \varepsilon^{-1/2} (\varepsilon \Delta u - \varepsilon^{-1} V'(u)) \right)^2 dx dt.$$

In [KRT], the behavior of the measures μ_ε^t and μ_ε are studied using tools from [HT] and [T]. It is proved that the measures do what one would expect: μ_ε^t converges to a finite sum of delta masses with coefficient $c_0 J$ where J is a positive integer and c_0 is given by (22). A delta mass with $J = 1$ at location x occurs when the limit function u_0 has a jump discontinuity at x and u_ε has a simple transition layer structure near x as $\varepsilon \downarrow 0$. Higher multiplicity (i.e., $J > 1$) of a delta mass is also possible if u_ε has multiple transition layers

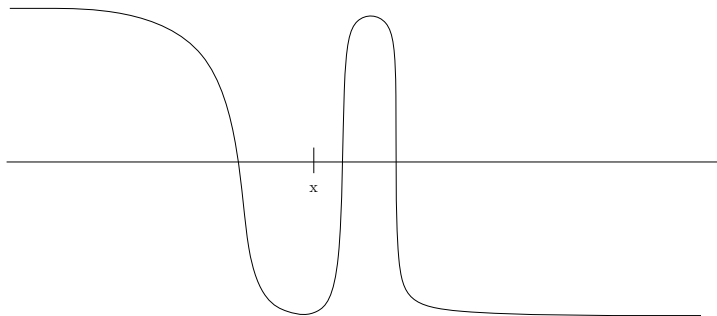


Figure 5: Three interfaces accumulating at the same point x in the limit. The energy will have a delta mass at x with weight $3c_0$.

converging to the same point x in the limit; cf. Figure 5. For J odd, the limit function u_0 has a jump discontinuity at x , while for J even there is no jump in u_0 at x (a “hidden” interface).

A lower bound on the action functional can be derived using the convergence of the measures $\mu_\varepsilon^t, \mu_\varepsilon, \nu_\varepsilon$. In [KRT], the lower bound is suboptimal for higher multiplicity interfaces; hence, an assumption of single multiplicity interfaces is made. In [TW], the single multiplicity assumption is removed, using an improved technique for the lower bound and properties of continuous functions in $d = 1$. The improved (optimal) bound can be described roughly in the following way. Let the sequence u_ε converge in $L^\infty(L^2)$ to u_0 with $u_0(\cdot, t) = \pm 1$ almost everywhere in space for every time. Suppose for simplicity that the associated measures μ_ε^t converge to

$$c_0 \sum_{j=1}^N \delta_{g_j(t)}$$

where N is constant for $t \in [0, T]$ and $g_1(t) \leq g_2(t) \leq \dots \leq g_N(t)$. (Higher multiplicity means that two or more interfaces are equal.) Then the action satisfies the lower bound:

$$S_\varepsilon(u_\varepsilon) \geq c_0 N + \frac{c_0}{4} \int_0^T \sum_{j=1}^N (\dot{g}_j(t))^2 dt.$$

To make things precise in Theorems 1 and 2, we introduce the notion of admissible functions and compatible measures. The limiting functional on the space of compatible measures is defined as

$$S^\mathcal{M}(\mu) = \frac{1}{2} \sum_{k=1}^M \left| \mu^{T_k^+} - \mu^{T_k^-} \right|_{TV} + \frac{c_0}{4} \sum_{k=1}^{M-1} \int_{(T_k, T_{k+1})} \sum_{j=1}^{N(k)} (\dot{g}_j)^2 dt,$$

where $|\cdot|_{TV}$ denotes the total variation norm and

$$\mu^{T_k^+} = \lim_{t \downarrow T_k} \mu^t, \quad \mu^{T_k^-} = \lim_{t \uparrow T_k} \mu^t.$$

The limiting functional on the space of *functions* is then defined by taking the infimum over compatible measures:

$$S_0(u) := \inf_{\mu_u} S^\mathcal{M}(\mu_u).$$

We refer the reader to [TW] for full details.

3.4 Progress in $d > 1$

The results from [ERV] and the upper bound constructions from [KORV] suggest that the Γ -limit of $S_\varepsilon(\cdot)$ in $d > 1$ should have the form

$$S_{nuc} + \frac{c_0}{4} \int_0^T \int_{\Gamma(t)} (v_n + \kappa)^2 d\sigma dt, \quad (24)$$

where S_{nuc} represents a nucleation cost for introducing a $(d - 1)$ -dimensional interface; $\Gamma(t)$ represents the interface at time t ; and v_n, κ denote the normal velocity and mean curvature of the interface, respectively. Such a limit would answer in the affirmative the question at the end of Subsection 3.1: The sharp interface limit of the action functional is, formally, the action functional of a stochastically perturbed version of mean curvature flow. Moreover, (24) suggests that action minimizers should use reverse and forward curvature flow; see (8), (9) above and the accompanying discussion. A sketch is given in Figure 6. This is consistent with numerical observations [ERV]. For additional details, see [KORV].

Convergence of (17) to (24) is related to a conjecture of DeGiorgi that says, roughly speaking,

$$\varepsilon^{-1} \int_{\Omega} (\varepsilon \Delta u - \varepsilon^{-1} V'(u))^2 dx \xrightarrow{\Gamma} c_0 \int_{\Gamma} \kappa^2 d\sigma.$$

This conjecture has recently been proved in two and three space dimensions in [RS, NT]; see also [BM, Mo] for earlier work.

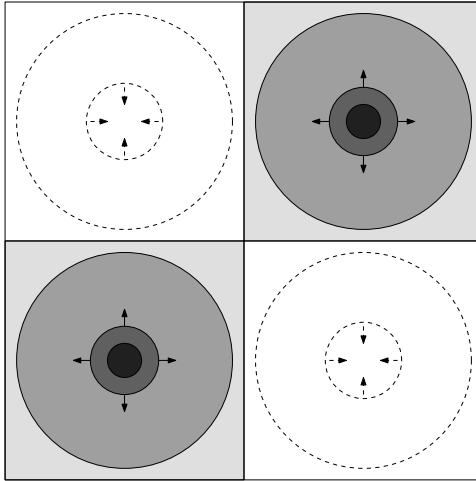


Figure 6: The sharp interface limit of a construction in $d = 2$. In the gray squares, interfaces grow by time-reversed curvature flow until a “checkerboard” pattern is formed. Then in the white squares, interfaces shrink by forward curvature flow.

4 Outlook

As explained above, the Allen Cahn problem has been studied here in part as a “warm up” problem with the full Landau Lifshitz Gilbert (LLG) equations as one long range goal or inspiration. Interesting numerical investigations of the magnetic problem include [ERV2,

LG]. To what extent can the analysis capture the phenomena observed in the numerics? Already in the Allen Cahn problem, we find that the spatial structure of action minimizers depends critically on parameter space. A similar phenomenon is analyzed in [DDP] for a nonlocal evolution equation in one space dimension. A natural next step is to analyze the action functional for other equations that are even more closely related to the LLG equations, for instance reduced models such as [K]. See also [N] for a generalization of the Allen Cahn result to the case in which V is a double-well potential with unequal wells.

Approaches based on large deviation theory can be useful for building efficient numerical schemes to quantify the effect of small noise [ERV, HV]. Although we have focussed on gradient systems, it is important to emphasize that gradient structure is not required by large deviation theory. See [HV] for examples. Numerical investigation of action minimization for physically relevant, nongradient systems such as spin-induced magnetic switching may produce interesting results.

There are many stochastic questions here that have been side-stepped. The theory of large deviations operates in the zero noise limit. We have considered *subsequently* sending the interface parameter to zero. Can one reverse the limit? Is there a stochastically perturbed sharp interface process that has S_0 as its action functional? An extra layer of questions arises in $d > 1$ about the role of the correlation length or regularization parameter of the noise.

Acknowledgements

We have described joint work with Bob Kohn, Felix Otto, Yoshihiro Tonegawa, and Eric Vanden-Eijnden, whose collaboration we gratefully acknowledge. We would also like to thank Weiqing Ren together with Weinan E and Eric Vanden-Eijnden for the numerical study that served as a basis for the work described in Section 3. Finally, we thank Geoff Grinstein and Roger Koch for inspiring discussions about magnetic switching phenomena. This work was partially supported by the National Science Foundation under Grant No. 0402762.

References

- [BM] Bellettini, G.; Mugnai, L. On the approximation of the elastica functional in radial symmetry. *Calc. Var. PDE* **24** (2005), 1–20.
- [BEGK] Bovier, B.; Eckhoff, M.; Gaynard, V.; Klein, M. Metastability in reversible diffusion processes I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc.* **6** (2004), 399–424.
- [B] Braides, A. *Γ -convergence for beginners*. Oxford University Press, Oxford, 2002.
- [BK] Bronsard, L.; Kohn, R. V. On the slowness of phase boundary motion in one space dimension. *Comm. Pure App. Math.* **43** (1990), 983–997.
- [CP] Carr, J.; Pego, R. L. Metastable patterns in solutions of $u_t = \epsilon^2 u_{xx} - f(u)$. *Comm. Pure App. Math.* **42** (1989), 523–576.
- [DZ] Da Prato, G.; Zabczyk, J. *Stochastic Equations in Infinite Dimensions*. Cambridge University Press, Cambridge, 1992.
- [DDP] De Masi, A; Dirr, N.; Presutti, E. Interface instability under forced displacements. *Ann. Henri Poincare* **7** (2006), 471–511.

- [DS] De Mottoni, P.; Schatzman, M. Evolution geometrique d’interfaces. *CRAS Ser. I Math.* **309** (1989), 453–458.
- [ERV] E, W.; Ren, W.; Vanden-Eijnden, E. Minimum action method for the study of rare events. *Comm. Pure App. Math.* **57** (2004) 637–656.
- [ERV2] E, W.; Ren, W.; Vanden-Eijnden, E. Energy landscape and thermally activated switching of submicron-size ferromagnetic elements. *J. App. Phys.* **93** (2003), 2275–2282.
- [ESS] Evans, L. C.; Soner, H. M.; Souganides, P. E. Phase transitions and generalized motion by mean curvature. *Comm. Pure App. Math.* **45** (1992), 1097–1123.
- [FJ] Faris, W. G.; Jona-Lasinio, G. Large fluctuations for a nonlinear heat equation with noise. *J. Phys. A: Math. Gen.* **15** (1982), 3025–3055.
- [FHS] Fogedby, H. C.; Hertz, J.; Svane, A. Domain wall propagation and nucleation in a metastable two-level system. *Phys. Rev. E* **70** (2004), 031105.
- [F] Freidlin, M. I. Random perturbations of reaction-diffusion equations: the quasideterministic approximation. *Trans. Amer. Math. Soc* **305** (1988), 665–697.
- [FW] Freidlin, M. I.; Wentzell, A. D.; *Random Perturbations of Dynamical Systems*. Second edition. Springer-Verlag, New York, 1998.
- [FH] Fusco, G.; Hale, J. K. Slow-motion manifolds, dormant instability, and singular perturbations. *J. Dynam. Differential Equations* **1** (1989), 75–94.
- [GOV] Galves, A.; Olivieri, E.; Vares, M. E. Metastability for a class of dynamical systems subject to small random perturbations. *Ann. Probab.* **15** (1987), 1288–1305.
- [HV] Heymann, M.; Vanden-Eijnden, E. The Geometric Minimum Action Method: A Least Action Principle on the Space of Curves, *Comm. Pure App. Math.*, to appear.
- [HT] Hutchinson, J.; Tonegawa, Y. Convergence of phase interfaces in the van der Waals-Cahn-Hilliard theory. *Calc. Var. PDE* **10** (2000), 49–84.
- [I] Ilmanen, T. Convergence of the Allen Cahn equation to Brakke’s motion by mean curvature. *J. Differential Geom.* **38** (1993), 417–461.
- [KORV] Kohn, R. V., Otto, F., Reznikoff, M. G., and Vanden-Eijnden, E. Action minimization and sharp-interface limits for the stochastic Allen-Cahn equation. *Comm. Pure Appl. Math.* **60** (2007), 393–438.
- [KRT] Kohn, R. V.; Reznikoff, M. G.; Tonegawa, Y. The sharp-interface limit of the action functional for Allen Cahn in one space dimension. *Calc. Var. PDE.* **25** (2006), 503–534.
- [KRV] Kohn, R. V.; Reznikoff, M. G.; Vanden-Eijnden, E. Magnetic elements at finite temperature and large deviation theory. *J. Nonlin. Sci.* **15** (2005), 223–253.
- [K] Kurzke, M. Boundary vortices in thin magnetic films. *Calc. Var. PDE* **26** (2006), 1–28.
- [LG] Liu, D.; Garcia-Cervera, C. Magnetic switching of ferromagnetic thin films under thermal perturbation. *J. Appl. Phys.* **98** (2005), 023903.
- [M] Modica, L. The gradient theory of phase transitions and the minimal interface criterion. *Arch. Rational Mech. Anal.* **98** (1987), 123–142.

- [MM] Modica L.; Mortola, S. Il limite nella Γ -convergenza di una famiglia di funzionali ellittici. *Boll. Un. Mat. Ital. A* **14** (1977), 526–529.
- [Mo] Moser, R. A higher order asymptotic problem related to phase transitions. *SIAM J. Math. Anal.* **37** (2005), 712–736.
- [N] Nagase, Y. The action minimization problem for an Allen Cahn equation with an unequal double-well potential, in preparation.
- [NT] Nagase, Y.; Tonegawa, Y. A singular perturbation problem with integral curvature bound, *Hiroshima Mathematical Journal*, to appear.
- [OR] Otto, F.; Reznikoff, M. G. Slow motion of gradient flows, accepted *J. Differential Equations*.
- [RS] Roeger, M.; Schaetzle, R. On a modified conjecture of DeGiorgi, *Mathematische Zeitschrift* **254** (2006), 675–714.
- [RKS] Rubinstein, J.; Sternberg, P.; Keller, J. B. Fast reaction, slow diffusion, and curve shortening. *SIAM J. Appl. Math.* **49** (1989), 116–133.
- [T] Tonegawa, Y. Phase field model with a variable chemical potential. *Proc. Roy. Soc. Edinburgh* **132A** (2002), 993–1019.
- [TW] Tonegawa, Y.; Westdickenberg, M. G. Higher multiplicity in the one dimensional Allen Cahn action functional, *Indiana University Mathematics Journal*, to appear.