Variational models for phase transitions
An approach via Γ-convergence

Giovanni Alberti

Introduction

This paper is an extended version of the lecture delivered at the Summer School on *Differential Equations and Calculus of Variations* (Pisa, September 16-28, 1996). That lecture was conceived as an introduction to the theory of Γ-convergence and in particular to the Modica-Mortola theorem; I have tried to reply the style and the structure of the lecture also in the written version. Thus first come few words on the definition and the meaning of Γ-convergence, and then we pass to the theorem of Modica and Mortola. The original idea was to describe both the mechanical motivations which underlay this result and the main ideas of its proof. In particular I have tried to describe a guideline for the proof which would adapt also to other theorems on the same line. I hope that this attempt has been successful. Notice that I never intended to give a detailed and exhaustive description of the many results proved in this field through the recent years, not even of the main ones. In particular the list of references is not meant to be complete, neither one should assume that the contributions listed here are the most relevant or significant.

This paper is organized as follows:

1. A brief introduction to Γ-convergence
2. The Cahn-Hilliard model for phase transitions and the Modica-Mortola theorem
3. The optimal profile problem and the proof of the Modica-Mortola theorem
4. Final remarks

Acknowledgements. I would like to thank Luigi Ambrosio for many useful remarks on a preliminary version. This note was written during a one-year visit at the Max Planck Institute for Mathematics in the Science in Leipzig, whose hospitality and support I gratefully acknowledge.

1. A brief introduction to Γ-convergence

The notion of Γ-convergence was introduced by E. De Giorgi and T. Franzoni in [16]; even though it is mainly intended as a notion of convergence for variational

Paper published in L. Ambrosio, N. Dancer: *Calculus of variations and partial differential equations. Topics on geometrical evolution problems and degree theory*, pp. 95-114 (edited by G. Buttazzo et al.), Springer-Verlag, Berlin-Heidelberg, 2000.

functionals on function spaces, it is more convenient to give its definition and main properties in a slightly more general setting, namely as a notion of convergence for functions on a metric space. Therefore in what follows X is a metric space, u an element of X, F a function from X into $[0,+\infty]$, and ε is a parameter which converges to 0. In the applications X will be a space of functions u on some open domain Ω of $\mathbb{R}^n$, and F a functional on X; typical examples are given by integral functionals on Sobolev or L^p spaces (cf. paragraph 1.3).

What we present here is a rather simplified version of the original definition. A detailed and systematic treatment of the general theory Γ-convergence, and many applications as well, can be found in G. Dal Maso's book [17] (see also [8], section 8).

Warning: Throughout this paper, instead of sequences of functions (and functionals) labelled by some integer parameter which tends to infinity, we consider families of functions labelled by a continuous parameter ε which tends to 0. Nevertheless we use the term "sequence" also to denote such ordered families (and, for instance, we write (u_ε) instead of $\{u_\varepsilon\}$). On this line, a subsequence of (u_ε) is any sequence (u_{ε_n}) such that $\varepsilon_n \to 0$ as $n \to \infty$, and we say that (u_ε) is pre-compact in the corresponding (metric) space X if every subsequence admits a sub-subsequence which converge in X. In proofs we often omit to relabel subsequences.

Definition 1. – *Let X be a metric space, and for $\varepsilon > 0$ let be given $F_\varepsilon : X \to [0,+\infty]$. We say that F_ε* Γ-converge to F on X as $\varepsilon \to 0$*, and we write $F_\varepsilon \xrightarrow{\Gamma} F$, if the following two conditions hold:*

(LB) *Lower bound inequality – for every $u \in X$ and every sequence (u_ε) s.t. $u_\varepsilon \to u$ in X there holds*

$$\liminf_{\varepsilon\to 0} F_\varepsilon(u_\varepsilon) \geq F(u) \ ; \tag{1.1}$$

(UB) *Upper bound inequality – for every $u \in X$ there exists (u_ε) s.t. $u_\varepsilon \to u$ in X and*

$$\lim_{\varepsilon\to 0} F_\varepsilon(u_\varepsilon) = F(u) \ . \tag{1.2}$$

Condition (LB) means that whatever sequence we choose to approximate u, the value of $F_\varepsilon(u_\varepsilon)$ is, in the limit, larger than $F(u)$; on the other hand condition (UB) implies that this bound is sharp, that is, there always exists a sequence (u_ε) which approximates u so that $F_\varepsilon(u_\varepsilon) \to F(u)$.

Remark 1. – When proving a Γ-convergence result, it is often convenient to reduce the amount of verifications and constructions. To this aim we notice that if (LB) holds, then equality (1.2) can be replaced by

$$\limsup_{\varepsilon\to 0} F_\varepsilon(u_\varepsilon) \leq F(u) \ . \tag{1.3}$$

Assume furthermore that we can find a set $\mathscr{D} \subset X$ which satisfies the following condition: for every $u \in X$ there exists an approximating sequence $(u_n) \subset \mathscr{D}$ such that $u_n \to u$ and $F(u_n) \to F(u)$; then a simple diagonal argument shows it is enough to verify condition (UB) for all $u \in \mathscr{D}$ only, and not for every $u \in X$. In fact one can push this argument a bit further, and just verify that for every $u \in \mathscr{D}$ and every $\eta > 0$ there exists a sequence $(u_\varepsilon) \subset X$ (but a subsequence is not enough!) such that $\limsup d(u_\varepsilon, u) \leq \eta$ and $\limsup F_\varepsilon(u_\varepsilon) \leq F(u) + \eta$.

Proposition 1. – *The notion of* Γ*-convergence enjoys the following properties:*

(i) *the* Γ*-limit F is always lower semicontinuous on X;*

(ii) *Stability under continuous perturbations – if $F_\varepsilon \xrightarrow{\Gamma} F$ and G is continuous, then $F_\varepsilon + G \xrightarrow{\Gamma} F + G$;*

(iii) *Stability of minimizing sequences – if $F_\varepsilon \xrightarrow{\Gamma} F$ and v_ε minimizes F_ε over X, then every cluster point of (v_ε) minimizes F over X.*

The proof of this proposition is left to the reader, and we pass to describe how this notion of variational convergence will be used.

1.1. Asymptotic behavior of minimizers and compactness. – Assume that for every $\varepsilon > 0$ we are given a function v_ε which minimizes the functional F_ε on X, and that we want to know what happens of v_ε as $\varepsilon \to 0$. Sometimes the minimizers v_ε can be written via some explicit formula from which we can deduce all information about the asymptotic behavior of v_ε when ε tends to 0. If no such formula is available, and indeed this is often the case, then we can exploit the fact that each v_ε solves the Euler-Lagrange equation associated with F_ε and try to understand which kind of limit equation is verified by a limit point v of (v_ε). Another possibility is to take the Γ-limit F of the functionals F_ε (if any exists), and then use statement (iii) of Proposition 1 to show that any limit point v of v_ε is in fact a minimizer of F, and in particular solves the Euler-Lagrange equation associated with F.

Notice that such a strategy makes sense only if we know *a priori* that the minimizing sequence (v_ε) is pre-compact in X (even the fact that F has some minimizer v does not imply that v is a limit point of v_ε). According to this viewpoint a Γ-convergence result for the functionals F_ε should always be paired with a *compactness result* for the corresponding minimizing sequences (v_ε). In fact one usually tries to prove the following asymptotical equi-coercivity of F_ε:

(C) *Compactness – let be given sequences (ε_n) and (u_n) such that $\varepsilon_n \to 0$ and $F_{\varepsilon_n}(u_n)$ is bounded; then (u_n) is pre-compact in X.*

1.2. Interesting rescalings. – If v_ε minimizes F_ε, then it minimizes also $\lambda_\varepsilon F_\varepsilon$ for every $\lambda_\varepsilon > 0$. This means that information about the limit points of (v_ε) can be recovered also by the Γ-limit of $\lambda_\varepsilon F_\varepsilon$, and different choices of the scaling factors λ_ε generate different Γ-limits, which give different information. It may well happen that the functionals F_ε converge to a constant functional F, so that the fact that every limit point v minimizes F actually gives no information about v, while the Γ-limit of the functionals $\lambda_\varepsilon F_\varepsilon$ may be less trivial for suitable choice of λ_ε (see for instance Remark 8). Therefore the problem arises of finding $\lambda_\varepsilon > 0$ so that the Γ-limit of the rescaled functionals $\lambda_\varepsilon F_\varepsilon$ gives the largest amount of information; sometimes this optimal rescaling is evident but sometimes it is not (compare for instance the situations described in paragraph 1.3 and Theorem 1).

We conclude this section with a simple but instructive example.

1.3. AN EXAMPLE FROM HOMOGENIZATION. – Let X be the class of all u in the Sobolev space $W^{1,2}(0,1)$ such that $u(0)=0$ and $u(1)=1$, endowed with the strong topology of $L^2(0,1)$. Let a be the 1-periodic function on $\mathbb{R}$ which is equal to α_1 on $[0,1/2)$ and to α_2 on $[1/2,1)$, with $0<\alpha_1<\alpha_2<+\infty$, and set

$$F_\varepsilon(u) := \int_0^1 a(x/\varepsilon)\,\big|\dot u(x)\big|^2 dx \; . \tag{1.4}$$

Then the functional F_ε Γ-converge on X to

$$F(u) := \alpha \int_0^1 |\dot u|^2 \quad \text{where} \quad \alpha := \frac{2\alpha_1\alpha_2}{\alpha_1+\alpha_2} \; . \tag{1.5}$$

This is a simple example of *homogenization*, (cf. [17], chapters 24 and 25, or [14]). The proof is quite instructive, but we just give a sketch, leaving the details to the interested reader.

1. Start with the constructive part of the proof, that is, with the upper bound inequality. Take u affine on $(0,1)$ and show that (1.2) can be fulfilled by suitable approximating functions u_ε which are affine on every interval of the type $\big[n\varepsilon,(n+1/2)\varepsilon\big)$ with $n=0,1,\dots$ (these are the intervals where $a(x/\varepsilon)$ is constant).
2. Extend the previous construction to every u which is piecewise affine on $(0,1)$.
3. Use a proper density argument to conclude the proof of the upper bound inequality (cf. Remark 1).
4. Try to understand why the approximation proposed in step 2 is optimal, and then prove the lower bound inequality.

REMARK 2. – The choice of the L^2-topology on X may look unnatural, and indeed a motivation is needed. Since $F_\varepsilon(u) \ge \alpha_1 \int |\dot u|^2$, when $F_\varepsilon(u_\varepsilon)$ is bounded in ε the functions u_ε are weakly pre-compact in $W^{1,2}(0,1)$, but not strongly. Hence the compactness condition (C) in paragraph 1.1 is verified if we endow X with the L^2-topology (recall that the weak topology of $W^{1,2}$ is not metrizable, and anyhow conditions (LB) and (UB) in Definition 1 do not change if we replace the L^2-topology with the weak $W^{1,2}$-topology).

REMARK 3. – The pointwise limit of $F_\varepsilon(u)$ as $\varepsilon\to 0$ is $\overline{F}(u) := \bar\alpha \int |\dot u|^2$ where $\bar\alpha$ is the average of α_1 and α_2, while the value of α in (2.5) is such that $1/\alpha$ is the average of $1/\alpha_1$ and $1/\alpha_2$; in particular $\alpha<\bar\alpha$. Notice that the if we endow X with the strong $W^{1,2}$-topology the Γ-limit of F_ε is $\overline{F}$. This shows that the choice of the topology on X does affect the Γ-limit. In view of paragraph 1.1 the right choice is the L^2-topology, because this way the compactness property (C) is verified.

REMARK 4. – The minimizers v_ε of F_ε over X can be easily computed (at least for $\varepsilon = 1/n$), and then also the limit of v_ε as $\varepsilon\to 0$ can be directly computed. It is interesting to perform such a calculation and then compare with the result obtained via Γ-convergence.

2. The Cahn-Hilliard model for phase transitions and the Modica-Mortola theorem

Consider a container which is filled with two immiscible and incompressible fluids (oil and water, or if you prefer two different phases of the same fluid). In the classical theory of phase transition it is assumed that, at equilibrium, the two fluids arrange themselves in order to minimize the area of the interface which separates the two phases (we neglect the interaction of the fluids with the wall of the container and the effect of gravity). This situation is modelled as follows: the container is represented by a bounded regular domain Ω in $\mathbb{R}^3$, and every configuration of the system is described by a function u on Ω which takes value 0 on the set which is occupied by the first fluid, and value 1 on the set occupied by the second fluid; the singular set of u (i.e., the set of discontinuity points of u) is the interface between the two fluids, and we denote it by Su. The space of all admissible configurations is given by all $u:\Omega\to\{0,1\}$ which satisfy $\int u = V$ where V is the total volume of the second fluid (we assume $0<V<\mathrm{vol}(\Omega)$). Finally we postulate an energy of the form

$$F(u) := \sigma\, \mathscr{H}^2(Su) \; , \tag{2.1}$$

where the parameter σ is called the *surface tension* between the two fluids and $\mathscr{H}^2$ is the two-dimensional Hausdorff measure (when A is a regular surface then $\mathscr{H}^2(A)$ is simply the total area of A). Therefore $F(u)$ is a surface energy distributed on the interface Su, and the equilibrium configuration is obtained by minimizing F over the space all admissible configurations.

An alternative way to study systems of two immiscible fluids is to assume that the transition is not given by a separating interface, but is rather a continuous phenomenon occurring in a thin layer which, on a macroscopic, level we identify with the interface. This means that we allow a fine mixture of the two fluids. In this case a configuration of the system is represented by a function $u:\Omega\to[0,1]$

where $u(x)$ denotes the average volume density of the second fluid at the point $x \in \Omega$ (thus $u(x) = 0$ means that the first fluid only is present at x, $u(x) = 1/2$ means that both fluids are present with the same rate, and so on). The space of all admissible configuration is the class X of all $u : \Omega \to [0,1]$ such that $\int u = V$ (recall that V is the total volume of the second fluid) and to every configuration u is associated the energy

$$E_\varepsilon(u) := \varepsilon^2 \int_\Omega |\nabla u|^2 + \int_\Omega W(u) \ , \tag{2.2}$$

where ε is small positive parameter and W is a continuous positive function which vanishes only at 0 and 1 (in short, a double-well potential). When we come to minimize E_ε, the term $\int W(u)$ favourites those configurations which take values close to 0 and 1 (phase separation), while the term $\varepsilon^2 \int |\nabla u|^2$ penalizes the spatial inhomogeneity of u. When ε is small the first term prevails, and the minimum of E_ε is attained by a function u_ε which takes mainly values close to 0 and 1 (and takes both, because of the volume constraint $\int u = V$) and the transition from 0 to 1 occurs in a thin layer (in fact with thickness of order ε). This model was proposed by J.W. Cahn and J.E. Hilliard in [15]. Notice that the energy E_ε was there obtained as a first order approximation of a more general one.

A connection between the classical model and the Cahn-Hilliard model was established by L. Modica [23], who proved that the minimizers of E_ε converge to minimizers of F. This was obtained by showing that suitable rescalings of the functionals E_ε Γ-converge to F. In order to state the precise Γ-convergence result we need to fix some notation. In what follows Ω is a bounded open set in $\mathbb{R}^N$ (and $N = 3$ is a particular case); we take V so that $0 < V < \mathrm{vol}(\Omega)$ and then we denote by X the class of all measurable functions $u : \Omega \to [0,1]$ such that $\int_\Omega u = V$, endowed with the L^1 norm. We also denote by $BV(\Omega, \{0,1\})$ the set of all functions $u : \Omega \to \{0,1\}$ with bounded variation, and Su is now the set of all *essential* singularities of u (for more details and precise definitions see [18], chapter 5).

Theorem 1. – (L. Modica and S. Mortola [24], see also [23]) *Set* $\sigma := 2\int_0^1 \sqrt{W(u)}\,du$, *and for every* $\varepsilon > 0$ *let*

$$F_\varepsilon(u) := \tfrac{1}{\varepsilon} E_\varepsilon(u) = \begin{cases} \varepsilon \int_\Omega |\nabla u|^2 + \frac{1}{\varepsilon}\int_\Omega W(u) & \text{if } u \in W^{1,2}(\Omega) \cap X, \\ +\infty & \text{elsewhere in } X, \end{cases} \tag{2.3}$$

and

$$F(u) := \begin{cases} \sigma\, \mathscr{H}^{N-1}(Su) & \text{if } u \in BV(\Omega, \{0,1\}) \cap X, \\ +\infty & \text{elsewhere in } X. \end{cases} \tag{2.4}$$

Then the functionals F_ε Γ*-converge to* F *in* X*, and the compactness condition* (C) *in paragraph 1.1 is satisfied.*

Corollary 1. – *If* v_ε *minimizes* F_ε *(or equivalently* E_ε*) on* X*, then the sequence* (v_ε) *is pre-compact in* X*, and every limit point* v *minimizes* F*.*

Remark 5. – Each functional F_ε is lower semicontinuous and coercive with respect to the strong topology of X, and then it has at least one minimizer. Minimizing F over X means finding a set $A \subset \Omega$ among those with prescribed (N-dimensional) volume V which minimizes the (($N-1$)-dimensional) area of $\partial A \cap \Omega$. In particular this implies that $\partial A \cap \Omega$ is an oriented surface with boundary in $\partial\Omega$ and constant mean curvature (by a well-known result of H. Federer [19] this surface is always analytic if $N \le 7$; if $N > 7$ singularities may appear, and the notions of surface and mean curvature must be intended in a particular weak sense which we do not specify here).

Remark 6. – When u is a function in $BV(\Omega, \{0,1\})$, Su is not the set of all points where u is discontinuous (that is, the topological boundary in Ω of the phase $\{u = 1\}$), but the set of all points where u is *essentially discontinuous*, that is, it has no approximate limit. This set agrees with the so-called *measure theoretic boundary* of $\{u = 1\}$ in Ω, and then $F(u)$ is finite if and only if the phase $\{u = 1\}$ (and the phase$\{u = 0\}$) is set of finite perimeter in Ω in the sense of Geometric Measure Theory (see [18], section 5.8). Notice that for such functions u, $F(u)$ is equal to the total variation $\int |Du|$ of the measure derivative Du multiplied by σ.

Remark 7. – The existence of a minimizer of E_ε over X may be proved via standard lower semicontinuity and compactness results. This minimizer may be not unique. Indeed if Ω is a ball centered in 0 and the minimizer of E_ε is unique, then it must be invariant under rotation, i.e., is radially symmetric. On the other hand a simple computation shows that F has no radially symmetric minimizer, and then we deduce that the minimizers of E_ε cannot be radially symmetric for values of ε arbitrary close to 0 (recall Corollary 1), and then the minimizer of E_ε is not unique for every ε sufficiently small.

Remark 8. – To a certain extent the rescaling $F_\varepsilon := \frac{1}{\varepsilon} E_\varepsilon$ given in Theorem 1 is optimal. Indeed one can easily check out the following table:

a. $E_\varepsilon \xrightarrow{\Gamma} E$ with $E(u) = \int_\Omega W(u)$;
b. if $\lambda_\varepsilon \to \infty$ and $\varepsilon\lambda_\varepsilon \to 0$ then $\lambda_\varepsilon E_\varepsilon \xrightarrow{\Gamma} E$ with $E(u) = 0$ when u takes the values 0 and 1 a.e. in Ω, and $E(u) = +\infty$ otherwise;
c. if $\varepsilon\lambda_\varepsilon \to \infty$ then $\lambda_\varepsilon E_\varepsilon \xrightarrow{\Gamma} E$ with $E(u) = +\infty$ everywhere in X.

In all these cases the set of minimizers of the Γ-limit strictly includes the minimizers of F.

3. The optimal profile problem and the proof of the Modica-Mortola theorem

In this section we give the main ideas of the proof of Theorem 1, and we also try to point out the underlying technical tools. The usual proof is indeed simple and elegant, but also quite specific (see [23], and paragraph 4.5); in particular it hardly adapts even to close generalizations of this statement. For these reasons I prefer to describe here a different approach, which should give a deeper insight in the structure of this theorem, and is probably more flexible. This is also an attempt to gather some important ideas which are scattered in the literature, and therefore not immediately available to non-experts. We remark that another very general approach to the proof of lower bound inequalities (in relaxation but also in Γ-convergence) based on an extensive use of blow-up techniques has been developed in [20]; see also [3] and [11] for applications to theorems of Modica-Mortola type.

First of all we notice that Theorem 1 reduces to the following three statements:

(i) *compactness – let be given sequences* (ε_n) *and* (u_n) *such that* $\varepsilon_n \to 0$ *and* $F_{\varepsilon_n}(u_n)$ *is bounded; then* (u_n) *is pre-compact in* $L^1(\Omega)$ *and every limit point belongs to* $BV(\Omega,\{0,1\})$;

(ii) *lower bound inequality – if* $u \in BV(\Omega,\{0,1\})$, $(u_\varepsilon) \subset W^{1,2}(\Omega)$ *and* $u_\varepsilon \to u$ *in* $L^1(\Omega)$ *then*

$$\liminf_{\varepsilon\to 0} F_\varepsilon(u_\varepsilon) \geq \sigma\, \mathscr{H}^{N-1}(Su) \; ; \tag{3.1}$$

(iii) *upper bound inequality – for every* $u \in BV(\Omega,\{0,1\})$ *exists* $(u_\varepsilon) \subset W^{1,2}(\Omega)$ *such that* $u_\varepsilon \to u$ *in* $L^1(\Omega)$, $\int u_\varepsilon = \int u$ *for every* ε *and*

$$\limsup_{\varepsilon\to 0} F_\varepsilon(u_\varepsilon) \leq \sigma\, \mathscr{H}^{N-1}(Su) \; . \tag{3.2}$$

Warning: in the following we consider only functions which take values in the interval $[0,1]$. For sequences of such functions convergence in measure is equivalent to convergence in L^p for any $p \in [1,\infty)$, and then we simply call it *strong convergence*. Similarly, all the weak L^p-topologies with $p \in [1,\infty)$ induce the same convergence, which is therefore referred to as *weak convergence*.

We first prove the three statements above in the one-dimensional case, and then we briefly show how to pass to the two-dimensional case, which requires the same amount of work as the passage to arbitrary dimension.

3a. The one-dimensional case

We can assume that Ω is an open bounded interval. In this case $BV(\Omega,\{0,1\})$ turns out to be the class of all $u : \Omega \to \{0,1\}$ with finitely many discontinuities (that is, piecewise constant on Ω), and $\mathscr{H}^0$ is simply the measure that counts

points. We often consider functions u such that $F_\varepsilon(u)$ is finite; then u belongs to $W^{1,2}(\Omega)$, and in particular it admits a continuous representant on $\overline{\Omega}$; unless differently stated we always refer to this representant. The basic ingredients of the proof are given in the following three paragraphs.

3.1. Localization of F_ε. – It is useful to consider F_ε also as a function of the integration domain; hence we set

$$F_\varepsilon(u,A) := \varepsilon \int_A |\dot u|^2 + \frac{1}{\varepsilon}\int_A W(u) \tag{3.3}$$

(for every measurable set A and every function u whose derivative belongs to $L^2(U)$ for some open set U which contains A). In particular $F_\varepsilon(u) = F_\varepsilon(u,\Omega)$. The functional in (3.3) is sometimes called the *localization* of the functional F_ε given in (2.3). Notice that $F_\varepsilon(u,A)$ is a positive measure with respect to the variable A; in the following we often use this property without further mention.

3.2. Scaling property of F_ε. – For every A and every u we set $u^\varepsilon(x) := u(\varepsilon x)$ and $\frac{1}{\varepsilon}A := \{x : \varepsilon x \in A\}$. Then we immediately obtain the following *scaling identity*

$$F_\varepsilon(u,A) = F_1\big(u^\varepsilon, \tfrac{1}{\varepsilon}A\big) \; . \tag{3.4}$$

In some sense we may say that the choice of scaling $F_\varepsilon := \frac{1}{\varepsilon}E_\varepsilon$ is the optimal one exactly because of identity (3.4).

3.3. The optimal profile problem. – We consider now the minimum problem

$$\bar\sigma := \inf\Big\{F_1(u,\mathbb{R}) \; : \; u : \mathbb{R} \to [0,1]\, , \lim_{x\to-\infty} u(x) = 0\, , \lim_{x\to+\infty} u(x) = 1\Big\} \; . \tag{3.5}$$

The number $\bar\sigma$ represents the minimal cost in term of the energy F_1 for a transition from the value 0 to the value 1 on the entire real line. The minimum problem (3.5) is therefore called the *optimal profile problem*, and a solution γ is an *optimal profile for transition* (with respect to the non-scaled energy F_1). We will show that the optimal profile problem is the key to the proof of Theorem 1, and in particular $\bar\sigma$ turns out to be equal to the constant σ in the statement of Theorem 1 (see Proposition 2). The connection between $\bar\sigma$ and the cost of transition relative to the energy F_ε is made by the following lemma:

Lemma 1. – *Let be given an interval* I *and a function* $u : I \to [0,1]$. *Assume that there exists* $a, b \in I$ *and* $\delta > 0$ *so that* $u(a) \leq \delta$ *and* $u(b) \geq 1-\delta$. *Then for every* $\varepsilon > 0$ *there holds*

$$F_\varepsilon(u,I) \geq \bar\sigma - O(\delta) \; , \tag{3.6}$$

where the error estimate $O(\delta)$ *depends on* δ, *but neither on* ε *nor on* u.

Proof. We can assume that $a < b$ and $I = (a,b)$, otherwise we just replace I with (a,b). By identity (3.4) it suffices to prove (3.6) when $\varepsilon = 1$. In order to compare $F_\varepsilon(u,I)$ with $\bar\sigma$ we extend u to the whole of $\mathbb{R}$ as shown in the picture.

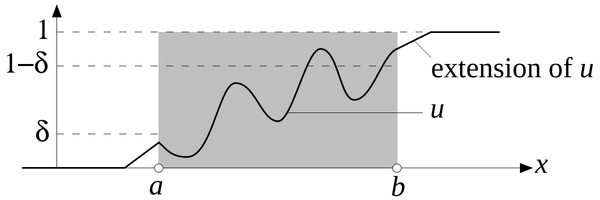

Fig.1. Extension of the function u out of $I = (a,b)$

Now an immediate estimate shows that $F_1(u, \mathbb{R}\setminus I) \le \delta + C\delta$ where C is the maximum value of W in $[0,1]$. Hence we set $O(\delta) := (1+C)\delta$ and we conclude

$$F_1(u,I) = F_1(u,\mathbb{R}) - F_1(u,\mathbb{R}\setminus I) \ge \bar\sigma - O(\delta)\ . \quad \square$$

Proposition 2. – *The minimum in* (3.5) *is attained and agree with the constant* σ *given in Theorem 1, that is,* $\bar\sigma = 2\int_0^1 \sqrt{W(u)}\,du$.

Proof. We first prove that $F_1(u,\mathbb{R}) \ge \sigma$ for every $u : \mathbb{R} \to [0,1]$ which tends to 1 (resp. to 0) at $+\infty$ (resp. at $-\infty$), and then we show that the equality is attained by some particular choice of u. We apply the inequality $a^2 + b^2 \ge 2ab$, with $a := \dot u(x)$ and $b := \sqrt{W(u(x))}$, to the definition of $F_1(u,\mathbb{R})$, and then we get

$$\begin{aligned} F_1(u,\mathbb{R}) &= \int_{-\infty}^{+\infty} \big[\dot u^2(x) + W(u(x))\big]\,dx \\ &\ge 2\int_{-\infty}^{+\infty} \sqrt{W(u(x))}\,\dot u(x)\,dx = 2\int_0^1 \sqrt{W(u)}\,du =: \sigma\ . \end{aligned} \tag{3.7}$$

Now we recall that equality holds in $a^2 + b^2 \ge 2ab$ when $a = b$, and then equality holds in (3.7) when u satisfies the differential equation $\dot u = \sqrt{W(u)}$. Thus we consider the Cauchy problem

$$\begin{cases} \dot u = \sqrt{W(u)}\,, \\ u(0) = 1/2\,. \end{cases} \tag{3.8}$$

The constant functions 1 and 0 are solution of $\dot u = \sqrt{W(u)}$ (because $W(0) = W(1) = 0$), and since $\sqrt{W}$ is continuous the problem (3.8) admits a global (increasing) solution on $\mathbb{R}$ (we have uniqueness when W is of class C^1). Since $W(u) > 0$ for $0 < u < 1$, this global solution must converge to 1 (resp. to 0) at $+\infty$ (resp. at $-\infty$). $\square$

Remark 9. – The possibility of an explicit computation of $\bar\sigma$ is quite specific of the form of the functional F_1. In many generalizations such a computation is not possible (cf. paragraphs 4.2 and 4.4). For this reason it is preferable to define σ and $\bar\sigma$ separately, and then show that $\sigma = \bar\sigma$. For the rest of the proof we will refer mainly to $\bar\sigma$, this being the "natural" constant to consider.

3.4. Proof of statements (i) and (ii). – We first sketch the idea of the proof. If we take a sequence (u_ε) such that $F_\varepsilon(u_\varepsilon) \le C < +\infty$, then in particular $\int_\Omega W(u_\varepsilon) \le C\varepsilon$, and this implies that the functions u_ε take values close to 0 or 1 outside an exceptional set with measure of order ε (recall that W is continuous and strictly positive between 0 and 1). If the sequence (u_ε) converges weakly but not strongly, then it must "oscillate" between values close to 0 and 1; on the other hand Lemma 1 shows that the cost of each oscillation (in term of the localized energy F_ε) is roughly of order $\bar\sigma$, and then the bound on $F_\varepsilon(u_\varepsilon)$ allows only for finitely many oscillations. Hence (u_ε) converges strongly to a limit function u which takes values 0 or 1 almost everywhere.

Let us compute now the number of transitions of u from 0 to 1 or viceversa. Passing to a subsequence, and modifying u in a set of measure 0, we may assume that u_ε converge to u everywhere. Then, if we take x_0 and x_1 so that $u(x_0) = 0$ and $u(x_1) = 1$, by (3.6) we get $F_\varepsilon\big(u_\varepsilon, (x_0,x_1)\big) \ge \bar\sigma - o(1)$. Hence the bound $F_\varepsilon(u_\varepsilon) \le C$ implies that u has at most $C/\bar\sigma$ transitions from 0 to 1, that is, $\mathscr{H}^0(Su) \le C/\bar\sigma$. Eventually we notice that passing to subsequences we can take the bound $C := \sup F_\varepsilon(u_\varepsilon)$ arbitrarily close to the lower limit of $F_\varepsilon(u_\varepsilon)$, and then (3.1) is proved.

The previous heuristic argument can be made rigorous in a very elegant and simple way by use of the notion of Young measure. To this aim we refer the interested reader to the concise paper [9]; a more detailed and exhaustive treatment of the subject can be found in [27], while a how-to-use guide will be provided in the first chapters of [25]. For our purposes it should be enough to recall that given an arbitrary sequence of functions $u_\varepsilon : \Omega \to [0,1]$, we may extract a subsequence (not relabeled) whose asymptotic behavior is captured by a certain family of probability measures $\{\nu_x : x \in \Omega\}$ called the *Young measure generated by* (u_ε). Every ν_x is a probability measure on the interval $[0,1]$ related with the asymptotic distribution of the values of u_ε, close to x. In particular the functions u_ε converge weakly to the function u which takes every $x \in \Omega$ into the center of mass of ν_x (that is, $u(x) := \int u\,d\nu_x(u)$), and we have strong convergence if and only if ν_x is a Dirac mass for a.e. $x \in \Omega$. Moreover for every test function $f : \Omega \times [0,1] \to \mathbb{R}$ which is continuous with respect to the second variable and bounded there holds

$$\int_\Omega f\big(x, u_\varepsilon(x)\big)\,dx \longrightarrow \int_\Omega \Big[\int_0^1 f(x,u)\,d\nu_x(u)\Big]\,dx\ . \tag{3.9}$$

Take now a sequence (u_ε) such that $F_\varepsilon(u_\varepsilon) \le C < +\infty$, and assume that it generates a Young measure $\{\nu_x : x \in \Omega\}$. Then $\int_\Omega W(u_\varepsilon) \to 0$, and if we apply (3.9) with $f(x,u) := W(u)$ we get

$$\int_\Omega \Big[\int_0^1 W(u)\, d\nu_x(u) \Big]\, dx = 0 \ .$$

We infer that for a.e. $x \in \Omega$ the measure ν_x is supported on $\{0,1\}$, and then it can be written as

$$\nu_x := \lambda(x) \cdot \delta_0 + \big(1 - \lambda(x)\big) \cdot \delta_1 \ , \tag{3.10}$$

for suitable $\lambda(x) \in [0,1]$ (as usual δ_t denotes the Dirac mass at t). Take now any interval $I \subset \Omega$ where λ is neither a.e. equal to 0 nor a.e. equal to 1. Then the functions u_ε must take in I values close to 0 and values close to 1, and then Lemma 1 yields $F_\varepsilon(u_\varepsilon, I) \ge \bar\sigma - o(1)$. Therefore the bound on $F_\varepsilon(u_\varepsilon)$ implies that we can find at most $C/\bar\sigma$ disjoint intervals of this type. Hence λ agrees a.e. with a function u which takes values 0 or 1 everywhere and has finitely many transition from 0 to 1; in short, $u \in BV(\Omega, \{0,1\})$. Hence ν_x is a Dirac mass at 0 for a.e. x such that $u(x) = 0$, and a Dirac mass at 1 for a.e. x such that $u(x) = 1$. This implies that the limit of the functions u_ε is u, and the convergence is strong; statement (i) is proved.

Since the number of transition of u from 0 to 1 is by definition $\mathscr{H}^0(Su)$, the previous argument provides the following bound:

$$C \ge \bar\sigma \mathscr{H}^0(Su) \ . \tag{3.11}$$

Inequality (3.11) implies (3.1) because, passing to subsequences, we can take the upper bound C arbitrarily close to the lower limit of $F_\varepsilon(u_\varepsilon)$.

3.5. Proof of statement (iii). – Once the proof in paragraph 3.4 and the meaning of the optimal profile problem (3.5) are well understood, the rest of the proof of Theorem 1, namely the upper bound inequality, is almost immediate. Assume for simplicity that Ω is an interval which contains the point 0 and that $u(x) = 1$ for $x \ge 0$ and $u(x) = 0$ for $x < 0$. Thus Su consists of the sole point 0, and we construct the approximating sequence (u_ε) which satisfies (3.2) by taking suitable scaling of the optimal profile γ; more precisely we set $u_\varepsilon(x) := \gamma(x/\varepsilon)$. Then $u_\varepsilon(x) \to u(x)$ for every $x \ne 0$ because the limit of γ at $+\infty$ is 1 and the limit at $-\infty$ is 0, and identity (3.4) yields

$$F_\varepsilon(u_\varepsilon) = F_1(\gamma, \tfrac{1}{\varepsilon}\Omega) \le F_1(\gamma, \mathbb{R}) = \bar\sigma \ .$$

So far we did not care about the constraint $\int u_\varepsilon = \int u$; in order to fit it, one has to slightly modify the previous definition, for instance by taking carefully chosen translations of u_ε. The construction of (u_ε) in the general case is sketched in the picture below; the details are left to the reader.

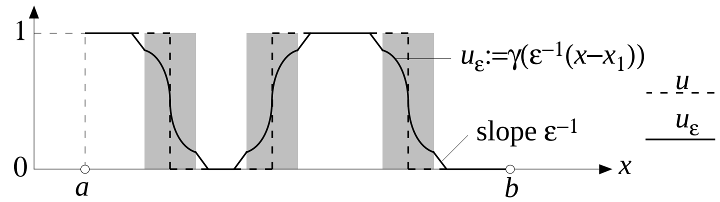

Fig.2. Construction of u_ε, with $\Omega = (a,b)$ and $Su = \{x_1, x_2, x_3\}$

3b. The general case

The proof of Theorem 1 in dimension larger than one is usually achieved by suitable adaptations of the proof for the one-dimensional case. Yet it is important to notice that statements (i) and (ii) can be deduced directly from the corresponding one-dimensional statements via a general slicing argument (we recall that statements (i) - (iii) are given at the beginning of section 3, and proved for the one-dimensional case in subsection 3a). This fact shows that to some extent the nature of Theorem 1 is one-dimensional; for instance it can hardly adapt to the so-called Ginzburg-Landau functionals (see paragraph 4.3). We remark that the use of slicing arguments to reduce to lower-dimensional statements has already been applied in many different situations; we just mention here the proof of the compactness theorem for integral currents due to B. White [28], the original proof of the compactness theorem for SBV functions in [7], and the rank-one property of derivatives of BV functions in [1]. For simplicity we give the proof in the two-dimensional case only.

3.6. Some notation. – We assume for the time being that Ω is a rectangle of the form $I \times J$, with I, J open intervals with length smaller than 1, and we write every $x \in \Omega$ as $x = (y,z)$ with $y \in I$, $z \in J$. For every function u defined on Ω and every $y \in I$ we denote by u^y the function on J defined by $u^y(z) := u(y,z)$, and for every $z \in J$ we denote by u^z the function on I defined by $u^z(y) := u(y,z)$. The functions u^y and u^z are called *one-dimensional slices* of u. We denote by $\overline{F}_\varepsilon(u, A)$ the one-dimensional functional given in (3.3) for every open interval A and every $u : A \to [0,1]$. We recall now that if $u \in W^{1,2}(\Omega)$ then $u^y \in W^{1,2}(J)$ for a.e. $y \in I$ and $u^z \in W^{1,2}(I)$ for a.e. $z \in J$, and

$$\frac{\partial u}{\partial z}(x) = \dot u^y(z) \ , \quad \frac{\partial u}{\partial y}(x) = \dot u^z(y) \quad \text{for a.e. } x \in \Omega$$

(see [18], section 4.9.2). Since $|\nabla u|^2 \ge \big|\frac{\partial u}{\partial z}\big|^2$, (resp. $\big|\frac{\partial u}{\partial y}\big|^2$), we immediately obtain the following *slicing inequalities* :

$$F_\varepsilon(u) \ge \int_I \overline{F}_\varepsilon(u^y, J)\, dy \qquad \Big(\text{resp. } \int_J \overline{F}_\varepsilon(u^z, I)\, dz\Big) \ . \tag{3.12}$$

3.7. A compactness criterion via slicing. – In order to deduce statement (i) from the corresponding one-dimensional compactness statement, we have to make a connection between the pre-compactness of a sequence (u_ε) of functions from Ω into $[0,1]$ and the pre-compactness of the slices (u^y_ε) and (u^z_ε). The simplest result on this line is the following criterion:

(C1) *assume that* $(u^y_\varepsilon) \Subset L^1(J)$ *for a.e.* $y \in I$ *and* $(u^z_\varepsilon) \Subset L^1(I)$ *for a.e.* $z \in J$*; then* $(u_\varepsilon) \Subset L^1(\Omega)$

(here $* \Subset **$ reads as "$*$ is pre-compact in $**$"). Unfortunately this result does not fit our purposes, but a sufficiently general statement is obtained by allowing for some "perturbations": we say that a sequence $(\bar u_\varepsilon)$ is *δ-close to* (u_ε) if $\|u_\varepsilon - \bar u_\varepsilon\|_1 < \delta$ for every ε, and then we have the following modification of (C1):

(C2) *assume that for every* $\delta > 0$ *there exist sequences* $(u_{\delta,\varepsilon})$ *and* $(\hat u_{\delta,\varepsilon})$ *δ-close to* (u_ε) *so that* $(u^y_{\delta,\varepsilon}) \Subset L^1(J)$ *for a.e.* $y \in I$ *and* $(\hat u^z_{\delta,\varepsilon}) \Subset L^1(I)$ *for a.e.* $z \in J$*; then* $(u_\varepsilon) \Subset L^1(\Omega)$.

The proof of these compactness criteria essentially relies on the characterization of pre-compact sets in the strong L^1-topology given by the Fréchet-Kolmogorov theorem; for a general statement and a detailed proof we refer to [6], section 6.3.

3.8. Proof of statements (i) and (ii). – Let be given (u_ε) so that $F_\varepsilon(u_\varepsilon) \le C < +\infty$. Then (3.12) yields

$$C \ge \int_I \overline{F}_\varepsilon(u^y_\varepsilon, J)\, dy \quad \text{for every } \varepsilon > 0 \tag{3.13}$$

(and a similar inequality holds for $\overline{F}_\varepsilon(u^z_\varepsilon, I)$). We fix now $\delta > 0$ and for every $\varepsilon > 0$ we take $u_{\delta,\varepsilon} : \Omega \to [0,1]$ so that

$$u^y_{\delta,\varepsilon} := \begin{cases} u^y_\varepsilon & \text{if } \overline{F}_\varepsilon(u^y_\varepsilon, J) \le C/\delta, \\ 0 & \text{otherwise.} \end{cases}$$

By (3.13) we have that $u^y_{\delta,\varepsilon} = u^y_\varepsilon$ for all $y \in I$ apart a set of measure smaller than δ, and then $\|u_\varepsilon - u_{\delta,\varepsilon}\|_1 \le \delta |I| \le \delta$. Hence the sequence $(u_{\delta,\varepsilon})$ is δ-close to (u_ε). Moreover for every $y \in I$ there holds $\overline{F}_\varepsilon(u^y_{\delta,\varepsilon}, J) \le C/\delta$ (recall that $\overline{F}_\varepsilon(0, I) = 0$) and then the sequence $(u^y_{\delta,\varepsilon})$ is pre-compact in $L^1(J)$ by the one-dimensional version of statement (i). In a similar way we can construct a sequence $(\hat u_{\delta,\varepsilon})$ δ-close to (u_ε) so that $(\hat u^z_{\delta,\varepsilon}) \Subset L^1(I)$ for a.e. $z \in J$, and therefore the compactness criterion (C2) shows that the sequence (u_ε) is pre-compact in $L^1(\Omega)$.

Assume now that (u_ε) converge to u in $L^1(\Omega)$. Then $u^y_\varepsilon \to u^y$ in $L^1(J)$ for a.e. $y \in I$, and if we apply Fatou's lemma to inequality (3.12) we get

$$\liminf_{\varepsilon \to 0} F_\varepsilon(u_\varepsilon) \ge \int_I \big[\liminf_{\varepsilon \to 0} \overline{F}_\varepsilon(u^y_\varepsilon, J) \big]\, dy\ .$$

Hence $\liminf \overline{F}_\varepsilon(u^y_\varepsilon, J)$ is finite for a.e. $y \in I$, and then u^y belongs to the space $BV(J, \{0,1\})$ by the one-dimensional version of (i). Moreover the the one-dimensional version of (ii) yields

$$\liminf_{\varepsilon \to 0} F_\varepsilon(u_\varepsilon) \ge \sigma \int_I \mathscr{H}^0(Su^y)\, dy\ , \tag{3.14}$$

and recalling that $\mathscr{H}^0(Su^y)$ is the total variation of u^y in $BV(J)$, (3.14) yields $\int_I \|\dot u^y\|\, dy < +\infty$, and in a similar way one gets $\int_J \|\dot u^z\|\, dz < +\infty$. We use now the following important fact: a function $u \in L^1(\Omega)$ belongs to $BV(\Omega)$ if (and only if) $u^y \in BV(J)$ for a.e. $y \in I$, $u^z \in BV(I)$ for a.e. $z \in J$, and $\int_I \|\dot u^y\|\, dy$ and $\int_J \|\dot u^z\|\, dz$ are finite (see [18], section 5.10.2).

Finally we recover the lower bound inequality (3.1) from (3.14). Assume for the moment that Su is a regular curve in Ω. Then the integral at the right hand side of (3.14) is just the measure of the projection of Su on I (keeping the multiplicity into account), which in general can be smaller than the length of Su. In fact the length of a curve C in $I \times J$ is close to the measure of its projection on I if and only if the normal to C is close to the projection axis (in this case the z-axis). Keeping this in mind we cover Su (up to a subset with small measure) by pairwise disjoint squares Q_i so that within each Q_i the normal to Su is close to one of the axes of Q_i.

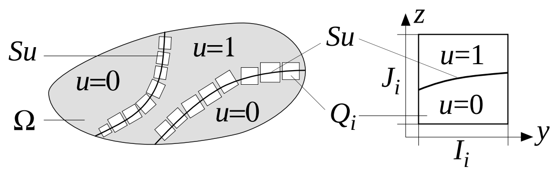

Fig.3. Covering the set Su with squares

Then, for every Q_i, inequality (3.14) becomes

$$\liminf_{\varepsilon \to 0} F_\varepsilon(u_\varepsilon, Q_i) \ge \sigma \int_{I_i} \mathscr{H}^0(Su^y)\, dy \simeq \sigma \mathscr{H}^1(Su \cap Q_i)\ ,$$

and taking the sum over all i,

$$\liminf_{\varepsilon \to 0} F_\varepsilon(u_\varepsilon) \ge \sum_i \liminf_{\varepsilon \to 0} F_\varepsilon(u_\varepsilon, Q_i) \ge \sigma \mathscr{H}^1(Su) - o(1)\ .$$

This argument can be made rigorous for every singular set Su by a careful use of the Besicovitch covering theorem (see [18], section 1.5) and a detailed description of the pointwise property of rectifiable sets and of the measure theoretic boundary of sets with finite perimeter (cf. [18], chapter 5).

3.9. Proof of statement (iii). – Unlike statements (i) and (ii), the proof of statement (iii) cannot be achieved by reduction to the one-dimensional case. On the other hand by Remark 1 it is enough to prove the upper bound inequality only for a suitable dense subset $\mathscr{D}$ in X. In this case we choose as $\mathscr{D}$ the class of all $u \in BV(\Omega, \{0,1\})$ whose singular set Su is a piecewise affine curve in $\mathbb{R}^2$ (a polyhedral surface of dimension $N-1$ when N is general); indeed every $u \in BV(\Omega, \{0,1\})$ can be approximated by a sequence $(u_n) \subset \mathscr{D}$ so that $\mathscr{H}^1(Su_n) \to \mathscr{H}^1(Su)$; this is an immediate consequence of a well-known approximation result for finite perimeter sets by smooth sets, see [21], Theorem 1.24 (in fact, another typical choice for $\mathscr{D}$ is given by the class of all u such that Su is a smooth curve with boundary included in $\partial\Omega$). Thus we take $u \in \mathscr{D}$, and given $\varepsilon > 0$ we construct u_ε as follows: we cover Su with disjoint rectangles R_i with width $\varepsilon^{2/3}$, up to a residual set with measure of order $\varepsilon^{2/3}$:

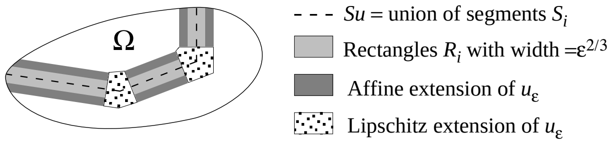

Fig.4. Covering of Su

In each rectangle R_i (see Fig. 4) we set $u_\varepsilon(x) := \gamma(x_i/\varepsilon)$ where x_i is the oriented distance of the point x from the segment S_i (so that it is positive on the side of S_i where $u = 1$ and negative on the side where $u = 0$). In the darker rectangles we take an affine extension of u_ε which agrees with u on the sides which border the white region, and we choose the width of each rectangle so that the slope of u_ε is $1/\varepsilon$ (therefore this width has order $o(\varepsilon)$). In the white region we take u_ε equal to u. Finally, we define u_ε in the interior of the remaining dotted regions by taking any Lipschitz extension with the same Lipschitz constant as on the boundary, which has order $O(\varepsilon^{-1})$. Hence, within each R_i (and in the corresponding darker rectangles) the function u_ε varies only in the direction ν_i normal to S_i, and $F_\varepsilon(u_\varepsilon, R_i) \le \sigma\mathscr{H}^1(S_i)$; while the contributions of the other regions vanish as $\varepsilon \to 0$. Thus $F_\varepsilon(u_\varepsilon) \le \sigma\mathscr{H}^1(Su) + o(1)$, and the proof of statement (iii) is completed.

4. Final remarks

Theorem 1 was conjectured by E. De Giorgi and proved by L. Modica and S. Mortola [24] in 1977, shortly after the definition of Γ-convergence was given in [16]; the connection with the Cahn-Hilliard model was established by L. Modica [23] only in 1987. Since then several results were given which extend Theorem 1 in different directions (cf. paragraphs 4.1, 4.3 and 4.4). The idea of defining $\bar\sigma$ via a suitable minimum problem involving only non-scaled functionals is common knowledge, and I was not able to trace the source. The idea of proving compact-

ness and lower bound inequality via Young measures (and Lemma 1) is essentially contained in [5].

Remark 10. – The minimum problem (3.5) leads to the Euler-Lagrange equation $2\ddot u - \dot W(u) = 0$. Hence solution of (3.5) are standing waves for the parabolic equation $u_t = 2u_{xx} - f(u)$ for certain choices of f. Standing and travelling waves for this equation have been widely studied in the literature, since connected with the asymptotic behavior of the solutions of the scalings of the Allen-Cahn equation $u_t = \Delta u - f(u)$; we refer the reader to [8] and the references therein.

Remark 11. – Notice that throughout the whole proof of Theorem 1, what we really need is the positivity of $\bar\sigma$, while it is not strictly necessary that the infimum in (3.5) is attained (in fact suitable modification of the proof of statement (iii) works even if no optimal profile is available). Nevertheless the existence of the optimal profile has a deeper meaning than it appears in the proof above. Indeed if (v_ε) is a sequence of minimizers of F_ε which converges to some $v \in BV(\Omega, \{0,1\})$, then the upper bound inequality is verified, and we would naturally conclude that if we blow-up the functions v_ε at some fixed singular point $\bar x$ of v by taking the functions $\gamma_\varepsilon(x) := v_\varepsilon(\varepsilon(x - \bar x))$, then γ_ε should more and more resemble an optimal profile. In other words we expect the optimal profiles to be the asymptotic shapes of the minimizers v_ε close to the discontinuity points of v. Yet a precise statement cannot be easily formulated in the current framework.

Remark 12. – The existence of a solution of the optimal profile problem (3.5) cannot be deduced by standard semicontinuity and compactness results: indeed not only the functional $F(\cdot, \mathbb{R})$ is translation invariant, but also its natural domain is the class of all $u : \mathbb{R} \to [0,1]$ such that $\nabla u \in L^2(\mathbb{R})$, and for such functions the limits at $\pm\infty$ are not always defined (take for instance $u(x) := \sin^2 \log(1 + x^2)$).

An alternative way to find a solution of (3.5) is via rearrangement. Given a function $u : \mathbb{R} \to [0,1]$ which tends to 1 at $+\infty$ and to 0 at $-\infty$, then each sublevel $E_t := \{x : u(x) > t\}$ with $t \in (0,1)$ can be written as the disjoint union of a bounded A_t and an half line $(b_t, +\infty)$; we define the *increasing rearrangement* of u as the function u^* whose sublevels are the half-lines $E^*_t := (b_t - a_t, +\infty)$, where a_t is the measure of A_t. This rearrangement operator decreases the functional $F_1(\cdot, \mathbb{R})$ among others, that is, $F_1(u, \mathbb{R}) \ge F_1(u^*, \mathbb{R})$ (see [22]). Hence in (3.5) we can restrict to the subclass of increasing functions u such that $\nabla u \in L^2(\mathbb{R})$ and $u(0) = 1/2$ (we add this constraint to work out the translation invariance of the functional), which is compact with respect to the strong convergence; then the existence of an optimal profile follows from the (strong) semicontinuity of $F_1(\cdot, \mathbb{R})$.

4.1. The vectorial case. – The mechanical model described at the beginning of section 2 applies to mixtures of two fluids only, but can be generalized to mixtures of an arbitrary number m of fluids. In this case every configuration of the macroscopic model can be described by a function $u : \Omega \to \{\alpha_1, \dots, \alpha_m\}$ where $\alpha_1, \dots, \alpha_m$ are arbitrarily chosen affinely independent points in $\mathbb{R}^{m-1}$ (each one corresponds to one fluid in the mixture), and the energy $F(u)$ in (2.1) must be rewritten as $F(u) := \sum_{i<j} \sigma_{ij} \mathscr{H}^2(S_{ij}u)$, where $S_{ij}u$ is the interface between the phases $\{u = \alpha_i\}$ and $\{u = \alpha_j\}$, and σ_{ij} is the corresponding surface tension. If V_i is the total volume of the phase α_i, then the admissible configurations are all u which satisfy the volume constraint $\int_\Omega u = \sum V_i \alpha_i$. Notice that the corresponding minimum problem is well-posed if the coefficients σ_{ij} satisfy the following *wetting conditions* : $\sigma_{ij} + \sigma_{jk} \geq \sigma_{ik}$ for all i, j, k. In fact, if any of these inequalities does not hold, then F is not lower semicontinuous on $BV(\Omega, \{\alpha_1, \dots, \alpha_m\})$ (it is worth trying to understand why).

In the continuous model u takes values in the convex hull T of $\{\alpha_1, \dots, \alpha_m\}$, and the associated energy $E_\varepsilon(u)$ is given as in (2.2) with W a continuous positive function on $\mathbb{R}^{m-1}$ which vanishes at $\alpha_1, \dots, \alpha_m$. If we modify (2.3) and (2.4) accordingly, then Theorem 1 holds, provided we set $\sigma_{ij} ;= \operatorname{dist}(\alpha_i, \alpha_j)$ where dist is the geodesic distance on T associated with the metric $\sqrt{W(u)}\, du$ (this result was first proved in [10], see also [12] for a more general result). The proof can be achieved by modifying the argument of section 3; the optimal profile problem (3.5) becomes now

$$\bar\sigma_{ij} := \inf \Big\{ F_1(u, \mathbb{R}) \ : \ u : \mathbb{R} \to [0,1]\, , \lim_{x \to -\infty} u(x) = \alpha_i\, , \lim_{x \to +\infty} u(x) = \alpha_j \Big\} \ .$$

Notice that one of the main technical difficulties in this proof was due to the lack of a "nice" dense subclass for $BV(\Omega, \{\alpha_1, \dots, \alpha_m\})$ (cf. the proof in paragraph 3.9).

4.2. The general anisotropic case. – Anisotropic functionals of type (2.2) are obtained for instance when we replace the Dirichlet integral $\int |\nabla u|^2$ with the quadratic form $\int \langle A \nabla u; \nabla u \rangle$ where A is a symmetric positive definite $n \times n$ matrix. This is the quadratic case of a larger class of anisotropic functionals considered in [13], [26] and then in [12]. Here the surface tension σ depends on the orientation of the interface, and the Γ-limit of F_ε is given by the integral over the interface Su of $\sigma(\nu)$ where ν is the normal to Su. For every direction e the value of $\sigma(e)$ is given by an N-dimensional version of the optimal profile problem (3.5). In some cases (including the quadratic case) it is still possible to apply some rearrangement theorem to prove that the optimal profile problem reduces to a one-dimensional minimization problem.

4.3. The Ginzburg-Landau functionals. – An important variation of the Cahn-Hilliard functionals are the Ginzburg-Landau functionals, which are defined as in (2.2) for all $u : \Omega \to \mathbb{R}^2$, by taking W a continuous positive function which vanishes on the unit circle $S^1 := \{|u| = 1\}$. In this case when ε tends to 0 the function u is forced to take values closer and closer to the unit circle, and then we expect that the Γ-limit of suitable rescalings of E_ε is finite on functions $u : \Omega \to S^1$ with singularities of co-dimension 2 (in particular when Ω has dimension 2 we expect point singularities). For this reason the asymptotic behavior of these functionals differs deeply from what described in Theorem 1. However no Γ-convergence result is available so far, while a complete description of the asymptotic behavior of the minimizing sequences (under some boundary constraint) was carried out when Ω has dimension 2 in [11].

4.4. A different type of interaction energy. – Another variation of the Cahn-Hilliard functionals is obtained by replacing the Dirichlet energy $\varepsilon^2 \int |\nabla u|^2$ in (2.2) with suitable scalings of a non-local interaction energy $I_\varepsilon(u)$. In [3] and [4] is considered the case

$$I_\varepsilon(u) := \iint J_\varepsilon(x' - x) \big(u(x') - u(x) \big)^2 dx' dx \ ,$$

where $J_\varepsilon(h) := \varepsilon^{-N} J(h/\varepsilon)$ and J is a positive interaction potential in $L^1(\mathbb{R}^N)$. These kind of functionals arises as scalings of the free energy of a continuum limit of spin systems on lattices, or Ising systems. Theorem 1 is still true, but now the surface tension σ is directly defined through the optimal profile problem (3.5), and cannot be computed explicitly in term of J and W; the existence of an optimal profile and the positivity of σ have been proved in this case via rearrangement, as described in Remark 12 (see [2]).

4.5. The original proof of Theorem 1. – The proof of statement (i) and (ii) of Theorem 1 given in [23] is very simple and elegant, and works directly in the N-dimensional case; it is quite interesting to compare it with the proof given in section 3. Given $u : \Omega \to [0,1]$, one applies the inequality $a^2 + b^2 \geq 2ab$ with $a := \sqrt{W(u)/\varepsilon}$ and $b := \sqrt{\varepsilon}\, |\nabla u|$ and obtains

$$\begin{aligned} F_\varepsilon(u) &= \int_\Omega \big(\varepsilon |\nabla u|^2 + \varepsilon^{-1} W(u) \big)\, dx \\ &\geq 2 \int_\Omega \sqrt{W(u)}\, |\nabla u|\, dx = \int_\Omega \big| \nabla (H(u)) \big|\, dx \end{aligned} \tag{4.1}$$

where $H : [0,1] \to \mathbb{R}$ satisfies $\dot H = 2\sqrt{W}$. Take now a sequence (u_ε) such that $F_\varepsilon(u_\varepsilon) \leq C < +\infty$. Then (4.1) implies that the functions $H(u_\varepsilon)$ are uniformly bounded in $BV(\Omega)$, and then pre-compact in $L^1(\Omega)$, and since H admits a continuous inverse, also (u_ε) is pre-compact in $L^1(\Omega)$, and every limit point must takes values 0 or 1 a.e. by the usual argument. Assume now that (u_ε) converge to some limit u. Since u takes values 0 and 1 only, then $H(u)$ takes values $H(0)$ and $H(1)$ only, and therefore the total variation of the measure derivative $D(H(u))$ is equal to the total variation of Du multiplied by a factor $H(1) - H(0)$, which is equal to σ (cf. Theorem 1). Hence (4.1) yields

$$\liminf_{\varepsilon \to 0} F_\varepsilon(u_\varepsilon) \geq \liminf_{\varepsilon \to 0} \int_\Omega \big| D(H(u_\varepsilon)) \big| \geq \big\| D(H(u)) \big\| = \sigma \|Du\| = \sigma \mathscr{H}^{N-1}(Su)$$

(the total variation of the measure derivatives is lower semicontinuous with respect to the strong convergence of functions). This completes the proof of statements (i) and (ii), while the proof of statement (iii) is quite similar to the one sketched in paragraph 3.9.

References

[1] G. Alberti: Rank-one property for derivatives of functions with bounded variation. *Proc. Royal Soc. Edin. A* **123** (1993), 239-274.

[2] G. Alberti, G. Bellettini: A non-local anisotropic model for phase transitions, I - The optimal profile problem. *Math. Ann.* **310** (1998), 527-560.

[3] G. Alberti, G. Bellettini: A non-local anisotropic model for phase transitions - Asymptotic behavior of rescaled energies. *European J. Applied Math.* **9** (1998), 261-284.

[4] G. Alberti, G. Bellettini, M. Cassandro, E. Presutti: Surface tension in Ising systems with Kac potential. *J. Stat. Phys.* **82** (1996), 743-796.

[5] G. Alberti, G. Bouchitté, P. Seppecher: Un résultat de perturbations singulières avec la norm $H^{1/2}$. *C. R. Acad. Sci. Paris, Ser. I* **319** (1994), 333-338.

[6] G. Alberti, G. Bouchitté, P. Seppecher: Phase transitions with line tension effect. *Arch. Rational Mech. Anal.* **144** (1998), 1-49.

[7] L. Ambrosio: A compactness theorem for a new class of functions of bounded variation. *Boll. Un. Mat. It. B (7)* **3** (1989), 857-881.

[8] L. Ambrosio: Geometric evolution problems, distance function and viscosity solutions. *Calculus of variations and partial differential equations. Topics on geometrical evolution problems and degree theory (Pisa, 1996)*, 5-93, edited by G. Buttazzo et al. Springer-Verlag, Berlin-Heidelberg 2000.

[9] J.M. Ball: A version of the fundamental theorem for Young measures. *PDE's and continuum models of phase transitions*, edited by M. Rascle et al., 207-215. Lecture Notes in Physics 344. Springer-Verlag, Berlin-Heidelberg 1989.

[10] S. Baldo: Minimal interface criterion for phase transitions in mixtures of Cahn-Hilliard fluids. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **7** (1990), 289-314.

[11] F. Bethuel, H. Brezis, F. Hélein: *Ginzburg-Landau vortices.* Progress in Nonlinear Differential Equations and Their Applications 13. Birkhäuser, Basel-Boston 1994.

[12] A.C. Barroso, I. Fonseca: Anisotropic singular perturbations - The vectorial case. *Proc. Royal Soc. Edin. A* **124** (1994), 527-571.

[13] G. Bouchitté: Singular perturbations of variational problems arising from a two-phase transition model. *Appl. Math. Opt.* **21** (1990), 289-315.

[14] A. Braides, A. Defranceschi: *Homogenization of Multiple Integrals.* Oxford Lecture Series in Mathematics and its Applications 12. Oxford University Press, Oxford 1998.

[15] J.W. Cahn, J.E. Hilliard: Free energy of a non-uniform system I - Interfacial free energy. *J. Chem. Phys.* **28** (1958), 258-267.

[16] E. De Giorgi, T. Franzoni: Su un tipo di convergenza variazionale. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* **58** (1975), 842-850.

[17] G. Dal Maso: *An introduction to* Γ*-convergence.* Progress in Nonlinear Differential Equations and Their Applications 8. Birkhäuser, Basel-Boston 1993.

[18] L.C. Evans and R.F. Gariepy: *Measure theory and fine properties of functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton 1992.

[19] H. Federer: The singular sets of area minimizing rectifiable currents with codimension one and of area minimizing flat chains modulo two with arbitrary codimension. *Bull. Amer. Math. Soc.* **76** (1970), 767-771.

[20] I. Fonseca, S. Müller: Quasiconvex integrands and lower semicontinuity in L^1. *SIAM J. Math. Anal.* **23** (1992), 1081-1098.

[21] E. Giusti: *Minimal Surfaces and Functions of Bounded Variation.* Monographs in Mathematics 80. Birkhäuser, Basel-Boston 1984.

[22] B. Kawohl: *Rearrangement and Convexity of Level Sets in PDE.* Lectures Notes in Mathematics 1150. Springer-Verlga, Berlin-Heidelberg 1985.

[23] L. Modica: The gradient theory of phase transition and the minimal interface criterion. *Arch. Rational Mech. Anal.* **98** (1987), 123-142.

[24] L. Modica, S. Mortola: Un esempio di Γ-convergenza. *Boll. Un. Mat. Ital. B (5)* **14** (1977), 285-299.

[25] S. Müller: Variational models for microstructure and phase transitions. *Calculus of variations and geometric evolution problems (Cetraro, 1996)*, 85-210, edited by S. Hildebrandt and M. Struwe. Lecture Notes in Mathematics, 1713. Springer-Verlag, Berlin; Centro Internazionale Matematico Estivo (C.I.M.E.), Florence, 1999.

[26] N. Owen, P. Sternberg: Nonconvex variational problems with anisotropic perturbations. *Nonlinear Anal. T.M.A.* **16** (1991), 705-719.

[27] M. Valadier: A course on Young measures. *Rend. Istit. Mat. Univ. Trieste* **26** (1994) suppl., 349-394.

[28] B. White: A new proof of the compactness theorem for integral currents, *Comm. Math. Helv.* **64** (1989), 207-220.

Giovanni Alberti
Dipartimento di Matematica "L. Tonelli"
Università di Pisa
via Buonarroti 2, 56127 Pisa (Italy)
`alberti@dm.unipi.it`