Relaxation and Regularization of Nonconvex Variational Problems

Robert V. Kohn[1]
Courant Institute of Mathematical Sciences
251 Mercer Street
New York, NY 10012, USA

and

Stefan Müller[2]
Institut für Angewandte Mathematik
Universität Bonn
D-5300 Bonn 1, Germany

In honor of Luigi Amerio's 80th birthday

Published as Rend. Sem. Mat. Fis. Univ.
Milano vol. 62 (1992) pp. 89-113

Abstract

We are interested in variational problems of the form $\min \int W(\nabla u)\,dx$, with W nonconvex. The theory of relaxation allows one to calculate the minimum value, but it does not determine a well-defined "solution" since minimizing sequences are far from unique. A natural idea for determining a solution is regularization, i.e. the addition of a higher order term such as $\epsilon|\nabla\nabla u|^2$. But what is the behavior of the regularized solution in the limit as $\epsilon \to 0$? Little is known in general. Our recent work [19, 20, 21] discusses a particular problem of this type, namely $\min_{u_y=\pm 1} \int\int u_x^2 + \epsilon|u_{yy}|\,dxdy$ with various boundary conditions. The present paper gives an expository overview of our methods and results.

[1]Partially supported by NSF grant DMS-9102829, AFOSR grant 90-0090, and ARO contract DAAL03-92-G-0011.
[2]Partially supported by NSF grant DMS-9002679 and by SFB 256 at the University of Bonn.

1 Introduction

This paper provides an expository review of our recent work [19, 20, 21]. The general setting is as follows. We are concerned with nonconvex (and non-quasiconvex) variational problems of the form

$$\min_{u=u_0 \text{at} \partial\Omega} \int_\Omega W(\nabla u)\, dx \ . \tag{1}$$

Such problems arise naturally in a variety of contexts, including structural optimization and the modelling of martensitic phase transformation [3, 4, 17, 22].

Since W is assumed to be neither convex nor quasiconvex, the minimum in (1) may not be achieved [9]. It is natural to be suspicious of a problem that "has no solution." The simplest way to restore existence is by regularization, i.e. by adding a term involving higher derivatives:

$$\min_{u=u_0 \text{at} \partial\Omega} \int_\Omega W(\nabla u) + \epsilon^2 |\nabla\nabla u|^2\, dx \ . \tag{2}$$

This type of regularization is standard in the literature on coherent phase transformation, see e.g. [15]. Another kind, not much different in practice, is the inclusion of surface energy at phase interfaces, see e.g. [12].

There are two different approaches to this class of problems. The first, more traditional one is to discard (1) as having no sense, and concentrate instead on (2). To make contact with (1), one should then consider the behavior of the minimizer u_ϵ as ϵ tends to zero. The Euler-Lagrange equation is a fourth-order partial differential equation. As $\epsilon \to 0$ there will be a sequence of bifurcations, and one should follow the "principal branch," i.e. the one representing the minimum rather than a saddle point (or local minimum). This program is conceptually simple, but hopelessly impractical except in one space dimension [5, 24].

A second, less traditional approach has been more successful in the multidimensional context. It addresses the $\epsilon = 0$ problem (1) directly, despite the possible nonexistence of a minimizer. Its goal is simply to identify the *minimum value* – which certainly makes sense – and to construct *examples of minimizing sequences*. That goal has been achieved in a remarkable list of examples. Central to this approach is the "relaxation" of (1), a new variational problem of the same type with W replaced by its "quasiconvexification," see e.g. [9, 10, 22, 23]. The case when u is scalar-valued is particularly simple, because in that context the integrand of the relaxed problem is just the ordinary convexification of W [11].

Despite its success, the viewpoint of relaxation has limitations. Knowing the relaxed problem provides some information about the "average character" of a minimizing sequence, but it leaves a great deal of ambiguity concerning the fine-scale structure. This is less than ideal for applications such as the modelling of martensitic phase transformation, where the goal is to predict or explain specific microstructures.

The source of this ambiguity is not the method of relaxation, but rather the formulation of the problem. Minimization of (1) simply does not determine the fine-scale structure of a minimizing sequence. To determine such details one must include some kind of "selection mechanism." Within the context of energy minimization, we are thus obliged to consider a perturbed problem such as (2).

But what, exactly, should be our goal? Seeking the "precise" fine-scale structure of a minimizing sequence seems too ambitious: we don't have a language suitable for describing it. Moreover, it is natural to try and bootstrap from knowledge of the relaxed problem. In

our opinion, the first goal should be to evaluate the *asymptotic behavior* of the energy, to principle order near $\epsilon = 0$. In other words, we seek an expansion of the form

$$\min_{u=u_0 \text{at} \partial\Omega} \int_\Omega W(\nabla u) + \epsilon^2 |\nabla\nabla u|^2 \, dx = E_0 + C\epsilon^\gamma + \cdots \ . \tag{3}$$

The constant E_0 is already known: it is the minimum value of the unperturbed problem (1), provided by the method of relaxation. The crucial new information is the exponent γ, the correction due to the presence of a regularization. Knowing the value of γ does not tell us the details of the microstructure. But it provides a criterion for distinguishing between minimizing sequences which are efficient and those which are not.

One might think that the next goal, after finding γ, should be to find the constant C, then to give further terms in an asymptotic expansion for the energy. Achieving this seems to be difficult. The example discussed below suggests a different goal. Besides finding γ, we are able to identify the *spatial distribution* of the energy. This leads to a fairly unambiguous picture of the minimizer u_ϵ.

The program just summarized is ambitious, and we are far from being able to execute it in general. What we have been been able to achieve is a fairly complete understanding of one simple example.

2 A nonconvex variational problem and its relaxation

Prior to regularization, our example is the variational problem

$$\min_{u=0 \text{ at } x=0} \int_0^1 \int_0^L u_x^2 + \lambda(u_y^2 - 1)^2 \, dxdy \ . \tag{4}$$

The unknown u is a scalar-valued function defined on the rectangle $[0, L] \times [0, 1]$. The parameter λ is positive and fixed. The integrand $W(\nabla u) = u_x^2 + \lambda(u_y^2 - 1)^2$ prefers $\nabla u = (0, 1)$ or $(0, -1)$. This preference is "incompatible" with the boundary condition $u = 0$ at $x = 0$.

It is easy to see that the minimum value of (4) is 0, and that this minimum is not achieved. To give an example of a minimizing sequence, consider any $h > 0$. On the subrectangle $h < x < L$ we take u_h to be a function of y alone, switching between $u_{hy} = 1$ and $u_{hy} = -1$ in layers of thickness h. In the strip $0 < x < h$ we may take u_h piecewise linear, interpolating between $u_h(h, \cdot)$ and the desired boundary value $u_h(0, \cdot) = 0$. This test function has energy of order h, tending linearly to zero as $h \to 0$.

The associated relaxed problem is

$$\min_{u=0 \text{ at } x=0} \int_0^1 \int_0^L u_x^2 + \lambda\Phi(u_y) \, dxdy \ , \tag{5}$$

where $\Phi(t)$ is the convexification of $(t^2 - 1)^2$. Notice that $\Phi(t) = 0$ for $|t| \leq 1$. It is easy to see that the unique minimizer of (5) is the function $u(x, y) \equiv 0$. It follows that every minimizing sequence $\{u_n\}$ for (4) converges weakly to zero. For such a sequence, ∇u_n must oscillate between the preferred values $(0, \pm 1)$, taking each value on volume fraction 50%. More precisely, the Young measure limit of ∇u_n is the sum of two Dirac masses, one at $(0, 1)$ and the other at $(0, -1)$, each with mass $1/2$. (See e.g. [2] or [26] for a discussion of Young measures.)

We shall also have use for the analogue of (4) with inhomogeneous boundary conditions on the left and right sides of the rectangle:

$$\min_{\substack{u=u_0(y)\text{ at }x=0\\ u=u_1(y)\text{ at }x=L}} \int_0^1 \int_0^L u_x^2 + \lambda(u_y^2-1)^2\,dxdy\ . \tag{6}$$

If the boundary functions u_0 and u_1 satisfy $|u_{0y}| \le 1$ and $|u_{1y}| \le 1$ then it is easy to calculate the value of the minimum using the method of relaxation. Indeed, the relaxed problem is the analogue of (5) with the new boundary condition:

$$\min_{\substack{u=u_0(y)\text{ at }x=0\\ u=u_1(y)\text{ at }x=L}} \int_0^1 \int_0^L u_x^2 + \lambda\Phi(u_y)\,dxdy\ .$$

The unique minimizer of the relaxed problem is easily seen to be the linear interpolant

$$u_*(x,y) = \frac{x}{L}u_1(y) + (1-\frac{x}{L})u_0(y)\ .$$

It has $|u_{*y}| \le 1$, so $\Phi(u_{*y}) = 0$. The minimum value of (6) is therefore

$$\int_0^1 \int_0^L (u_*)_x^2\,dxdy = \frac{1}{L}\int_0^1 |u_1(y)-u_0(y)|^2\,dy\ . \tag{7}$$

It is often convenient to focus on the limiting case $\lambda \to \infty$. In this limit the "penalty" $\lambda(u_y^2-1)^2$ becomes the constraint $|u_y| = 1$. The $\lambda = \infty$ version of (4) is thus

$$\min_{\substack{u=0\text{ at }x=0\\ u_y=\pm 1\ a.e.}} \int_0^1 \int_0^L u_x^2\,dxdy\ . \tag{8}$$

We like (8) because it is basically geometric in character. To specify an admissible test function, one has only to give the sets where $u_y = 1$ and $u_y = -1$, and the values of u along a single horizontal line; the function $u(x,y)$ is then fully determined, by integration in y. The test functions admissible for the constrained problem (8) are naturally also admissible for the unconstrained problem (4). We believe that minimizing sequences constructed in the constrained setting will generally also be efficient in the unconstrained setting.

It is easy to relax the constrained problem, either by passing to the limit $\lambda \to \infty$ in (5) or else by convexifying the constraint set $u_y = \pm 1$. In particular, the relaxation of the constrained problem with boundary condition $u = u_0(y)$ at $x = 0$, $u = u_1(y)$ at $x = L$ is

$$\min_{\substack{u=u_0(y)\text{ at }x=0\\ u=u_1(y)\text{ at }x=L\\ |u_y|\le 1}} \int_0^1 \int_0^L u_x^2\,dxdy\ . \tag{9}$$

The boundary data must satisfy $|u_{0y}| \le 1$ and $|u_{1y}| \le 1$, to be consistent with the constraint $|u_y| \le 1$. The minimizer is again the linear interpolant u_*, and the minimum of value of the relaxed problem is again (7).

3 Regularization

Our goal is to understand the role of regularization in determining minimizing sequences for the nonconvex problem (4) or its constrained analogue (8).

One could regularize (4) by adding a term $\epsilon^2|\nabla\nabla u|^2$, as suggested in the introduction. We can make do with less, however: it suffices to add just a term involving u_{yy}^2:

$$\min_{u=0 \text{ at } x=0} \int_0^1 \int_0^L u_x^2 + \lambda(u_y^2 - 1)^2 + \epsilon^2 u_{yy}^2 \, dxdy \ . \tag{10}$$

The rigorous justification for considering (10) a "regularization" of (4) is the fact that (10) achieves its minimum, for any fixed $\epsilon > 0$ [20]. As heuristic justification, we observe that u_x has no incentive to oscillate, while u_y likes to jump between $+1$ and -1. The role of the extra term $\epsilon^2 u_{yy}^2$ is to penalize such jumps. We believe that the inclusion of additional terms in u_{xx}^2 and u_{xy}^2 would not change the minimizers significantly. Indeed, the "interfaces" (where u_y changes from $+1$ to -1) must be approximately horizontal, if u_x is to remain small. This should force u_{xx} and u_{xy} to be much smaller than u_{yy} on average.

The regularization of the constrained problem (8) must be done differently. In view of its geometric character, the natural thing is to add some kind of "surface energy." The simplest choice from the analytical viewpoint is

$$\min_{\substack{u=0 \text{ at } x=0 \\ u_y = \pm 1 \ a.e.}} \int_0^1 \int_0^L u_x^2 + \epsilon |u_{yy}| \, dxdy \ . \tag{11}$$

The term $|u_{yy}|$ in (11) is a singular measure supported on the "interface" where u_y changes from 1 to -1. It is in effect an anisotropic surface energy: $\int\int |u_{yy}|$ is twice the length of the vertical projection of the interface. The rigorous justification of (11) is, once again, the existence of minimizers [20].

We believe that the minimizers of (10) resemble those of (11), in the limit as $\epsilon \to 0$. As evidence, we observe that if u_y makes a transition from -1 to $+1$ in a strip $y_0 - \delta < y < y_0 + \delta$, then along any vertical line $x = x_0$

$$\int_{y_0-\delta}^{y_0+\delta} \lambda(u_y^2 - 1)^2 + \epsilon^2 u_{yy}^2 \, dy \geq \int_{y_0-\delta}^{y_0+\delta} 2\sqrt{\lambda}\epsilon |u_y^2 - 1||u_{yy}| \, dy = \int_{y_0-\delta}^{y_0+\delta} c_0 \epsilon |u_{yy}| \, dy \ ,$$

with $c_0 = \sqrt{\lambda} \int_{-1}^{+1} |t^2 - 1| \, dt$. This inequality becomes sharp if $\epsilon |u_{yy}| = \sqrt{\lambda}|u_y^2 - 1|$, i.e. if u_y has the proper profile in the layer. It follows that if u is locally a function of y alone, then (11) with ϵ replaced by $c_0\epsilon$ gives a sharp lower bound for (10). Our minimizers will not be functions of y alone, but their $x-$dependence will be relatively weak, so we still expect this correspondence between (10) and (11) to be approximately valid.

The rigorous theory presented in [20] and summarized below is primarily concerned with the constrained problem (11). It is also possible to address the unconstrained problem (10) using similar methods; that will be presented elsewhere.

4 Asymptotic behavior of the minimum value

According to the viewpoint explained above, the first task is to evaluate the minimum value of the energy as a function of ϵ. Since the relaxed problem has minimum energy zero, we expect a relation of the form

$$\text{minimum energy } \sim C\epsilon^{\gamma} \tag{12}$$

for either (10) or (11).

In Section 2 we described a minimizing sequence for (4), using piecewise linear u_h which depend only on y except in a layer near $x = 0$. These functions cannot be used directly for (10), because u_{yy} is singular at the edges of the "pieces." But that defect is easily fixed by introducing boundary layers of width approximately ϵ. Taking $\lambda = 1$ for simplicity, one finds that the resulting energy is of order $h + \epsilon L/h$. Minimization with respect to h yields $h \approx \sqrt{\epsilon L}$, and an energy of order $\epsilon^{1/2}L^{1/2}$.

The preceding construction suggests that $\gamma = 1/2$ in (12). This is *wrong*. As we shall explain shortly, the correct exponent γ is not $1/2$ but rather $2/3$. In particular, the minimizing sequence u_ϵ selected by the regularization $\epsilon^2|u_{yy}|^2$ is rather different from the one described above.

It is easy to see what is wrong with u_h. The source of the spatial oscillations is the incompatibility between the preferred gradients $u_y = \pm 1$ and the boundary condition $u = 0$ at $x = 0$. Far from the boundary, however, there is no need for such fine oscillations, and they cost a lot of "regularization energy". So it is better for the oscillations to coarsen as x increases. The coarsening costs some "bulk energy," since u_x can no longer be exactly zero. But the savings in "regularization energy" is sufficient to make up for this cost.

Let us describe a better test function v, working for simplicity with the constrained problem (11) rather than (10). We suppose that along each line $x =$ constant, the graph of $v(x, \cdot)$ is a (more or less) regular sawtooth with $v_y = \pm 1$ and period $h(x)$. Then the typical size of $v(x, y)$ is $h(x)$, so v_x is of order h_x. The energy of v is therefore

$$\int_0^1 \int_0^L u_x^2 + \epsilon|u_{yy}| \sim \int_0^L h_x^2 + \epsilon h^{-1}\, dx \ . \tag{13}$$

We require $h = 0$ at $x = 0$, to accommodate the boundary condition on u. The choice

$$h(x) = c\epsilon^{1/3}x^{2/3} \tag{14}$$

makes the two terms h_x^2 and ϵh^{-1} scale similarly, and it leads to an energy of order $\epsilon^{2/3}L^{1/3}$. (The choice (14) is not optimal, i.e. it does not minimize the right hand side of (13), but it suffices to get the optimal scaling law for the energy.) The preceding discussion resembles one used by Hubert in a different context [16]. It is admittedly somewhat heuristic; see [19] for a more detailed construction. The power law (14) suggests that at distance $x = \ell$ from the left hand boundary, a minimizer should have roughly $\epsilon^{-1/3}\ell^{-2/3}$ transitions between $u_y = -1$ and $u_y = +1$. The construction in [19] achieves this through spatially self-similar branching.

Specific test functions can only give upper bounds on the energy, hence lower bounds on the exponent γ. Our construction shows that the energy can be made as small as $C\epsilon^{2/3}L^{1/3}$, i.e. that $\gamma \geq 2/3$. Might there be other constructions which would give lower values of the energy (larger values of γ)? The answer is no. This is the content of the following result:

Theorem 1 *There exists a positive constant C with the following property: for any function $u(x,y)$ with $u=0$ at $x=0$ and $u_y=\pm 1$ almost everywhere,*

$$\int_0^1 \int_0^L u_x^2 + \epsilon |u_{yy}| \, dxdy \geq C\epsilon^{2/3} L^{1/3} \ .$$

A rigorous proof will be found in [20]. Here we give just the idea. Consider any function $u(x,y)$ satisfying the hypotheses of the theorem. Since $u_y = \pm 1$, the restriction of u to any line $x =$ constant has a graph which looks like a (not necessarily regular) sawtooth. Let $N(x)$ be the number of "teeth". If $N(x)$ is uniformly large then there is a lot of "surface energy," i.e. $\epsilon \int |u_{yy}| \, dxdy = 2\epsilon \int N(x) \, dx$ is large. Suppose on the other hand that $N(x_0)$ is small for some x_0. Then the function $y \mapsto u(x_0, y)$ is a sawtooth function with few teeth, so it must make rather large excursions from 0; thus $\int_0^1 u^2(x_0,y) \, dy$ is large. But then it follows that there is a lot of "bulk energy:" if $u=0$ at $x=0$ and u is large in L^2 at $x=x_0$, then $\int\int u_x^2 \, dxdy$ must be large. In summary, if the "surface energy" $\int\int \epsilon |u_{yy}|$ is small then the "bulk energy" $\int\int u_x^2$ is large. Quantifying this argument leads directly to the conclusion of Theorem 1.

5 Spatial distribution of the energy

It is natural to conjecture that the minimizer of (11) resembles the test function v constructed in Section 4. We cannot prove this in full, but we can show that the spatial distribution of the energy is roughly as suggested by the construction:

Theorem 2 *Let $u = u_\epsilon$ achieve the minimum of (11). Then there is a constant C such that for any $\ell \in (0, L)$,*

$$\int_0^1 \int_0^\ell u_x^2 + \epsilon |u_{yy}| \, dxdy \leq C\epsilon^{2/3} \ell^{1/3} \ ; \tag{15}$$

also, there exists another constant c such that for any $\ell \in (0,L)$,

$$\int_0^1 (u_x^2 + \epsilon |u_{yy}|)(\ell, y) \, dy \geq c\epsilon^{2/3} \ell^{-2/3} \ . \tag{16}$$

Theorems 1 and 2 are very different: the former is an assertion about *any* function u which is admissible for (11); the latter is an assertion about the function u_ϵ which achieves the minimum. It is thus natural that the proof of Theorem 2 should depend on the use of a suitable comparison function. We sketch the main ingredients, referring to [20] for details.

The first step is to compare u with a test function of the form

$$\tilde{u} = \begin{cases} u & \text{for } x \geq \ell \\ w & \text{for } x \leq \ell \ , \end{cases}$$

where w satisfies $w_y = \pm 1$ a.e. and $w = u$ at $x = \ell$. Such a comparison yields

$$\int_0^1 \int_0^\ell u_x^2 + \epsilon |u_{yy}| \, dxdy \leq \int_0^1 \int_0^\ell w_x^2 + \epsilon |w_{yy}| \, dxdy \ . \tag{17}$$

To be of use, w should make the right hand side as small as possible. This is the role of:

Lemma 3 *For any $\delta > 0$, there exists $w(x,y)$ satisfying $w(0,y) = 0$, $w(\ell,y) = u(\ell,y)$, $w_y = \pm 1$ a.e., and*

$$\int_0^1 \int_0^\ell w_x^2 + \epsilon |w_{yy}| \, dxdy \leq \frac{1+\delta}{\ell} \int_0^1 |u(\ell,y)|^2 \, dy + C_\delta \epsilon^{2/3} \ell^{1/3} \ . \tag{18}$$

The estimate (18) is consistent with our general program. Setting $\delta = 0$ in the first term gives the minimum value of the associated *relaxed problem*, a special case of (9). The second term is the correction due to surface energy, and it satisfies the same scaling law as in Theorem 1. Thus the lemma suggests that the exponent γ may be insensitive to the boundary condition; we shall return to this later, in Section 8.

Continuing the argument for Theorem 2, we use (17), (18), and the elementary inequality

$$\frac{1}{\ell} \int_0^1 u^2(\ell,y) \, dy \leq \int_0^1 \int_0^\ell u_x^2 \, dxdy$$

to see that

$$\int_0^1 \int_0^\ell u_x^2 + \epsilon |u_{yy}| \, dxdy \leq (1+\delta) \int_0^1 \int_0^\ell u_x^2 \, dxdy + C_\delta \epsilon^{2/3} \ell^{1/3} \ . \tag{19}$$

At this point we have a problem. The standard idea would be to absorb the term $(1+\delta) \int u_x^2$ into the left hand side. This appears to fail, because the corresponding term on the left has only a coefficient of 1. What saves us is a sort of "equipartition of energy," whereby the two terms on the left hand side of (19) are roughly comparable in magnitude:

Lemma 4 *If u is any stationary point for (11), then*

$$\frac{d}{dx} \int_0^1 u_x^2(x,y) - \epsilon |u_{yy}|(x,y) \, dy = 0 \ . \tag{20}$$

This is the conservation law associated to translation invariance in x. A similar relation holds for stationary points of (10): they must satisfy

$$\frac{d}{dx} \int_0^1 u_x^2(x,y) - \left(\lambda (u_y^2 - 1)^2 + \epsilon^2 u_{yy}^2 \right) (x,y) \, dy = 0 \ . \tag{21}$$

The proof of (21) is utterly elementary: just multiply the Euler-Lagrange equation for (10) by u_x, integrate in y, then integrate by parts. The proof of (20) is a little more technical, since the Euler-Lagrange equation for (11) is awkward to formulate. The successful argument is based on variations of the form $t \mapsto u(x + t\phi(x), y)$.

We make use of (20) as follows. Let c_0 be the value of

$$\int_0^1 u_x^2 \, dy - \int_0^1 \epsilon |u_{yy}| \, dy \ ,$$

which is independent of x. If u is a minimizer rather than just a stationary point, then Theorem 1 gives

$$\int_0^1 \int_0^L u_x^2 \, dxdy \leq C \epsilon^{2/3} L^{1/3} \ , \quad \int_0^1 \int_0^L \epsilon |u_{yy}| \, dxdy \leq C \epsilon^{2/3} L^{1/3} \ ,$$

and it follows that $|c_0| \le C\epsilon^{2/3}L^{-2/3}$. Therefore we have

$$\begin{aligned}\int_0^1 \int_0^\ell u_x^2 \, dxdy &= \int_0^1 \int_0^\ell \epsilon |u_{yy}| \, dxdy + c_0 \ell \\ &\le \int_0^1 \int_0^\ell \epsilon |u_{yy}| \, dxdy + C\epsilon^{2/3}\ell^{1/3}\end{aligned}$$

for any $\ell \le L$. Combining this result with (19) we get

$$\int_0^1 \int_0^\ell u_x^2 + \epsilon |u_{yy}| \, dxdy \le \frac{1+\delta}{2} \int_0^1 \int_0^\ell u_x^2 \, dxdy + \frac{1+\delta}{2} \int_0^1 \int_0^\ell \epsilon |u_{yy}| \, dxdy + C\epsilon^{2/3}\ell^{1/3} \ ,$$

which clearly implies (15).

Given (15), the second assertion (16) is relatively easy. The argument used for Theorem 1 shows that if there is little "surface energy" at $x = \ell$ then there must be a lot of "bulk energy" in the rectangle $0 < x < \ell$. But we know from (15) that the bulk energy $\int_0^1 \int_0^\ell u_x^2 \, dxdy$ is small. So the surface energy $\int_0^1 \epsilon |u_{yy}|(\ell, y) \, dy$ must be large. Quantification of this idea leads to

$$\int_0^1 \epsilon |u_{yy}|(\ell, y) \, dy \ge C\epsilon^{2/3}\ell^{-2/3} \ , \tag{22}$$

which clearly implies (16). Our sketch of the proof of Theorem 2 is now complete.

6 Convergence rate

Since $u \equiv 0$ is the unique solution of the relaxed problem, it is obvious that the minimizers of the regularized problem must tend to 0 in L^2 as $\epsilon \to 0$. Our results yield a precise rate for this convergence:

Theorem 5 *There are constants c and C with the following property: if $u = u_\epsilon$ achieves the minimum of (11) with ϵ sufficiently small, then*

$$c\epsilon^{2/3}L^{7/3} \le \int_0^1 \int_0^L u^2(x, y) \, dxdy \le C\epsilon^{2/3}L^{7/3} \ .$$

Proof: For the upper bound we begin with the elementary elementary inequality

$$\int_0^1 u^2(\ell, y) \, dy \le \ell \int_0^1 \int_0^\ell u_x^2 \, dxdy \ ,$$

which holds for any $0 \le \ell \le L$ and any function $u(x, y)$ with $u = 0$ at $x = 0$. If u minimizes (11) then the right hand side is controlled by (15), and it follows that $\int_0^1 u^2(\ell, y) \, dy \le C\epsilon^{2/3}\ell^{4/3}$. Integration in x leads to the desired upper bound on $\int \int u^2 \, dxdy$.

For the lower bound, we recall that $y \mapsto u(x, y)$ is a "sawtooth" function with slopes ± 1. If the L^2 norm is small then there must be many "teeth," resulting in a lot of surface energy. We quantify this as follows. Let $N(x)$ be the number of transitions at x, i.e. the number of times that $y \mapsto u_y(x, y)$ changes sign. One can show (see Lemma 2.7 of [20]) that

$$\int_0^1 u^2(x, y) dy \ge C(N(x) + 1)^{-2} \ .$$

Integration in x and an application of Jensen's inequality yields

$$\begin{aligned}\frac{1}{L}\int_0^1\int_0^L u^2(x,y)\,dxdy &\geq C\frac{1}{L}\int_0^L (N(x)+1)^{-2}\,dx \\ &\geq C\left(\frac{1}{L}\int_0^L N(x)\,dx+1\right)^{-2} .\end{aligned}$$

But

$$\begin{aligned}\int_0^L N(x)\,dx &= \frac{1}{2}\int_0^1\int_0^L |u_{yy}|\,dxdy \\ &\leq C\epsilon^{-1/3}L^{1/3}\end{aligned}$$

by Theorem 1. If ϵ is small then this gives

$$\frac{1}{L}\int_0^L N(x)\,dx+1 \leq C\epsilon^{-1/3}L^{-2/3}$$

and we conclude that

$$\frac{1}{L}\int_0^1\int_0^L u^2(x,y)\,dxdy \geq C\epsilon^{2/3}L^{4/3} .$$

This is equivalent to the desired lower bound.

Notice that in proving Theorem 5, we established somewhat more. We actually obtained an upper bound on $\int u^2\,dy$ for every x. We suppose the analogous pointwise lower bound should be true, but we do not have a proof. However, by using Theorem 2 one can give a similar lower bound on $\int_0^1\int_0^\ell u^2\,dxdy$ for any $0\leq\ell\leq L$.

7 Soft rather than hard boundary condition

In our examples (10)–(11), minimizers have fine scale structure because the boundary condition $u=0$ at $x=0$ is incompatible with the preferred gradients $\nabla u=(0,\pm 1)$. For the application to twinning in martensite, the "hard" boundary condition $u=0$ is not very realistic. The true physical boundary condition is "softer," i.e. it involves a penalization of $|u|$ at the boundary. We argue in [19] that the variational problem

$$\min_{u_y=\pm 1}\int_0^1\int_0^L u_x^2+\epsilon|u_{yy}|\,dxdy+\beta\left(\int_0^1 u^2(0,y)\,dy\right)^{1/2} \tag{23}$$

provides a reasonable model. Note that (11) is just the formal limit of (23) as $\beta\to\infty$.

The test function v constructed in Section 4 is also admissible for the penalized functional (23), and it shows that

$$\min(23)\leq C\epsilon^{2/3}L^{1/3} ,$$

regardless of the value of β. But now a test function similar to the u_h of Section 3 can also be used, and it leads to the conclusion

$$\min(23)\leq C\beta^{1/2}\epsilon^{1/2}L^{1/2} .$$

Clearly either estimate can be stronger, depending on how β compares with $(\epsilon/L)^{1/3}$.

We believe that what happens is this. If $\beta \ll (\epsilon/L)^{1/3}$ then the minimizer of (23) is more or less independent of x, i.e. it resembles u_h. As β increases with ϵ/L held fixed, there will be a first "twin–branching" bifurcation, then another, and so forth. In the limit $\beta \gg (\epsilon/L)^{1/3}$ the minimizer accumulates increasingly complicated structure, and comes to resemble the test function v described in Section 4 for the hard boundary condition. While we cannot prove that this is the right picture, we prove in [20] that our constructions correctly estimate the minimum energy for all values of β:

Theorem 6 *For any $\beta > 0$, the minimum value of (23) is bounded above and below by a constant times* $\min\{\epsilon^{2/3}L^{1/3}\ ,\ \beta^{1/2}\epsilon^{1/2}L^{1/2}\}$.

8 Other boundary conditions

Our model problems (10)–(11) are special, in that the associated relaxed problems have minimum energy 0. What about other boundary conditions, for which the energy stays bounded away from zero as $\epsilon \to 0$? We have not studied this question systematically, but we have considered in some detail the boundary condition $u = -y$ at $x = 0$, $u = +y$ at $x = L$:

$$\min_{\substack{u=-y\,\mathrm{at}x=0\\ u=+y\,\mathrm{at}x=L\\ |u_y|=\pm 1\ a.e.}} \int_0^1 \int_0^L u_x^2 + \epsilon|u_{yy}|\,dxdy\ . \tag{24}$$

This example is interesting, because neither end condition alone is incompatible with the energy. Rather, refinement is forced by the interaction of two end conditions. The solution of the associated relaxed problem (9) is $u_0(x,y) = (\frac{2x}{L} - 1)y$, which has energy $\int\int u_{0x}^2 = \frac{4}{3L}$. The Young measure limit of any minimizing sequence has mass $\frac{x}{L}$ at $\nabla u = (0,1)$ and mass $1 - \frac{x}{L}$ at $\nabla u = (0,-1)$.

As explained in the introduction, we expect a result of the form

$$\min(24) \sim \frac{4}{3L} + C\epsilon^{\gamma}$$

for the regularized problem. It turns out that the value of γ is still 2/3. Thus, the correction due to regularization seems to be rather insensitive to the specific choice of boundary conditions.

Theorem 7 *There are constants c and C such that*

$$c\epsilon^{2/3}L^{1/3} \le \min(24) - \frac{4}{3L} \le C\epsilon^{2/3}L^{1/3}$$

for ϵ sufficiently close to 0.

The proof will be given in [21]. We also have a result on the L^2 convergence rate, analogous to that of Theorem 5: if u_ϵ achieves the minimum of (24) and u_0 solves the relaxed problem, then

$$c\epsilon^{2/3}L^{7/3} \le \int_0^1 \int_0^L |u_\epsilon - u_0|^2\,dxdy \le C\epsilon^{2/3}L^{7/3}\ .$$

9 Further directions

These examples represent just the bare beginnings of a theory. We comment briefly on some recent, related work and on directions which may be ripe for analysis:

Scalar problems with three or more preferred gradients.

Our $W(\nabla u) = u_x^2 + \lambda(u_y^2 - 1)^2$ prefers just two gradients, $\nabla u = (0, \pm 1)$. What about other choices of W, which prefer three or more possible values of ∇u? We expect the answer to depend on whether the unregularized problem achieves its minimum or not. If it does, then minimizing sequences should develop microstructure only on a lower-dimensional set (e.g. near the boundary); this would surely be reflected in the correction due to surface energy. For affine boundary conditions, we know from the work of and Cellina [6, 7] and Friesecke [13] whether or not the unregularized problem achieves its minimum. Chipot and Müller have recently analyzed the convergence of a finite element minimization scheme [8]; the effect of finite mesh size h is similar to that of nonzero surface energy ϵ (both parameters have the dimensions of length, and serve to set a length scale for the microstructure). Swart and Holmes did careful numerical simulations of viscoelastodynamics, which provide examples of (possibly local) minimizers for problems such as (4); see [25].

It seems to make a difference whether $W(\nabla u)$ prefers finitely many gradients or a continuum of possible values. A simple example of the latter type is $W(\nabla u) = (|\nabla u|^2 - 1)^2$, which prefers $|\nabla u| = 1$. Such an integrand was considered by Aviles and Giga in connection with the modelling of smectic liquid crystals [1]. The same type of integrand has arisen in the recent work of Gioia and Ortiz on the blistering of thin films [14]. The latter work suggests a possible connection between the limit of small surface energy and the theory of viscosity solutions. So far, however, we know of no rigorous result in that direction.

Vector-valued problems

If u is vector-valued rather than scalar-valued, then relaxation is no longer the same as convexification and whole new phenomena arise. A typical example is

$$W(\nabla u) = \begin{cases} 1 + |\nabla u|^2 & \text{if } \nabla u \neq 0 \\ 0 & \text{if } \nabla u = 0, \end{cases} \tag{25}$$

with $u : I\!R^2 \to I\!R^2$. The relaxation was computed explicitly by Kohn and Strang in [22], where one also finds an interpretation of (25) in terms of optimal design. For some boundary conditions – for example, if $u(x) = x$ at $\partial\Omega$ – the variational problem $\int_\Omega W(\nabla u)\, dx$ has *both* generalized solutions (obtained, for example, "sequential lamination") *and* classical ones (obtained, for example, by the "concentric sphere construction"). We expect that the effect of regularization will be to prefer a classical solution, but this has yet to be proved. Much the same problem arises in the modelling of coherent precipitates; see [18] for a discussion of elastic energy minimization in that setting.

References

[1] P. Aviles and Y. Giga, *A mathematical problem related to the physical theory of liquid crystals,* in Proc. Centre Math. Anal., Australian Nat. Univ. (J. Hutchinson and L. Simon, eds.) 12 (1987) 1-16.

[2] J.M. Ball, *A version of the fundamental theorem for Young measures,* in Lecture Notes in Physics 344, M. Rascle et al. eds., Springer (1989) 207-215.

[3] J.M. Ball and R.D. James, *Fine phase mixtures as minimizers of the energy,* Arch. Rat. Mech. Anal. 100 (1987) 13-52.

[4] J.M. Ball and R.D. James, *Proposed experimental tests of a theory of fine microstructure and the two- well problem,* Phil. Trans. Roy. Soc. Lon. 338A (1992) 389-450.

[5] J. Carr, M.E. Gurtin, and M. Slemrod, *Structured phase transitions on a finite interval,* Arch. Rat. Mech. Anal. 86 (1984) 317-351.

[6] A. Cellina, *On minima of a functional of the gradient: necessary conditions,* Nonlin. Anal. – Theory, Methods, and Applns. 20 (1993) 337-341.

[7] A. Cellina, *On minima of a functional of the gradient: sufficient conditions,* Nonlin. Anal. – Theory, Methods, and Applns. 20 (1993) 343-347.

[8] M. Chipot and S. Müller, in preparation.

[9] B. Dacorogna, *Direct Methods in the Calculus of Variations,* Springer, 1989.

[10] G. Dal Maso, *An Introduction to* Γ*-Convergence,* Birkhauser, 1993.

[11] I. Ekeland and R. Temam, *Analyse Convexe et Problèmes Variationnels,* Dunod, 1974.

[12] I. Fonseca, *Phase transitions of elastic solid materials,* Arch. Rat. Mech. Anal. 107 (1989) 195-223.

[13] G. Friesecke, *A necessary and sufficient condition for nonattainment and formation of microstructure almost everywhere in scalar variational problems,* preprint.

[14] M. Ortiz and G. Gioia, *The morphology and folding patterns of buckling-driven thin-film blisters,* preprint.

[15] B. Horovitz, G. Barsch, and J. Krumhansl, *Twin bands in martensites: statics and dynamics,* Phys. Rev. B 43 (1991) 1021-1033.

[16] A. Hubert, *Zur Theorie der Zweiphasigen Domänenstrukturen in Supraleitern und Ferromagneten,* Phys. Status Solidi 24 (1967) 669-682.

[17] R.V. Kohn, *The relaxation of a double-well energy,* Continuum Mech. Thermodyn. 3 (1991) 193-236.

[18] R.V. Kohn and J. Lu, in preparation.

[19] R.V. Kohn and S. Müller, *Branching of twins near an austenite-twinned-martensite interface*, Phil. Mag. A 66 (1992) 697-715.

[20] R.V. Kohn and S. Müller, *Surface energy and microstructure in coherent phase transitions,* Comm. Pure Appl. Math., to appear.

[21] R.V. Kohn and S. Müller, in preparation.

[22] R.V. Kohn and G. Strang, *Optimal design and relaxation of variational problems I–III,* Comm. Pure Appl. Math. 34 (1987) 113-137, 139-182, 353-377.

[23] R.V. Kohn and M. Vogelius, *Relaxation of a variational method for impedance computed tomography,* Comm. Pure Appl. Math. 40 (1987) 745-777.

[24] S. Müller, *Singular perturbation as a selection criterion for periodic minimizing sequences,* Calc. of Variations and PDE 1 (1993) 169-204.

[25] P. Swart and P. Holmes, *Energy minimization and the formation of microstructure in dynamic antiplane shear,* Arch. Rat. Mech. Anal. 121 (1992) 37-85.

[26] L. Tartar, *Compensated compactness and applications to partial differential equations,* in Nonlinear Analysis and Mechanics, R. Knops, ed., Pitman (1978) 136-212.