

Heteroclinic orbits for a higher order phase transition problem

PETER W. BATES* and XIAOFENG REN†

Department of Mathematics, Brigham Young University, Provo, Utah 84602, USA

(Received 4 April 1995; received in revised form 16 July 1996)

A standard model for one-dimensional phase transitions is the second-order semilinear equation with bistable nonlinearity, where one seeks a solution which connects the two stable values. From an Ising-like model but which includes long-range interaction, one is led to consider the equation where the second-order operator is replaced by one of arbitrarily high order. Others have found the desired heteroclinic solutions for such equations, under the assumption that the higher-order terms have small coefficients, by employing singular perturbation methods for dynamical systems. Here, without making any assumption on the sizes of the coefficients, we obtain such heteroclinic solutions by using variational methods under the assumption that the nonlinearity arises from a potential having two wells of equal depths.

1 Introduction

Our purpose is to establish the existence of transition states which minimize a ‘truncated’ free energy, given by (1.10) below. As we shall show, this truncated energy arises naturally and generalizes the usual Allen–Cahn energy by allowing derivatives of arbitrarily high order.

In the theory of phase transitions from a microscopic viewpoint, one may consider a d -dimensional lattice having associated with each of its vertices either a $+1$ or a -1 (called spins). If $\nu(r)$ is the probability that the spin at site r is $+1$, then the expected value of the spin there is $u(r) = +1\nu + (-1)(1 - \nu) = 2\nu - 1 \in [-1, 1]$.

The Helmholtz free energy of a distribution u over the lattice is given by

$$F = H - TS, \quad (1.1)$$

where $H(u)$ is the internal interaction energy, T is the absolute temperature and $S(u)$ is the entropy. Following Khachatryan (see [1] and [2], for instance), the internal energy takes the form

$$H(u) = -\frac{1}{2} \sum_{r, r'} J(r-r') u(r) u(r'), \quad (1.2)$$

where the sum is over the lattice sites and $J(r-r') = J(r'-r) \geq 0$ gives the pairwise interaction energies. This is similar to the Ising model where nearest neighbour interactions give a contribution to the energy. Here, however, we allow $J(r) > 0$ for all r . In fact this

* Supported in part by NSF grant DMS-9305044. The hospitality of the Isaac Newton Institute, where this work was completed, is also appreciated.

† Current address: Department of Mathematics, Utah State University, Logan, UT 84322, USA.

formulation with continuous spins is usually justified by dividing a basic lattice up into relatively small blocks each of which has a large number of vertices; $u(r)$ is then the average spin at a vertex within the block centred at position r . Thus, our lattice could be a lattice of blocks with ‘block spins’ which take values in $[-1, 1]$, all values being possible in the limit as the number of particles in a block tends to infinity. The interaction, then, is between blocks and arises from a long-range (e.g. Kac) potential.

In the Bragg–Williams approximation (see [3] and [4]) the total entropy takes the form of the sum over the lattice of

$$-K[(1+u(r))\log(1+u(r))+(1-u(r))\log(1-u(r))],$$

where K is a positive constant. This has minima at -1 and 1 and naturally penalizes values of u outside the interval $[-1, 1]$.

In the analysis to follow we will consider a rearrangement of (1.1), changing H to

$$\frac{1}{4} \sum_{r,r'} J(r-r')(u(r)-u(r'))^2, \quad (1.3)$$

so that F is written

$$\begin{aligned} & \frac{1}{4} \sum_{r,r'} J(r-r')(u(r)-u(r'))^2 \\ & + \sum_r (TK[(1+u(r))\log(1+u(r))+(1-u(r))\log(1-u(r))] - cu^2(r)), \end{aligned} \quad (1.4)$$

where $c = (1/2)\sum_r J(r)$, which we assume is finite; the total free energy is unchanged. Just as in the Ising model, phase transitions are possible only for low temperatures. Note that when TK is large the second term in (1.4) is the sum over the lattice of a function which has a unique minimum and in this case, no phase transition is supported. At low temperature, however, this function has two wells of equal depth and so each of two homogeneous states minimizes the energy. We consider this case but replace the second term in (1.4) by

$$\frac{1}{4} \sum_r W(u(r)) \quad (1.5)$$

and assume that W is a double-well function which is everywhere defined. For computational ease later, we rescale u so that the minima of the double-well function occur at $u = -1, 1$. Taking the continuum limit gives

$$F(u) = \frac{1}{4} \left[\iint J(x-y)(u(x)-u(y))^2 dx dy + \int W(u(x)) dx \right]. \quad (1.6)$$

We assume J is non-negative, and normalize so that $\int J(x) dx = 1$.

The passage to the continuum limit to obtain (1.6) can be made rigorous using the theory of Gibbs measures, as noted in [5] (see also [6]). If one considers an interacting particle system with Glauber dynamics with Kac potential, letting the size of the system and the range of the potential tend to infinity in an appropriate way leads rigorously to the evolution equation

$$u_t = \tanh(\beta J * u) - u,$$

where β is proportional to inverse temperature and $*$ means convolution (see [7–9]). Then

one can verify that (1.6), with the correct choice of W , is a Lyapunov functional associated with this evolution equation. On the other hand, if one considers the gradient flow associated with (1.6) one obtains the evolution equation

$$u_t = J * u - u - W'(u).$$

For a general class of functions W , this equation has been studied in [10], where the existence and stability of travelling waves has been established. So far no particle dynamics has been rigorously associated with this latter evolution equation, however.

Returning to our current programme, following [11], we change variable in the first integral in (1.6) using $\eta = (x - y)/2$, $\xi = (x + y)/2$, then expand $u(x) = u(\xi + \eta)$ and $u(y) = u(\xi - \eta)$ about ξ to get the formal expression

$$2^d \iint J(2\eta) \left(\sum_{k=1}^{\infty} \sum_{|\alpha|=2k-1} \frac{1}{\alpha!} \partial^\alpha u(\xi) \eta^\alpha \right)^2 d\xi d\eta, \quad (1.7)$$

where both integrals are over R^d , $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d)$, $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_d$, $\partial^\alpha = \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \dots \partial_{x_d}^{\alpha_d}$, $\alpha! = \alpha_1! \alpha_2! \dots \alpha_d!$, and $\eta^\alpha = \eta_1^{\alpha_1} \eta_2^{\alpha_2} \dots \eta_d^{\alpha_d}$.

We truncate the infinite sum at $k = N$ and integrate with respect to η . Note that the truncation at $k = 1$ gives the usual energy for the Allen–Cahn equation of phase transition. For this paper we will consider only the case where $d = 1$ which will simplify the expression somewhat. The higher dimensional case is considered elsewhere (see [12]). This ‘truncated’ free energy is given by

$$F_N = \frac{1}{4} \iint \left[\sum_{i,j=1}^N c_{ij} \frac{d^{2i-1} u}{dx^{2i-1}}(\xi) \frac{d^{2j-1} u}{dx^{2j-1}}(\xi) + W(u(\xi)) \right] d\xi, \quad (1.8)$$

where

$$c_{ij} = 2 \int J(2\eta) \frac{\eta^{2i+2j-2}}{(2i-1)!(2j-1)!} d\eta. \quad (1.9)$$

Note that the matrix (c_{ij}) is positive definite, since

$$\begin{aligned} \sum_{i,j=1}^N c_{ij} \xi_i \xi_j &= 2 \int J(2\eta) \sum_{i,j=1}^N \frac{\eta^{2i-1} \xi_i}{(2i-1)!} \frac{\eta^{2j-1} \xi_j}{(2j-1)!} d\eta \\ &= 2 \int J(2\eta) \left(\sum_{i=1}^N \frac{\eta^{2i-1} \xi_i}{(2i-1)!} \right)^2 d\eta. \end{aligned}$$

Motivated by this, we will study a higher-order variational problem

$$I = \inf_{u \in \mathcal{L}} \int_R \left[\sum_{i,j=1}^N A_{ij} D^{2i-1} u D^{2j-1} u + W(u) \right] dx. \quad (1.10)$$

The function $W: R \rightarrow R$ is a ‘balanced’ double-well potential, i.e. the wells are of equal depths. An example is $W(u) = (u^2 - 1)^2$ but we allow more general functions. To be more precise, we assume only that W is continuous on R , C^2 near -1 and 1 , $W \geq 0$, -1 and 1 are the only global minima of W and that $W(-1) = W(1) = 0$, $W''(-1) > 0$ and $W''(1) > 0$. We also assume that

$$\liminf_{|y| \rightarrow \infty} W(y) > 0.$$

We assume the matrix $[A_{ij}]$ is positive definite and real. The function class \mathcal{U}_1 is defined as follows: We first define

$$= \left\{ u \in C(R) : D^{2i-1}u \in L^2(R), i = 1, \dots, N, \int_R W(u(x)) dx < \infty \right\}, \quad (1.11)$$

where $D^{2i-1}u$ denotes the $(2i-1)$ st distributional derivative of u . Because $Du \in L^2(R)$, we see for every $u \in \mathcal{U}_1$, there is a bound on the modulus of continuity. Also since $\int_R W(u) < \infty$, u has limits -1 or 1 at $-\infty$ at ∞ . We therefore decompose \mathcal{U}_1 into $\mathcal{U}_1, \mathcal{U}_2, \mathcal{U}_3$ and \mathcal{U}_4 , where

$$\left. \begin{aligned} \mathcal{U}_1 &= \{u \in \mathcal{U}_1 : \lim_{x \rightarrow -\infty} u(x) = -1, \lim_{x \rightarrow \infty} u(x) = 1\}, \\ \mathcal{U}_2 &= \{u \in \mathcal{U}_1 : \lim_{x \rightarrow -\infty} u(x) = 1, \lim_{x \rightarrow \infty} u(x) = -1\}, \\ \mathcal{U}_3 &= \{u \in \mathcal{U}_1 : \lim_{x \rightarrow -\infty} u(x) = 1, \lim_{x \rightarrow \infty} u(x) = 1\}, \\ \mathcal{U}_4 &= \{u \in \mathcal{U}_1 : \lim_{x \rightarrow -\infty} u(x) = -1, \lim_{x \rightarrow \infty} u(x) = -1\}. \end{aligned} \right\} \quad (1.12)$$

Clearly, $\mathcal{U}_1 = \mathcal{U}_1 \cup \mathcal{U}_2 \cup \mathcal{U}_3 \cup \mathcal{U}_4$ and $\mathcal{U}_i \cap \mathcal{U}_j = \emptyset$ if $i \neq j$. We look for a minimizer in \mathcal{U}_1 , since \mathcal{U}_2 is a similar subclass and 1 (-1 , respectively) is the minimizer in \mathcal{U}_3 (\mathcal{U}_4 , respectively). We denote

$$\int_R \left[\sum_{i,j=1}^N A_{ij} D^{2i-1}u D^{2j-1}u + W(u) \right] dx$$

by $E(u)$, the energy of u .

The Euler–Lagrange equation for the minimizers of (1.10) is

$$\left. \begin{aligned} \sum_{i,j=1}^N 2A_{ij} D^{2i+2j-2}u - W'(u) &= 0 \quad \text{in } R, \\ \lim_{x \rightarrow -\infty} u(x) &= -1, \quad \lim_{x \rightarrow \infty} u(x) = 1. \end{aligned} \right\} \quad (1.13)$$

If $N = 1$, one recovers the steady state Allen–Cahn equation

$$\left. \begin{aligned} u'' - W'(u) &= 0 \quad \text{in } R, \\ \lim_{x \rightarrow -\infty} u(x) &= -1, \quad \lim_{x \rightarrow \infty} u(x) = 1. \end{aligned} \right\} \quad (1.14)$$

If one removes the condition that W is balanced, one can look for a travelling wave solution, that is, a function u and a number c such that

$$\left. \begin{aligned} u'' + cu' - W'(u) &= 0 \quad \text{in } R, \\ \lim_{x \rightarrow -\infty} u(x) &= -1, \quad \lim_{x \rightarrow \infty} u(x) = 1. \end{aligned} \right\} \quad (1.15)$$

Fife & McLeod [13] proved that such a travelling wave is unique up to translation and the family of translates is globally stable with respect to the associated parabolic equation. Note that if W is balanced, multiplying the equation by u' and integrating shows that $c = 0$.

Gardner & Jones [14] studied the case where $N = 2$ and sought travelling wave solutions when W is not necessarily balanced. They treated the resulting sixth-order equation as a perturbation of the second-order equation, and proved the existence of a travelling wave solution if the coefficients of the higher-order derivatives are sufficiently small. This result was generalized later by Bates *et al.* [15] to the case of arbitrary N . However, the smallness condition on the coefficients of the higher-order derivatives cannot be removed in their approach, since it employs geometric singular perturbation theory for dynamical systems.

The main purpose of this paper is to show that problem (1.10) possesses an energy minimizer. Therefore, equation (1.13) can be solved without the smallness condition on the coefficients of the higher-order derivatives. Unfortunately, our variational approach cannot handle the case where W is not balanced. We will prove:

Theorem 1.1 *Every minimizing sequence in \mathcal{I}_1 gives rise to an energy minimizer in the same class after a proper translation. If in addition W is C^1 on \mathbb{R} , this minimizer is a heteroclinic solution of the corresponding Euler–Lagrange equation.*

2 Proof of the theorem

Let $\{u_n\}$ be a minimizing sequence of E in \mathcal{I}_1 . We first show that $\int_{\mathbb{R}} W(u_n)$ does not vanish.

Lemma 2.1 *There exist $\lambda, 0 < \lambda < \infty$ such that*

$$\int_{\mathbb{R}} W(u_n) \geq \lambda$$

for all n .

Proof Since $0 \leq I < \infty$, $\limsup_{n \rightarrow \infty} \int_{\mathbb{R}} W(u_n) < \infty$. Also we may assume that the Du_n 's are bounded in L^2 uniformly, so the moduli of continuity of the u_n 's are bounded uniformly by a constant. Since $u_n \in \mathcal{I}_1$, $u_n = 0$ at some point x_n . Therefore, for every $\epsilon > 0$ there is $\delta > 0$, independent of n , such that $|x - x_n| < \delta$ implies $|u_n(x)| < \epsilon$. The continuity of W implies that for every $\eta > 0$ there exists $\epsilon > 0$ such that if $|y| < \epsilon$, then $W(y) \in (W(0) - \eta, W(0) + \eta)$. For this ϵ choose δ as above. Then for all x such that $|x - x_n| < \delta$ we have $|u_n(x)| < \epsilon$, and in turn we deduce $W(u_n(x)) \in (W(0) - \eta, W(0) + \eta)$. Therefore, $\int_{\mathbb{R}} W(u_n) \geq \int_{x_n - \delta}^{x_n + \delta} W(u_n) dx \geq 2\delta(W(0) - \eta) > 0$ if we choose η small. \square

We need a criterion to give a necessary condition for a possible phase transition to occur. This criterion plays an essential role in our analysis of $\{u_n\}$.

Lemma 2.2 *For every energy minimizing sequence $\{u_n\}$, $\epsilon_1 > 0$, and $L > 0$ there exists $\delta > 0$ such that if for any n*

$$\int_a^b W(u_n) dx < \delta \quad \text{and} \quad b - a > L,$$

then $u_n(x)$ is either in the ϵ_1 -neighbourhood of -1 for all $x \in (a, b)$ or in the ϵ_1 -neighbourhood of 1 for all $x \in (a, b)$.

Proof Suppose the contrary. There exist $\epsilon_1 > 0$, $L > 0$, $\{\delta_k\}$, $\delta_k \rightarrow 0$ as $k \rightarrow \infty$, $\{u_{n_k}\}$ from the minimizing sequence, $[a_k, b_k]$, $b_k - a_k > L$, and $\{x_k\}$, with $x_k \in [a_k, b_k]$, such that

$$\int_{a_k}^{b_k} W(u_{n_k}) dx \leq \delta_k \quad \text{as } k \rightarrow \infty$$

and $|u_{n_k}(x_k) + 1| \geq \epsilon_1$, $|u_{n_k}(x_k) - 1| \geq \epsilon_1$.

Because the Du_{n_k} 's are bounded uniformly in $L^2(R)$, we have a bound independent of k on the moduli of continuity of the u_{n_k} 's. Therefore, we can find η independent of k such that $|u_{n_k}(x) - 1| \geq \epsilon_1/2$ and $|u_{n_k}(x) + 1| \geq \epsilon_1/2$ for all $x \in (x_k - \eta, x_k + \eta)$, so we deduce

$$\begin{aligned} \int_{a_k}^{b_k} W(u_{n_k}) dx &\geq \int_{(x_k - \eta, x_k + \eta) \cap [a_k, b_k]} W(u_{n_k}) dx \\ &\geq \eta \min_{x \in (x_k - \eta, x_k + \eta) \cap [a_k, b_k]} W(u_{n_k}(x)) \\ &\geq \eta \min_{y \in R \setminus (-1 - \epsilon_1/2, -1 + \epsilon_1/2) \cup (1 - \epsilon_1/2, 1 + \epsilon_1/2)} W(y) > 0. \end{aligned}$$

This contradicts

$$\int_{a_k}^{b_k} W(u_{n_k}) dx \leq \delta_k \rightarrow 0$$

as $k \rightarrow \infty$. \square

We now adapt the concentration-compactness principle of Lions [16]. The idea is to show that the minimizing nature of the sequence $\{u_n\}$ forces these functions to have exactly one transition and this transition to have almost minimal energy and therefore to be localized. By translating these functions so that the transitions all occur in a compact interval, we obtain a convergent minimizing sequence. In Lemma 2.1 we may take

$$\lambda = \liminf_{n \rightarrow \infty} \int_R W(u_n) dx$$

and then take $\{u_n\}$ to be a subsequence along which this is achieved. Consider $\{W(u_n)\}$ as a non-negative bounded sequence in $L^1(R)$. Define

$$Q_n(t) = \sup_{y \in R} \int_{y+B(t)} W(u_n(x)) dx \quad (2.1)$$

for $t \geq 0$ where $B(t)$ denotes the ball centred at the origin with radius t and $y + B(t)$ denotes the translation of $B(t)$ by y . $\{Q_n\}$ is a sequence of non-negative, non-decreasing, uniformly bounded functions on $[0, \infty)$ and

$$\lim_{n \rightarrow \infty} \lim_{t \rightarrow \infty} Q_n(t) = \lambda. \quad (2.2)$$

By a classical result about monotone functions, there exist a subsequence of $\{u_n\}$, again denoted by $\{u_n\}$, and a non-negative, non-decreasing function $Q(t)$ such that

$$\lim_{n \rightarrow \infty} Q_n(t) = Q(t)$$

for all $t \geq 0$. Let $\alpha = \lim_{t \rightarrow \infty} Q(t)$. Then $\alpha \in [0, \lambda]$ by (2.2) since $Q_n(t)$ is non-decreasing. Furthermore, by (2.2) and definition (2.1) we may make the following three statements:

I: If $\alpha = \lambda$, there exist a subsequence of $\{u_n\}$, again denoted by $\{u_n\}$, and $y_n \in \mathbb{R}$ such that the sequence $\{W(u_n)\}$ has the property that for every $\epsilon > 0$ there exist $r > 0$ and a positive integer n_1 so that

$$\int_{y_n+B(r)} W(u_n(x)) dx > \lambda - \epsilon$$

for $n \geq n_1$.

II: If $\alpha = 0$, then for all $r > 0$,

$$\limsup_{n \rightarrow \infty} \int_{y \in \mathbb{R}} \int_{y+B(r)} W(u_n(x)) dx = 0.$$

III: If $\alpha \in (0, \lambda)$, then there exists a subsequence of $\{u_n\}$ again denoted by $\{u_n\}$ such that for every $\epsilon > 0$ there exist $r > 0$, $r_n \rightarrow \infty$ as $n \rightarrow \infty$, and a positive integer n_1 such that for $n \geq n_1$

$$\alpha - \epsilon < Q_n(r), \quad Q_n(r_n) < \alpha + \epsilon.$$

These three statements can be found in the proof of Lemma I.1 [16]. We will prove the main theorem by showing that only the case $\alpha = \lambda$ can occur when $\{u_n\}$ is an energy minimizing sequence. We will then deduce that certain translates of the u_n 's approach an energy minimizer in \mathbb{R}^1 .

Lemma 2.3 *The case $\alpha = 0$ cannot occur.*

This follows from the proof of Lemma 2.1, taking $y = x_n$ and $r = \delta$ (see the last line of that proof).

Lemma 2.4 *The case $\alpha \in (0, \lambda)$ cannot occur.*

We postpone the proof of this lemma to the next section. We are now left with only the case $\alpha = \lambda$. The next lemma proves the main theorem.

Lemma 2.5 *The case $\alpha = \lambda$ implies that there exists $u \in \mathbb{R}^1$ such that*

$$E(u) = \inf_{v \in \mathbb{R}^1} E(v).$$

Proof Statement I above implies that there exist $y_n \in \mathbb{R}$ such that for every $\epsilon > 0$ there exist $r \in \mathbb{R}$ and a positive integer n_1 such that when $n \geq n_1$

$$\int_{y_n+B(r)} W(u_n) > \lambda - \epsilon. \tag{2.3}$$

Take intervals $E_1 = (-\infty, -r)$ and $E_2 = (r, \infty)$. For large n we deduce from (2.3) that

$$\int_{y_n+E_1} W(u_n) dx < 2\epsilon \quad \text{and} \quad \int_{y_n+E_2} W(u_n) dx < 2\epsilon.$$

Lemma 2.2 implies that u_n stays close to -1 on $y_n + E_1$ and close to 1 on $y_n + E_2$ uniformly in n . If we consider the translations $u_n(\cdot - y_n)$ of u_n , we see that the $u_n(\cdot - y_n)$'s stay close to -1 on E_1 and close to 1 on E_2 uniformly in n . The L^2 bound on $Du_n(\cdot - y_n)$ implies that the sequence $\{u_n(\cdot - y_n)\}$ is equicontinuous. Since $u_n(x_n - y_n) = 0$ for some x_n in $B(r)$, the L^2 bound on $Du_n(\cdot - y_n)$ also gives a uniform upper bound for $|u_n(\cdot - y_n)|$ on every compact subset of R . Using Ascoli–Arzela's theorem on equicontinuous families of functions, we can pass to the limit along a subsequence of $\{u_n(\cdot - y_n)\}$ and conclude that there is a function u such that

$$u_n(\cdot - y_n) \rightarrow u \quad \text{in } C(K) \quad \text{for every compact set } K \subset R. \quad (2.4)$$

The upper bound on $D^{2i-1}u_n(\cdot - y_n)$, $i = 1, 2, \dots, N$, implies that after passing to a further subsequence,

$$D^{2i-1}u_n(\cdot - y_n) \rightarrow D^{2i-1}u \quad \text{weakly in } L^2(R) \quad \text{for } i = 1, 2, \dots, N. \quad (2.5)$$

Since $[A_{ij}]$ is positive definite, $\sum_{i,j=1}^N A_{ij} \xi_i \xi_j$ is convex in $\xi_1, \xi_2, \dots, \xi_N$. Hence

$$\int_R \sum_{i,j=1}^N A_{ij} D^{2i-1}u D^{2j-1}u \, dx$$

is weakly lower semi-continuous in \quad , so (2.5) implies

$$\begin{aligned} \liminf_{n \rightarrow \infty} \int_R \sum_{i,j=1}^N A_{ij} D^{2i-1}u_n(x - y_n) D^{2j-1}u_n(x - y_n) \, dx \\ \geq \int_R \sum_{i,j=1}^N A_{ij} D^{2i-1}u(x) D^{2j-1}u(x) \, dx, \end{aligned} \quad (2.6)$$

and Fatou's lemma together with (2.4) implies

$$\lambda = \liminf_{n \rightarrow \infty} \int_R W(u_n(x - y_n)) \, dx \geq \int_R W(u(x)) \, dx. \quad (2.7)$$

Therefore $u \in \quad$. The uniform estimate on the intervals E_1 and E_2 implies $u \in \quad_1$. So u is an energy minimizer in \quad_1 . \square

Once we have an energy minimizer u , we know that u solves the corresponding Euler–Lagrange equation and the proof of Theorem 1.1 is complete. It would be natural to expect that the minimizer is monotone but we do not make such a claim. Note that there is no maximum principle or positive operator involved in (1.13) but the fact that we have a minimizer means the transition is made as efficiently as possible.

Remark 2.6 *The convergence in (2.4) is actually*

$$u_n(\cdot - y_n) \rightarrow u \quad \text{in } C(R).$$

The global convergence of $u_n(\cdot - y_n)$ to u follows from the convergence on compact sets and the fact that $u_n(\cdot - y_n)$ stays uniformly close to -1 or 1 outside large compact sets.

Corollary 2.7 *The convergence in (2.5) is strong convergence.*

Proof From (2.6) and (2.7) we see

$$\lim_{n \rightarrow \infty} E(u_n) \geq E(u).$$

Since $u \in \mathcal{U}_1$ and $\{u_n\}$ is an energy minimizing sequence, we have

$$\lim_{n \rightarrow \infty} E(u_n) = E(u).$$

This together with (2.6), (2.7) and the positivity of $[A_{ij}]$, gives the result in their usual way. \square

3 Proof of Lemma 2.4

It follows from statement III of §2 that for every small ϵ , we have $r > 0$ and a sequence $r_n \rightarrow \infty$, such that $Q_n(r), Q_n(r_n) \in (\alpha - \epsilon, \alpha + \epsilon)$ for large n . Therefore, we can find y_n such that

$$\int_{y_n + B(r)} W(u_n) dx \in (\alpha - \epsilon, \alpha + \epsilon) \quad (3.1)$$

and

$$\int_{R \setminus (y_n + B(r_n))} W(u_n) dx \in (\lambda - \alpha - \epsilon, \lambda - \alpha + \epsilon) \quad (3.2)$$

for large n .

Let us denote the two intervals of $(y_n + B(r_n)) \setminus (y_n + B(r))$ by $y_n + F_{1n}$ and $y_n + F_{2n}$ with the first lying to the left of the second. Clearly

$$|F_{1n}| = |F_{2n}| \rightarrow \infty$$

as $n \rightarrow \infty$ where $|S|$ denotes the length of an interval of S . Also denote $(-\infty, -r)$ by G_1 and (r, ∞) by G_2 . Clearly $F_{1n} \subset G_1$ and $F_{2n} \subset G_2$. For n large enough

$$\int_{y_n + F_{1n}} W(u_n) dx < 2\epsilon, \quad \text{and} \quad \int_{y_n + F_{2n}} W(u_n) dx < 2\epsilon. \quad (3.3)$$

Choosing ϵ small enough, for large n we see from Lemma 2.2 that the u_n 's stay close to -1 or 1 on $y_n + F_{1n}$ and $y_n + F_{2n}$. In view of Lemma 2.2, we will prove Lemma 2.4 by considering three cases.

Case 1 Suppose, after choosing a proper subsequence if necessary, the u_n 's stay close to -1 on $y_n + F_{2n}$.

We will show in this case that the u_n 's already have large energies on $y_n + G_2$ so they cannot constitute an energy minimizing sequence. Let ξ be a non-negative smooth cut-off function such that

$$\left. \begin{aligned} \xi(x) &= 0 & \text{if } x < 0, \\ \xi(x) &= 1 & \text{if } x > 1, \end{aligned} \right\} \quad (3.4)$$

$0 \leq \xi(x) \leq 1$, and $|D^k \xi(x)| \leq C_1$ for $k = 1, 2, \dots, 2N-1$. Set

$$v_n(x) = (u_n(x) + 1) \xi\left(\frac{x - (y_n + (r + r_n)/2)}{d}\right) - 1,$$

where d is a large number to be specified later. Note that $v_n \in \mathcal{U}_1$.

We fix d and take n large enough so that $d < (r_n - r)/2$. Then $v_n(x) = -1$ for $x \notin y_n + G_2$. We will compare $\int_{y_n + G_2} W(u_n) dx$ with $\int_R W(v_n) dx$. Since for $x \in y_n + G_2$, u_n and v_n only differ on $y_n + F_{2n}$, and on that interval v_n is closer to -1 than u_n , we have

$$\int_R W(v_n) dx \leq \int_{y_n + G_2} W(u_n) dx. \quad (3.5)$$

We also compare

$$\int_{y_n + G_2} A_{ij} D^{2i-1} u_n D^{2j-1} u_n dx \quad \text{with} \quad \int_R A_{ij} D^{2i-1} v_n D^{2j-1} v_n dx.$$

Indeed, we need only to compare

$$\int_{y_n + F_{2n}} A_{ij} D^{2i-1} u_n D^{2j-1} u_n dx \quad \text{with} \quad \int_{y_n + F_{2n}} A_{ij} D^{2i-1} v_n D^{2j-1} v_n dx$$

since the two integrands agree on $(y_n + G_2) \setminus (y_n + F_{2n})$ and $D^{2i-1} v_n$ vanishes on the interval to the left of $y_n + F_{2n}$. Consider

$$\begin{aligned} A_{ij} D^{2i-1} v_n D^{2j-1} v_n &= A_{ij} D^{2i-1} \left[(u_n(x) + 1) \xi \left(\frac{x - (y_n + (r + r_n)/2)}{d} \right) \right] \\ &\quad \times D^{2j-1} \left[(u_n(x) + 1) \xi \left(\frac{x - (y_n + (r + r_n)/2)}{d} \right) \right] \\ &= A_{ij} \left[\sum_{k=1}^{2i-1} \frac{C_{2i-1}^k}{d^k} D^{(2i-1)-k} (u_n(x) + 1) D^k \xi \left(\frac{x - (y_n + (r + r_n)/2)}{d} \right) \right] \\ &\quad \times \left[\sum_{l=1}^{2j-1} \frac{C_{2j-1}^l}{d^l} D^{(2j-1)-l} (u_n(x) + 1) D^l \xi \left(\frac{x - (y_n + (r + r_n)/2)}{d} \right) \right]. \end{aligned}$$

The first term from the product is

$$A_{ij} \xi^2 \left(\frac{x - (y_n + (r + r_n)/2)}{d} \right) D^{2i-1} u_n D^{2j-1} u_n \quad (3.6)$$

and the other terms assume the form

$$\begin{aligned} A_{ij} \frac{C_{2i-1}^k C_{2j-1}^l}{d^{k+l}} D^{(2i-1)-k} (u_n(x) + 1) D^{(2j-1)-l} (u_n(x) + 1) \\ \times D^k \xi \left(\frac{x - (y_n + (r + r_n)/2)}{d} \right) D^l \xi \left(\frac{x - (y_n + (r + r_n)/2)}{d} \right), \end{aligned}$$

where at least one of k and l is not equal to 0. These are bounded by

$$\frac{C}{d^{k+l}} |D^{(2i-1)-k} (u_n(x) + 1) D^{(2j-1)-l} (u_n(x) + 1)|.$$

We therefore need

Lemma 3.1 *There exists $C > 0$ and a positive integer n_1 such that for $n \geq n_1$*

$$\int_{y_n + F_{2n}} \frac{1}{d^{k+l}} D^{(2i-1)-k}(u_n(x)+1) D^{(2j-1)-l}(u_n(x)+1) dx \leq \frac{C}{d}$$

for all $i, j \in \{1, 2, \dots, N\}$, $k \in \{0, 1, 2, \dots, 2i-1\}$ and $l \in \{0, 1, 2, \dots, 2j-1\}$ with $k+l \geq 1$.

Proof In view of Lemma 2.2, we know that $u_n(x)+1$ is close to 0 uniformly for $x \in y_n + F_{2n}$. According to our assumptions that $W''(-1) > 0$ and $W''(1) > 0$, we find

$$(u_n(x)+1)^2 \leq CW(u_n(x))$$

for $x \in y_n + F_{2n}$ where C is independent of large n . We then obtain

$$\int_{y_n + F_{2n}} (u_n(x)+1)^2 \leq C,$$

where C is again independent of large n . We also know that

$$\int_R [D^{2i-1} u_n(x)]^2 dx \leq C, \quad i = 1, 2, \dots, N. \quad (3.7)$$

We need an estimate on

$$\int_R [D^{2i} u_n(x)]^2 dx, \quad i = 1, 2, \dots, N-1. \quad (3.8)$$

We can achieve such an estimate by directly appealing to the following interpolation inequality for intermediate derivatives.

Lemma 3.2 *Let $-\infty \leq a < b \leq \infty$, let $1 \leq p \leq \infty$, and let $0 < \epsilon_0 < \infty$. There exists a finite constant $K = K(\epsilon_0, p, b-a)$, depending continuously on $b-a$ for $0 < b-a \leq \infty$, such that for every ϵ_1 satisfying $0 < \epsilon_1 \leq \epsilon_0$, and for every function f twice continuously differentiable on the open interval (a, b)*

$$\int_a^b |f'(x)|^p dx \leq K \epsilon_1 \int_a^b |f''(x)|^p dx + K \epsilon_1^{-1} \int_a^b |f(x)|^p dx.$$

Moreover, if $b-a = \infty$, then $K = K(p)$ can be found so that the inequality holds for every positive number ϵ_1 .

The proof of this inequality can be found in [17, Lemma 4.10, p. 70].

Applying this lemma, we obtain that there is a universal constant K such that for every $f \in C^2(R)$

$$\int_R (f'(x))^2 dx \leq K \int_R [(f''(x))^2 + (f(x))^2] dx.$$

With the help of this estimate the desired estimate on (3.8) follows from (3.7). We now have a uniform bound on

$$\int_{y_n + F_{2n}} [D^k(u_n(x)+1)]^2 dx, \quad k = 0, 1, \dots, 2N-1.$$

Lemma 3.1 then follows from this bound by Hölder's inequality. \square

We sum over i and j and find

$$\int_{y_n+F_{2n}} \sum_{i,j=1}^N A_{ij} D^{2i-1} v_n D^{2j-1} v_n \, dx \leq \int_{y_n+F_{2n}} \xi^2 \sum_{i,j=1}^N A_{ij} D^{2i-1} u_n D^{2j-1} u_n \, dx + O\left(\frac{1}{d}\right).$$

Therefore

$$\int_R \sum_{i,j=1}^N A_{ij} D^{2i-1} v_n D^{2j-1} v_n \, dx \leq \int_{y_n+G_2} \sum_{i,j=1}^N A_{ij} D^{2i-1} u_n D^{2j-1} u_n \, dx + O\left(\frac{1}{d}\right).$$

Together with (3.5) and the fact that $v_n \in \mathcal{V}_1$, we deduce

$$\begin{aligned} I &\leq \int_R \left[\sum_{i,j=1}^N A_{ij} D^{2i-1} v_n D^{2j-1} v_n + W(v_n) \right] dx \\ &\leq \int_{y_n+G_2} \left[\sum_{i,j=1}^N A_{ij} D^{2i-1} u_n D^{2j-1} u_n + W(u_n) \right] dx + O\left(\frac{1}{d}\right). \end{aligned}$$

If we choose d large enough, we obtain

$$\begin{aligned} E(u_n) &\geq \int_{R \setminus (y_n+G_2)} W(u_n) \, dx + \int_{y_n+G_2} \left[\sum_{i,j=1}^N A_{ij} D^{2i-1} u_n D^{2j-1} u_n + W(u_n) \right] dx \\ &\geq \alpha + I - 2\epsilon, \end{aligned}$$

which implies that $\{u_n\}$ cannot be an energy minimizing sequence. This contradiction rules out the possibility that the u_n 's stay close to -1 on $y_n + F_{2n}$.

Case 2 Suppose, after choosing a subsequence if necessary, the u_n 's stay close to 1 on $y_n + F_{1n}$.

This possibility can be ruled out by the same argument as in the previous case. Instead of considering the interval $y_n + F_{2n}$, one needs to consider the interval $y_n + F_{1n}$.

Case 3 Suppose, after choosing a subsequence if necessary, the u_n 's stay close to -1 on $y_n + F_{1n}$ and close to 1 on $y_n + F_{2n}$.

To rule out this possibility, we will truncate u_n outside $y_n + B(r_n)$ and replace it by w_n which joins -1 to 1 smoothly. It turns out they have approximately equal energies, being at least close to I , but the two tails of u_n have extra energies bounded away from 0. Therefore such u_n 's cannot constitute an energy minimizing sequence. We now carry out the details. Set

$$\begin{aligned} w_n(x) &= \left\{ \left[(u_n(x) + 1) \xi \left(\frac{x - (y_n - (r + r_n)/2)}{d} \right) - 1 \right] - 1 \right\} \\ &\quad \times \xi \left(\frac{-x + (y_n + (r + r_n)/2)}{d} \right) + 1, \quad (3.9) \end{aligned}$$

where $d < (r_n - r)/2$ is a large number to be specified later.

We first compare $\int_{y_n+B(r_n)} W(u_n) \, dx$ with $\int_R W(w_n) \, dx$. Observe that u_n and w_n only differ on $(y_n + G_1) \cup (y_n + G_2)$ for large n and $W(w_n) = 0$ outside $y_n + B(r_n)$. Also on $y_n + F_{1n}$,

where u_n is close to -1 , w_n is closer to -1 , and on $y_n + F_{2n}$, where u_n is close to 1 , w_n is closer to 1 . Therefore, we deduce

$$\int_R W(w_n) dx \leq \int_{y_n+B(r_n)} W(u_n) dx. \quad (3.10)$$

We then compare

$$\int_{y_n+B(r_n)} A_{ij} D^{2i-1} u_n D^{2j-1} u_n dx \quad \text{with} \quad \int_R A_{ij} D^{2i-1} w_n D^{2j-1} w_n dx.$$

We consider the integrals on three different intervals, $y_n + F_{1n}$, $y_n + F_{2n}$ and $y_n + B(r)$. Since u_n and w_n agree on $y_n + B(r)$ for large n , we only consider $y_n + F_{1n}$ and $y_n + F_{2n}$.

On $y_n + F_{1n}$,

$$\xi\left(\frac{-x + (y_n + (r + r_n)/2)}{d}\right) = 1$$

for large n . Therefore

$$w_n(x) = [u_n(x) + 1] \xi\left(\frac{x - (y_n - (r + r_n)/2)}{d}\right) - 1.$$

On $y_n + F_{2n}$,

$$\xi\left(\frac{x - (y_n - (r + r_n)/2)}{d}\right) = 1$$

for large n . Therefore

$$w_n(x) = [u_n(x) - 1] \xi\left(\frac{-x + (y_n + (r + r_n)/2)}{d}\right) + 1.$$

Using the same argument as that in case 1, we find

$$\int_R \sum_{i,j=1}^N A_{ij} D^{2i-1} w_n D^{2j-1} w_n dx \leq \int_{y_n+B(r_n)} \sum_{i,j=1}^N A_{ij} D^{2i-1} u_n D^{2j-1} u_n dx + O\left(\frac{1}{d}\right).$$

Together with (3.10) and the fact that $w_n \in \mathcal{I}_1$ we deduce

$$\begin{aligned} I &\leq \int_R \left[\sum_{i,j=1}^N A_{ij} D^{2i-1} w_n D^{2j-1} w_n + W(w_n) \right] dx \\ &\leq \int_{y_n+B(r_n)} \sum_{i,j=1}^N [A_{ij} D^{2i-1} u_n D^{2j-1} u_n + W(u_n)] dx + O\left(\frac{1}{d}\right). \end{aligned}$$

If we choose d large enough, we obtain

$$\begin{aligned} E(u_n) &\geq \int_{R \setminus (y_n+B(r_n))} W(u_n) dx \\ &\quad + \int_{y_n+B(r_n)} \left[\sum_{i,j=1}^N A_{ij} D^{2i-1} u_n D^{2j-1} u_n + W(u_n) \right] dx \geq \lambda - \alpha + I - 2\epsilon, \end{aligned}$$

which implies that $\{u_n\}$ cannot be an energy minimizing sequence if we choose ϵ small. This rules out the third possibility and proves Lemma 2.4.

4 Conclusion

We have formulated a model for the equilibrium theory of phase transitions based on the Helmholtz free energy for a lattice system having possibly long-range rather than Ising nearest-neighbour interaction. The continuized version of the interaction energy was expanded and truncated at an arbitrarily high level. The term in the free energy that is associated with the entropy of mixing was taken to be based on a general smooth function which retained the property that two distinct homogeneous states minimized the free energy.

Our task was to seek transition states which minimized this free energy functional. We found such minimizers through the use of the concentration-compactness principle within the calculus of variations.

Note that the resulting Euler–Lagrange equation has the form of the usual Allen–Cahn equation but with the Laplace operator replaced by an elliptic operator of arbitrarily high order. While previous studies of the second-order Allen–Cahn equation have usually relied on the maximum principle or phase plane analysis, these are not available for the higher-order version considered here. However, in the variational approach the presence of the higher derivatives was actually helpful in obtaining a priori bounds needed to find a minimizer.

Interesting questions remain such as those surrounding the relationship between the discrete lattice model and the continuized and ‘truncated’ model. Also, the effect of our replacing the usual (Bragg–Williams) entropy density function, which is undefined outside a certain interval, by one defined everywhere should be explored. We expect, but have not shown, that the minimizer is a monotone function. In addition, the question of existence of travelling wave solutions when the potential has wells of unequal depths remains open. In the absence of rigorous results, computational experiments would be a welcome addition to our understanding of these issues.

References

- [1] KHACHATURYAN, A. G. (1968) Microscopic theory of diffusion in crystalline solid solutions and the time evolution of the diffuse scattering of X rays and thermal neutrons. *Soviet Phys. Solid State*, **9**, 2040–2046.
- [2] KHACHATURYAN, A. G. (1983) *Theory of Structural Transformations in Solids*. Wiley-Interscience.
- [3] BRAGG, W. L. & WILLIAMS, E. J. (1934) Effect of thermal agitation on atomic arrangement in alloys. *Proc. Roy. Soc. London Ser. A*, **145**, 699–730.
- [4] HAASEN, P. (1978) *Physical Metallurgy*. Cambridge University Press.
- [5] BELLETINI, G. & PRESUTTI, E. In preparation.
- [6] LEBOWITZ, J. L. & PENROSE, O. (1966) Rigorous treatment of the van der Waals–Maxwell theory of liquid-vapor transition. *J. Math. Phys.* **7**, 98–113.
- [7] DEMASI, A., ORLANDI, E., PRESUTTI, E. & TRIOLO, L. (1994) Glauber evolution with Kac potentials: I. Mesoscopic and macroscopic limits, interface dynamics. *Nonlinearity*, **7**, 633–696.
- [8] KATSOUKAKIS, M. & SOUGANIDIS, P. E. (1994) Interacting particle systems and generalized front propagation. *Arch. Rat. Mech. Anal.*, **127**, 133–157.

- [9] SOUGANIDIS, P. E. (1997) Interface dynamics in phase transitions. *Proceedings of the 1994 ICM*. To appear.
- [10] BATES, P. W., FIFE, P. C., REN, X. & WANG, X. (1997) Traveling waves in a convolution model for phase transitions. *Arch. Rat. Mech. Anal.* To appear.
- [11] BATES, P. W., FIFE, P. C., GARDNER, R. & JONES, C. K. R. T. (1977) Phase field models for hypercooled solidification. *Physica*. To appear.
- [12] BATES, P. W. & REN, X. (1996) Transition layer solutions of a higher order equation in an infinite tube. *Comm. Partial Differential Equations*, **21**, 195–220.
- [13] FIFE, P. & MCLEOD, J. B. (1977) The approach of solutions of nonlinear diffusion equations to traveling front solutions. *Arch. Rat. Mech. Anal.*, **65**, 335–361.
- [14] GARDNER, R. & JONES, C. K. R. T. (1989) Traveling waves of a perturbed diffusion equation arising in a phase field model. *Indiana Math. J.*, **38** (4), 1197–1222.
- [15] BATES, P. W., FIFE, P. C., GARDNER, R. & JONES, C. K. R. T. (1977) The existence of traveling wave solutions of a generalized phase-field model. *SIAM J. Math. Anal.* **28** (1), 60–93.
- [16] LIONS, P. L. (1984) The concentration-compactness principle in the calculus of variations: The locally compact case. Part I and Part II. *Ann. Inset. H. Poincare Sect. A (N.S.)* **1**, 109–145, 223–283.
- [17] ADAMS, R. A. (1975) *Sobolev Spaces*, Academic Press.