Carnegie Mellon University — Center for Nonlinear Analysis

GAMMA CONVERGENCE AND APPLICATIONS TO PHASE TRANSITIONS

GIOVANNI LEONI

1. Liquid–Liquid Phase Transitions

Consider a fluid confined into a container $\Omega \subset \mathbb{R}^N$. Assume that the total mass of the fluid is m, so that admissible density distributions $u : \Omega \to \mathbb{R}$ satisfy the constraint $\int_\Omega u(\boldsymbol{x})\, d\boldsymbol{x} = m$. The total energy is given by the functional $u \mapsto \int_\Omega W(u(\boldsymbol{x}))\, d\boldsymbol{x}$, where $W : \mathbb{R} \to [0, \infty)$ is the energy per unit volume. Assume that W supports two phases $a < b$, that is, W is a *double-well potential*, with $\{t \in \mathbb{R} : W(t) = 0\} = \{a, b\}$. Then any density distribution u that renders the body stable in the sense of Gibbs is a minimizer of the following problem

$$(\mathcal{P}_0) \qquad \min \left\{ \int_\Omega W(u(\boldsymbol{x}))\, d\boldsymbol{x} : \int_\Omega u(\boldsymbol{x})\, d\boldsymbol{x} = m \right\}.$$

If $\mathcal{L}^N(\Omega) = 1$ and $a < m < b$, then given any measurable set $E \subset \Omega$ with

$$\mathcal{L}^N(E) = \frac{m-a}{b-a}, \tag{1}$$

the function $u = b\chi_E + a\chi_{\Omega \setminus E}$ is a solution of problem $(\mathcal{P}_0)$. Here $\mathcal{L}^N$ stands for the N-dimensional Lebesgue measure. This lack of uniqueness is due the fact that interfaces between the two phases a and b are not penalized by the total energy. The physically preferred solutions should be the ones that arise as limiting cases of a theory that penalizes interfacial energy, so it is expected that these solutions should minimize the surface area of $\partial E \cap \Omega$.

In the van der Walls–Cahn–Hilliard theory of phase transitions [CH1958], [Ro1979], [VdW1893], the energy depends not only on the density u but also on its gradient, precisely,

$$\mathcal{G}_\varepsilon(u) := \int_\Omega W(u(\boldsymbol{x}))\, d\boldsymbol{x} + \varepsilon^2 \int_\Omega |\nabla u(\boldsymbol{x})|^2\, d\boldsymbol{x}. \tag{2}$$

Note that the gradient term penalizes rapid changes of the density u, and thus it plays the role of an interfacial energy. Stable density distributions u are now solutions of the minimization problem

$$(\mathcal{P}_\varepsilon) \qquad \min \left\{ \int_\Omega W(u(\boldsymbol{x}))\, d\boldsymbol{x} + \varepsilon^2 \int_\Omega |\nabla u(\boldsymbol{x})|^2\, d\boldsymbol{x} \right\},$$

where the minimum is taken over all smooth functions u satisfying $\int_\Omega u(\boldsymbol{x})\, d\boldsymbol{x} = m$. In 1983 Gurtin [Gu1985] conjectured that the limits, as $\varepsilon \to 0$, of solutions u_ε of $(\mathcal{P}_\varepsilon)$ are solutions u_0 of $(\mathcal{P}_0)$ with minimal surface area, that is, if $u_0 = a\chi_{E_0} + b\chi_{\Omega \setminus E_0}$, then

$$\text{surface area of } E_0 \le \text{surface area of } E \tag{3}$$

for every measurable set with $\mathcal{L}^N(E) = \frac{m-a}{b-a}$. Moreover, he also conjectured that

$$\mathcal{G}_\varepsilon(u_\varepsilon) \sim \varepsilon \text{ surface area of } E_0. \tag{4}$$

Using results of Modica and Mortola [MM1977][1], this conjecture was proved independently for $N \geq 2$ by Modica [Mo1987] and by Sternberg [St1988] in the setting of Γ-convergence. The one-dimensional case $N = 1$ had been studied by Carr, Gurtin, and Slemrod in [CGS1984].

1.1. Γ-Convergence.

The notion of gamma convergence was introduced by De Giorgi in [DG1975] (see also [Br2002], [DM1993]).

Definition 1.1. *Let (Y, d) be a metric space and consider a sequence $\{\mathcal{F}_n\}$ of functions $\mathcal{F}_n : Y \to [-\infty, \infty]$. We say that $\{\mathcal{F}_n\}$* Γ-converges *to a function $\mathcal{F} : Y \to [-\infty, \infty]$ if the following properties hold:*

(i) ***(Liminf Inequality)*** *For every $y \in Y$ and every sequence $\{y_n\} \subset Y$ such that $y_n \to y$,*

$$\mathcal{F}(y) \leq \liminf_{n\to+\infty} \mathcal{F}_n(y_n). \tag{5}$$

(ii) ***(Limsup Inequality)*** *For every $y \in Y$ there exists $\{y_n\} \subset Y$ such that $y_n \to y$ and*

$$\limsup_{n\to\infty} \mathcal{F}_n(y_n) \leq \mathcal{F}(y). \tag{6}$$

The function $\mathcal{F}$ is called the Γ*-limit* of the sequence $\{\mathcal{F}_n\}$.

Exercise 1.2. *Let (Y, d) be a metric space and consider a sequence $\{\mathcal{F}_n\}$ of functions $\mathcal{F}_n : Y \to (-\infty, \infty)$. Assume that there exists*

$$\min_{z\in Y} \mathcal{F}_n(z) = \mathcal{F}_n(y_n),$$

that $\{\mathcal{F}_n\}$ Γ-converges to $\mathcal{F}$, and that $y_n \to y$ for some $y \in Y$. Prove that there exists $\min_{z\in Y} \mathcal{F}(z)$ *and that*

$$\mathcal{F}(y) = \min_{z\in Y} \mathcal{F}(z) = \lim_{n\to\infty} \min_{y\in Y} \mathcal{F}_n(y).$$

In applications the main challenges are

- Finding the appropriate rescaling of $\mathcal{F}_n$ and an appropriate metric d. These usually follow by studying equibounded sequences and by a compactness argument.
- Identifying the Γ-limit $\mathcal{F}$.
- Proving (i) and (ii).

[1] In [MM1977] Modica and Mortola studied the Γ-convergence of the sequence of functionals

$$\int_{\mathbb{R}^N} \frac{1}{\varepsilon} \sin^2(\pi u(\boldsymbol{x}))\, d\boldsymbol{x} + \varepsilon \int_{\mathbb{R}^N} |\nabla u(\boldsymbol{x})|^2\, d\boldsymbol{x}$$

with respect to the convergence in $L^1(\mathbb{R}^N)$.

Exercise 1.3. *Take $Y = L^2(\Omega)$ and assume that*

$$W(t) = |t-a|\,|b-t|.$$

Study the Γ*-convergence of the family of functionals*

$$\mathcal{G}_\varepsilon(u) := \begin{cases} \displaystyle\int_\Omega \left(W(u) + \varepsilon^2|\nabla u|^2\right) d\boldsymbol{x} & \text{if } u \in W^{1,2}(\Omega) \text{ and } \int_\Omega u\, d\boldsymbol{x} = m, \\ \infty & \text{otherwise in } L^2(\Omega). \end{cases}$$

Exercise 1.4. *Take $Y = L^2(\Omega)$ and assume that*

$$W(t) = |t-a|\,|b-t|.$$

Study the Γ*-convergence of the family of functionals*

$$\mathcal{G}_\varepsilon(u) := \begin{cases} \displaystyle\int_\Omega \left(W(u) + \varepsilon^2|\nabla u|^2\right) d\boldsymbol{x} & \text{if } u \in W^{1,2}(\Omega) \text{ and } \int_\Omega u\, d\boldsymbol{x} = m, \\ \infty & \text{otherwise in } L^2(\Omega) \end{cases}$$

with respect to weak convergence in $L^2(\Omega)$.[2]

1.2. **Compactness.** In view of (4), for $\varepsilon > 0$ we consider the rescaled functional

$$\mathcal{F}_\varepsilon : W^{1,2}(\Omega) \to [0, \infty]$$

defined by

$$\mathcal{F}_\varepsilon(u) := \int_\Omega \left(\frac{1}{\varepsilon}W(u) + \varepsilon|\nabla u|^2\right) d\boldsymbol{x}, \tag{7}$$

where the double well potential $W : \mathbb{R} \to [0, \infty)$ satisfies the following hypotheses:

- (H_1) W is continuous, $W(t) = 0$ if and only if $t \in \{a, b\}$ for some $a, b \in \mathbb{R}$ with $a < b$.
- (H_2) There exist $L > 0$ and $T > 0$ such that

  $$W(t) \geq L\,|t|.$$

  for all $t \in \mathbb{R}$ with $|t| \geq T$.

Definition 1.5. *Let $\Omega \subset \mathbb{R}^N$ be an open set. We define the* space of functions of bounded variation *$BV(\Omega)$ as the space of all functions $u \in L^1(\Omega)$ whose distributional first-order partial derivatives are finite signed Radon measures; that is, for all $i = 1, \ldots, N$ there exists a finite signed measure $\lambda_i : \mathcal{B}(\Omega) \to \mathbb{R}$ such that*

$$\int_\Omega u \frac{\partial \phi}{\partial x_i}\, d\boldsymbol{x} = -\int_\Omega \phi\, d\lambda_i \tag{8}$$

for all $\phi \in C_c^\infty(\Omega)$. The measure λ_i is called the weak, *or* distributional, partial derivative *of u with respect to x_i and is denoted $D_i u$.*

See [AFP2000], [Le2009] for more information about functions of bounded variation.

[2] Which means that in the definition of Γ-convegence, you should replace $y_n \to y$ with $u_n \rightharpoonup u$ in $L^2(\Omega)$.

Theorem 1.6 (Compactness)**.** *Let $\Omega \subset \mathbb{R}^N$ be an open set with finite measure. Assume that the double-well potential W satisfies conditions (H_1) and (H_2). Let $\varepsilon_n \to 0^+$ and let $\{u_n\} \subset W^{1,2}(\Omega)$ be such that*

$$M := \sup_n \mathcal{F}_{\varepsilon_n}(u_n) < \infty. \tag{9}$$

Then there exist a subsequence $\{u_{n_k}\}$ of $\{u_n\}$ and $u \in BV(\Omega; \{a,b\})$ such that

$$u_{n_k} \to u \text{ for in } L^1(\Omega).$$

Proof. We begin by showing that $\{u_n\}$ is bounded in $L^1(\Omega)$ and equi-integrable. By (9) and (H_2),

$$\int_{\{|u_n|\geq T\}} |u_n|\, d\boldsymbol{x} \leq \frac{1}{L}\int_{\{|u_n|\geq T\}} W(u_n(x))\, d\boldsymbol{x} \leq \frac{M}{L}\varepsilon_n. \tag{10}$$

Since Ω has finite measure,

$$\begin{aligned} L\int_\Omega |u_n|\, d\boldsymbol{x} &= L\int_{\{|u_n|\geq T\}} |u_n|\, d\boldsymbol{x} + L\int_{\{|u_n|< T\}} |u_n|\, d\boldsymbol{x} \\ &\leq \int_{\{|u_n|\geq T\}} W(u_n)\, d\boldsymbol{x} + LT\mathcal{L}^N(\Omega) \leq M\varepsilon_n + LT\mathcal{L}^N(\Omega). \end{aligned}$$

Thus, $\{u_n\}$ is bounded in $L^1(\Omega)$. We claim that $\{u_n\}$ is equi-integrable. Indeed, let $\gamma > 0$ be fixed and find $N_\varepsilon \in \mathbb{N}$ so large that $\frac{M}{L}\varepsilon_n \leq \frac{1}{2}\gamma$ for all $n \geq N_\varepsilon$. Then, by (10),

$$\int_{\{|u_n|\geq T\}} |u_n|\, d\boldsymbol{x} \leq \frac{1}{2}\gamma \tag{11}$$

for all $n \geq N_\varepsilon$, while if $E \subset \Omega$ is measurable,

$$\int_{E\cap\{|u_n|< T\}} |u_n|\, d\boldsymbol{x} \leq T\mathcal{L}^N(E) \leq \frac{1}{2}\gamma, \tag{12}$$

provided that $\mathcal{L}^N(E) \leq \frac{1}{2T}\gamma$. It follows from (11) and (12) that for every measurable set $E \subset \Omega$ with $\mathcal{L}^N(E) \leq \frac{1}{2T}\gamma$,

$$\int_E |u_n|\, d\boldsymbol{x} = \int_{E\cap\{|u_n|> T\}} |u_n|\, d\boldsymbol{x} + \int_{E\cap\{|u_n|< T\}} |u_n|\, d\boldsymbol{x} \leq \frac{1}{2}\gamma + \frac{1}{2}\gamma$$

for all $n \geq N_\varepsilon$. Finally, since the finite family $\{u_1, \ldots, u_{N_\varepsilon}\}$ is equi-integrable (exercise), there exists $\delta_1 > 0$ such that

$$\int_E |u_n|\, d\boldsymbol{x} \leq \gamma$$

for all $n \leq N_\varepsilon$ and for every measurable set $E \subset \Omega$ with $\mathcal{L}^N(E) \leq \delta_1$. It suffices to take $\delta := \min\left\{\delta_1, \frac{1}{2T}\gamma\right\}$. This implies that $\{u_n\}$ is equi-integrable.

Since Ω has finite measure, in view of the Vitali's convergence theorem and of the Egoroff convergence theorem, to obtain strong convergence of a subsequence, it suffices to prove pointwise convergence of a subsequence.

For $K > 0$ define

$$W_1(t) := \min\{W(t), K\}, \quad t \in \mathbb{R} \tag{13}$$

and

$$f(t) := 2\int_a^t \sqrt{W_1(s)}\,ds, \quad t \in \mathbb{R}. \tag{14}$$

Since $0 \le W_1 \le W$, for every $n \in \mathbb{N}$, we have

$$\mathcal{F}_{\varepsilon_n}(u_n) \ge 2\int_\Omega \sqrt{W_1(u_n(\boldsymbol{x}))}|\nabla u_n(\boldsymbol{x})|\,d\boldsymbol{x} = \int_\Omega |\nabla (f\circ u_n)(\boldsymbol{x})|\,d\boldsymbol{x}. \tag{15}$$

Note that the function f is Lipschitz continuous and we are using the chain rule in $W^{1,1}(\Omega)$ (see Theorem 6.8) Then by (9),

$$\sup_n \int_\Omega |\nabla (f\circ u_n)|\,d\boldsymbol{x} \le M. \tag{16}$$

Moreover, since $\operatorname{Lip} f \le 2\sqrt{K}$, and $f(a) = 0$,

$$|f(u_n(\boldsymbol{x}))| = |f(u_n(\boldsymbol{x})) - f(a)| \le 2\sqrt{K}\,|u_n(\boldsymbol{x}) - a|$$

for $\mathcal{L}^N$ a.e. $\boldsymbol{x} \in \Omega$ and for all $n \in \mathbb{N}$. Since $\{u_n\}$ is bounded in $L^1(\Omega)$, it follows that the sequence $\{f\circ u_n\}$ is bounded in $L^1(\Omega)$. By the Rellich–Kondrachov theorem, there exist a subsequence of $\{u_n\}$ (not relabeled) and a function $w \in BV(\Omega)$ such that

$$w_n := f\circ u_n \to w \text{ in } L^1_{\text{loc}}(\Omega).$$

By taking a further subsequence, if necessary, without loss of generality, we may assume that $w_n(\boldsymbol{x}) \to w(\boldsymbol{x})$ and that $W(u_n(\boldsymbol{x})) \to 0$ for $\mathcal{L}^N$ a.e. $\boldsymbol{x} \in \Omega$. Since the function $W_1(t) > 0$ for all $t \ne a, b$, it follows from (14) that the function f is strictly increasing and continuous. Thus, its inverse f^{-1} is continuous and

$$u_{n_k}(\boldsymbol{x}) = f^{-1}(w_k(\boldsymbol{x})) \to f^{-1}(w(\boldsymbol{x})) =: u(\boldsymbol{x})$$

for $\mathcal{L}^N$ a.e. $\boldsymbol{x} \in \Omega$. It follows by (H_1) and the fact that $W(u_{n_k}(\boldsymbol{x})) \to 0$ for $\mathcal{L}^N$ a.e. $\boldsymbol{x} \in \Omega$, that $u(\boldsymbol{x}) \in \{a, b\}$ for $\mathcal{L}^N$ a.e. $\boldsymbol{x} \in \Omega$.

In turn, $w(\boldsymbol{x}) \in \{f(a), f(b)\} = \{0, f(b)\}$ for $\mathcal{L}^N$ a.e. $\boldsymbol{x} \in \Omega$, and so we may write

$$w = f(b)\,\chi_E \tag{17}$$

for a set $E \subset \Omega$. Since $w \in BV(\Omega)$ and Ω has finite measure, we have that $\chi_E \in BV(\Omega)$. Hence,

$$u = b\chi_E + a(1 - \chi_E) \tag{18}$$

belongs to $BV(\Omega)$. □

Remark 1.7. *Theorem 1.6 was proved by Modica* [Mo1987] *and by Sternberg* [St1988] *under the stronger assumption that*

$$\frac{1}{c}|t|^p \le W(t) \le c\,|t|^p$$

for all $|t| \geq T$ and for some $c > 0$ and $p \geq 2$. The weaker hypothesis (H_2) is due to Fonseca and Tartar [FT1989].

1.3. **Liminf Inequality.** In view of the previous theorem, the metric convergence in the definition of Γ-convergence should be $L^1(\Omega)$. Thus, we extend $\mathcal{F}_\varepsilon$ to $L^1(\Omega)$ by setting

$$\mathcal{F}_\varepsilon(u) := \begin{cases} \int_\Omega \left(\frac{1}{\varepsilon}W(u) + \varepsilon|\nabla u|^2\right) d\boldsymbol{x} & \text{if } u \in W^{1,2}(\Omega) \text{ and } \int_\Omega u\, d\boldsymbol{x} = m, \\ \infty & \text{otherwise in } L^1(\Omega). \end{cases} \tag{19}$$

Let $\varepsilon_n \to 0^+$. Under appropriate hypotheses on W and Ω, we will show that the sequence of functionals $\{\mathcal{F}_{\varepsilon_n}\}$ Γ-converges to the functional

$$\mathcal{F}(u) := \begin{cases} c_W \operatorname{P}(E,\Omega) & \text{if } u \in BV(\Omega;\{a,b\}) \text{ and } \int_\Omega u\, d\boldsymbol{x} = m, \\ \infty & \text{otherwise in } L^1(\Omega), \end{cases} \tag{20}$$

where

$$c_W := 2\int_a^b \sqrt{W(t)}\, dt \tag{21}$$

and $E := \{\boldsymbol{x} \in \Omega : u(\boldsymbol{x}) = b\}$.

For $u \in BV(\Omega)$ we set

$$Du := (D_1 u, \ldots, D_N u).$$

Thus, if $u \in BV(\Omega)$, then $Du \in \mathcal{M}_\mathrm{b}(\Omega;\mathbb{R}^N)$, and since $\mathcal{M}_\mathrm{b}(\Omega;\mathbb{R}^N)$ may be identified with the dual of $C_0(\Omega;\mathbb{R}^N)$, we have that

$$\begin{aligned} |Du|(\Omega) &:= \|Du\|_{\mathcal{M}_\mathrm{b}(\Omega;\mathbb{R}^N)} \\ &= \sup\left\{\sum_{i=1}^N \int_\Omega \Phi_i dD_i u : \Phi \in C_0(\Omega;\mathbb{R}^N),\ \|\Phi\|_{C_0(\Omega;\mathbb{R}^N)} \leq 1\right\} < \infty. \end{aligned}$$

Definition 1.8. *Let $\Omega \subset \mathbb{R}^N$ be an open set and let $u \in L^1_{\mathrm{loc}}(\Omega)$. The* variation *of u in Ω is defined by*

$$V(u,\Omega) := \sup\left\{\sum_{i=1}^N \int_\Omega \frac{\partial \Phi_i}{\partial x_i} u\, d\boldsymbol{x} : \Phi \in C_c^\infty(\Omega;\mathbb{R}^N),\ \|\Phi\|_{C_0(\Omega;\mathbb{R}^N)} \leq 1\right\}.$$

Exercise 1.9. *Let $\Omega \subset \mathbb{R}^N$ be an open set and let $u \in L^1_{\mathrm{loc}}(\Omega)$. Prove the following.*

(i) *If the distributional gradient Du of u belongs to $\mathcal{M}_\mathrm{b}(\Omega;\mathbb{R}^N)$, then*

$$|Du|(\Omega) = V(u,\Omega).$$

(ii) *If $V(u,\Omega) < \infty$, then the distributional gradient Du of u belongs to $\mathcal{M}_\mathrm{b}(\Omega;\mathbb{R}^N)$. In particular, if $u \in L^1(\Omega)$, then u belongs to $BV(\Omega)$ if and only if $V(u,\Omega) < \infty$. Hint: Use the Riesz representation theorem in $C_0(\Omega;\mathbb{R}^N)$.*

(iii) *If* $\{u_n\} \subset L^1_{\text{loc}}(\Omega)$ *is a sequence of functions converging to* u *in* $L^1_{\text{loc}}(\Omega)$*, then*

$$V(u,\Omega) \le \liminf_{n\to\infty} V(u_n,\Omega).$$

The previous example shows that characteristic functions of smooth sets belong to $BV(\Omega)$*. More generally, we have the following.*

Definition 1.10. *Let* $E \subset \mathbb{R}^N$ *be a Lebesgue measurable set and let* $\Omega \subset \mathbb{R}^N$ *be an open set. The* perimeter *of* E *in* Ω*, denoted* $\mathrm{P}(E,\Omega)$*, is the variation of* χ_E *in* Ω*, that is,*

$$\begin{aligned}\mathrm{P}(E,\Omega) &:= V(\chi_E,\Omega)\\ &= \sup\left\{\sum_{i=1}^N \int_E \frac{\partial \Phi_i}{\partial x_i}\,d\boldsymbol{x} :\ \Phi \in C_c^\infty\left(\Omega;\mathbb{R}^N\right),\ \|\Phi\|_{C_0(\Omega;\mathbb{R}^N)} \le 1\right\}.\end{aligned}$$

The set E *is said to have* finite perimeter in Ω *if* $\mathrm{P}(E,\Omega) < \infty$*.*

If $\Omega = \mathbb{R}^N$, we write

$$\mathrm{P}(E) := \mathrm{P}\left(E,\mathbb{R}^N\right).$$

Remark 1.11. *In view of Exercise 1.9, if* $\Omega \subset \mathbb{R}^N$ *is an open set and* $E \subset \mathbb{R}^N$ *is a Lebesgue measurable set with* $\mathcal{L}^N(E\cap\Omega) < \infty$*, then* χ_E *belongs to* $BV(\Omega)$ *if and only if* $\mathrm{P}(E,\Omega) < \infty$*.*

We are now ready to study the Γ-convergence of the sequence of functionals (19).

Theorem 1.12 (Liminf inequality)**.** *Let* $\Omega \subset \mathbb{R}^N$ *be an open set with finite measure. Assume that the double-well potential* W *satisfies conditions* (H_1) *and* (H_2)*. Let* $\varepsilon_n \to 0^+$*. Let* $\varepsilon_n \to 0^+$ *and let* $\{u_n\} \subset L^1(\Omega)$ *be such that* $u_n \to u$ *in* $L^1(\Omega)$*. Then*

$$\liminf_{n\to\infty} \mathcal{F}_{\varepsilon_n}(u_n) \ge \mathcal{F}(u), \tag{22}$$

where $\mathcal{F}_n$ *and* $\mathcal{F}$ *are the functionals defined in* (19) *and* (20)*, respectively.*

Proof. Consider a sequence $\{u_n\} \subset L^1(\Omega)$ such that $u_n \to u$ in $L^1(\Omega)$ for some $u \in L^1(\Omega)$. If

$$\liminf_{n\to\infty} \mathcal{F}_{\varepsilon_n}(u_n) = \infty,$$

then there is nothing to prove, thus we assume that

$$\liminf_{n\to\infty} \mathcal{F}_{\varepsilon_n}(u_n) < \infty. \tag{23}$$

Let $\{\varepsilon_{n_k}\}$ be a subsequence of $\{\varepsilon_n\}$ such that

$$\liminf_{n\to\infty} \mathcal{F}_{\varepsilon_n}(u_n) = \lim_{k\to\infty} \mathcal{F}_{\varepsilon_{n_k}}(u_{n_k}) < \infty.$$

Then $\mathcal{F}_{\varepsilon_{n_k}}(u_{n_k}) < \infty$ for all k sufficiently large. Hence, $u_{n_k} \in W^{1,2}(\Omega)$ for all k sufficiently large. By Theorem 1.6, $u \in BV(\Omega;\{a,b\})$. Finally, by extracting a further subsequence, not relabelled, we can assume that $u_n(\boldsymbol{x}) \to u(\boldsymbol{x})$ for $\mathcal{L}^N$ a.e. $\boldsymbol{x} \in \Omega$.

Hence, in what follow, without loss of generality, we will assume that (23) holds, that $\{u_n\} \subset W^{1,2}(\Omega)$, that $u \in BV(\Omega;\{a,b\})$, that $\liminf_{n\to+\infty} \mathcal{F}_{\varepsilon_n}(u_n)$ is actually a limit, and that $\{u_n\}$ converges to u in $L^1(\Omega)$ and pointwise $\mathcal{L}^N$ a.e. in Ω.

Step 1: We begin by truncating the sequence $\{u_n\}$. Consider the Lipschitz function

$$h(t) := \begin{cases} b & \text{if } t \geq b, \\ t & \text{if } a < t < b, \\ a & \text{if } t \leq a. \end{cases}$$

Note that

$$h'(t) = \begin{cases} 0 & \text{if } t > b, \\ 1 & \text{if } a < t < b, \\ 0 & \text{if } t < a. \end{cases}$$

By the chain rule in Sobolev spaces, the functions $v_n := h \circ u_n$ are still in $W^{1,2}(\Omega)$ and for $\mathcal{L}^N$ a.e. $\boldsymbol{x} \in \Omega$,

$$\nabla (h \circ u_n)(\boldsymbol{x}) = \begin{cases} \nabla u_n(\boldsymbol{x}) & \text{if } a < u_n(\boldsymbol{x}) < b, \\ 0 & \text{otherwise,} \end{cases}$$

where we used the fact that $\nabla u_n(\boldsymbol{x}) = 0$ for $\mathcal{L}^N$ a.e. $\boldsymbol{x} \in \Omega$ such that $u_n(\boldsymbol{x}) = a$ or $u_n(\boldsymbol{x}) = b$. Hence, $|\nabla v_n| \leq |\nabla u_n|$. Moreover, since $W \geq 0$ and $W(a) = W(b) = 0$, if $a < u_n(\boldsymbol{x}) < b$, we have that $W(v_n(\boldsymbol{x})) = W(u_n(\boldsymbol{x}))$, otherwise $W(v_n(\boldsymbol{x})) = 0 \leq W(u_n(\boldsymbol{x}))$. Hence,

$$\int_\Omega \left(\frac{1}{\varepsilon} W(v_n) + \varepsilon |\nabla v_n|^2 \right) d\boldsymbol{x} \leq \int_\Omega \left(\frac{1}{\varepsilon} W(u_n) + \varepsilon |\nabla u_n|^2 \right) d\boldsymbol{x}.$$

Finally, since $u \in BV(\Omega; \{a, b\})$, we have that $h \circ u = u$ and so using the fact that $\operatorname{Lip} h \leq 1$,

$$|h(u_n(\boldsymbol{x})) - u(\boldsymbol{x})| = |h(u_n(\boldsymbol{x})) - h(u(\boldsymbol{x}))| \leq |u_n(\boldsymbol{x}) - u(\boldsymbol{x})|,$$

which shows that $v_n \to u$ in $L^1(\Omega)$ and $v_n(\boldsymbol{x}) \to v(\boldsymbol{x})$ for $\mathcal{L}^N$ a.e. $\boldsymbol{x} \in \Omega$.

Step 2: Consider the function W_1 and f defined in (13) and (14), respectively, where

$$K := \max_{t \in [a,b]} W(t).$$

Note that $W_1(t) = W(t)$ for all $t \in [a, b]$. As in the proof of Theorem 1.6 we have that

$$\begin{aligned} \mathcal{F}_{\varepsilon_n}(v_n) &\geq 2 \int_\Omega \sqrt{W_1(v_n(\boldsymbol{x}))} |\nabla v_n(\boldsymbol{x})| \, d\boldsymbol{x} \\ &= \int_\Omega |\nabla (f \circ v_n)(\boldsymbol{x})| \, d\boldsymbol{x}, \end{aligned}$$

that $f \circ v_n \to f \circ u$ in $L^1(\Omega)$ and $(f \circ v_n)(\boldsymbol{x}) \to (f \circ u)(\boldsymbol{x})$ for $\mathcal{L}^N$ a.e. $\boldsymbol{x} \in \Omega$. By the lower semicontinuity of the seminorms in BV, we have that

$$\begin{aligned} \liminf_{n \to \infty} \mathcal{F}_{\varepsilon_n}(v_n) &\geq \liminf_{n \to \infty} \int_\Omega |\nabla (f \circ v_n)(\boldsymbol{x})| \, d\boldsymbol{x} = \liminf_{n \to \infty} |D(f \circ v_n)|(\Omega) \\ &\geq |D(f \circ u)|(\Omega). \end{aligned}$$

Write

$$u = b\chi_E + a(1 - \chi_E),$$

for a set $E \subset \Omega$, so that

$$f \circ u = f(b)\,\chi_E + f(a)(1-\chi_E) = f(b)\,\chi_E.$$

Hence,

$$\begin{aligned}|D(f \circ u)|(\Omega) &= |D(f(b)\,\chi_E)|(\Omega) = f(b)\,|D(\chi_E)|(\Omega) \\ &= c_W\,\mathrm{P}(E,\Omega).\end{aligned}$$

This concludes the proof. □

1.4. **Limsup Inequality.** To prove the limsup inequality, we will need stronger assumptions on the set Ω.

Theorem 1.13 (Limsup inequality)**.** *Let $\Omega \subset \mathbb{R}^N$ be an open bounded set with Lipschitz boundary. Assume that the double-well potential W satisfies condition (H_1). Then for every $u \in L^1(\Omega)$ there exists a sequence $\{u_n\} \subset L^1(\Omega)$ be such that $u_n \to u$ in $L^1(\Omega)$. Then*

$$\limsup_{n\to\infty} \mathcal{F}_{\varepsilon_n}(u_n) \le \mathcal{F}(u), \tag{24}$$

where $\mathcal{F}_n$ and $\mathcal{F}$ are the functionals defined in (19) *and* (20)*, respectively.*

Proof. If $\mathcal{F}(u) = \infty$, then we can take $u_n := u$ for all n. Thus, assume that $\mathcal{F}(u) < \infty$, so that $u \in BV(\Omega;\{a,b\})$ and $\int_\Omega u\,d\boldsymbol{x} = m$. Write

$$u = b\chi_E + a(1-\chi_E).$$

Step 1: Sketch of the proof for $\mathbf{N = 1}$: Assume first that $N = 1$ and that $\Omega = (-\ell,\ell)$ and that the function u takes the form

$$g_0(t) := \begin{cases} a & \text{if } t < 0, \\ b & \text{if } t \ge 0. \end{cases} \tag{25}$$

We would like to approximate g_0 with a Lipschitz function g_ε such that $g_\varepsilon(-\ell) = a$, $g_\varepsilon(\ell) = b$ and minimizing the one dimensional functional

$$\int_{-\ell}^{\ell} \left(\frac{1}{\varepsilon}W(g) + \varepsilon|g'|^2\right)\,dt.$$

The Euler–Lagrange equation of this functional is $2\varepsilon^2 g'' = W'(g)$. To see this, assume that g is a minimizer and take $h \in C_c^2(-\ell,\ell)$. and consider the function $g + sh$. Then

$$\mathcal{F}_{\varepsilon_n}(g+sh) \ge \mathcal{F}_{\varepsilon_n}(g)$$

for all s and so the function

$$s \mapsto \mathcal{F}_{\varepsilon_n}(g+sh)$$

has a minimum at $s=0$. In turn,

$$0=\frac{d\mathcal{F}_{\varepsilon_n}}{ds}(g+sh)\bigg|_{s=0}=\int_{-\ell}^{\ell}\left(\frac{1}{\varepsilon}W'(g+sh)h+\varepsilon 2\left(g'+sh'\right)h'\right)\,dt\bigg|_{s=0}$$
$$=\int_{-\ell}^{\ell}\left(\frac{1}{\varepsilon}W'(g)h+\varepsilon 2g'h'\right)\,dt=\int_{-\ell}^{\ell}\left(\frac{1}{\varepsilon}W'(g)-\varepsilon 2g''\right)h\,dt$$

for all $h\in C_c^2\left(-\ell,\ell\right)$. A density argument gives $2\varepsilon^2 g''=W'\left(g\right)$. Multiplying by g' and integrating gives

$$\varepsilon^2|g'|^2=c_\varepsilon+W(g).$$

The constant c_ε cannot be zero. Indeed, if $c_\varepsilon=0$ and if $g\left(t_0\right)=a$ or $g\left(t_0\right)=b$, then since $W\left(a\right)=W\left(b\right)=0$, then g would be a constant. On the other hand we need g to go from a to b as fast as possible. We take $c_\varepsilon=\varepsilon$. Hence, we define

$$\varphi_\varepsilon\left(z\right):=\int_a^z\frac{\varepsilon}{\sqrt{\varepsilon+W\left(s\right)}}\,ds$$

if $a\leq z\leq b$. Since φ_ε is strictly increasing, it has an inverse $\varphi_\varepsilon^{-1}:[0,\varphi_\varepsilon\left(b\right)]\to[a,b]$. Moreover, $\varphi_\varepsilon^{-1}\left(0\right)=a$, $\varphi_\varepsilon^{-1}\left(\varphi_\varepsilon\left(b\right)\right)=b$ and

$$\frac{d\varphi_\varepsilon^{-1}}{dt}\left(t\right)=\frac{1}{\varphi_\varepsilon'\left(\varphi_\varepsilon^{-1}\left(t\right)\right)}=\frac{\sqrt{\varepsilon+W\left(\varphi_\varepsilon^{-1}\left(t\right)\right)}}{\varepsilon},$$

which is what we wanted. Finally, since $W\geq 0$,

$$\varphi_\varepsilon\left(b\right)=\int_a^b\frac{\varepsilon}{\sqrt{\varepsilon+W\left(s\right)}}\,ds\leq\int_a^b\frac{\varepsilon}{\sqrt{\varepsilon}}\,ds=\left(b-a\right)\varepsilon^{1/2}.$$

Extend $\varphi_\varepsilon^{-1}\left(t\right)$ to be a for $t<0$ and b for $t>\varphi_\varepsilon\left(b\right)$. The function φ_ε^{-1} has all the desired properties, except the mass constraint.

Step 2: Assume that $N\geq 2$ and that that E is an open set with ∂E a nonempty compact hypersurface of class C^2 and that E meets the boundary of Ω transversally, that is, $\mathcal{H}^{N-1}\left(\partial E\cap\partial\Omega\right)=0$. Here $\mathcal{H}^{N-1}$ is the $(N-1)$-dimensional Hausdorff measure. Let's rewrite u as follows

$$u\left(\boldsymbol{x}\right)=g_0\left(\mathrm{d}_E\left(\boldsymbol{x}\right)\right),$$

where g_0 is the function (25) and d_E is the *signed distance* of E, that is,

$$\mathrm{d}_E\left(\boldsymbol{x}\right):=\begin{cases}\operatorname{dist}\left(\boldsymbol{x},\partial E\right) & \text{if } \boldsymbol{x}\in E,\\ -\operatorname{dist}\left(\boldsymbol{x},\partial E\right) & \text{if } \boldsymbol{x}\in\mathbb{R}^N\setminus E.\end{cases}$$

The idea is now to consider the function

$$v_\varepsilon\left(\boldsymbol{x}\right)=\varphi_\varepsilon^{-1}\left(\mathrm{d}_E\left(\boldsymbol{x}\right)\right).$$

The problem is that v_ε does not satisfy the mass constraint $\int_\Omega v_\varepsilon(\boldsymbol{x})\,d\boldsymbol{x} = m$. To solve this problem, observe that for every $t\in\mathbb{R}$, $\varphi_\varepsilon^{-1}(t)\le g_0(t)$, while $g_0(t)\le\varphi_\varepsilon^{-1}(t+\varphi_\varepsilon(b))$. Hence,

$$\int_\Omega \varphi_\varepsilon^{-1}(\mathrm{d}_E(\boldsymbol{x}))\,d\boldsymbol{x} \le \int_\Omega g_0(\mathrm{d}_E(\boldsymbol{x}))\,d\boldsymbol{x} = \int_\Omega u(\boldsymbol{x})\,d\boldsymbol{x} = m,$$

$$\int_\Omega \varphi_\varepsilon^{-1}(\mathrm{d}_E(\boldsymbol{x})+\varphi_\varepsilon(b))\,d\boldsymbol{x} \ge \int_\Omega g_0(\mathrm{d}_E(\boldsymbol{x}))\,d\boldsymbol{x} = \int_\Omega u(\boldsymbol{x})\,d\boldsymbol{x} = m.$$

By the continuity of the function

$$s\in[0,\varphi_\varepsilon(b)]\mapsto \int_\Omega \varphi_\varepsilon^{-1}(\mathrm{d}_E(\boldsymbol{x})+s)\,d\boldsymbol{x}$$

and the intermediate value theorem, we may find $s_\varepsilon\in[0,\varphi_\varepsilon(b)]$ such that

$$\int_\Omega \varphi_\varepsilon^{-1}(\mathrm{d}_E(\boldsymbol{x})+s_\varepsilon)\,d\boldsymbol{x} = m.$$

Hence, we can now define $g_\varepsilon(t) := \varphi_\varepsilon^{-1}(t+s_\varepsilon)$ and

$$u_\varepsilon(\boldsymbol{x}) := g_\varepsilon(\mathrm{d}_E(\boldsymbol{x})). \tag{26}$$

The function d_E is a Lipschitz function with $|\nabla\,\mathrm{d}_E(\boldsymbol{x})|=1$ for $\mathcal{L}^N$ a.e. $\boldsymbol{x}\in\mathbb{R}^N$ (see Propositions 5.2 and 5.4). Moreover, using the fact that $\mathcal{H}^{N-1}(\partial E\cap\partial\Omega)=0$ and that ∂E is of class C^2, we have that (see Lemma 5.8),

$$\lim_{r\to 0}\mathcal{H}^{N-1}(\{\boldsymbol{x}\in\Omega:\ \mathrm{d}_E(\boldsymbol{x})=r\}) = \mathcal{H}^{N-1}(\Omega\cap\partial E).$$

Hence, by the coarea formula for Lipschitz functions (see Theorem 1.14)), we have

$$\begin{aligned}
&\int_\Omega\left(\frac{1}{\varepsilon}W(u_\varepsilon(\boldsymbol{x}))+\varepsilon|\nabla u_\varepsilon(\boldsymbol{x})|^2\right)d\boldsymbol{x}\\
&\quad=\int_\Omega\left(\frac{1}{\varepsilon}W(g_\varepsilon(\mathrm{d}_E(\boldsymbol{x})))+\varepsilon|g_\varepsilon'(\mathrm{d}_E(\boldsymbol{x}))|^2\right)d\boldsymbol{x}\\
&\quad=\int_{-s_\varepsilon}^{\varphi_\varepsilon(b)-s_\varepsilon}\left(\frac{1}{\varepsilon}W(g_\varepsilon(r))+\varepsilon|g_\varepsilon'(r)|^2\right)\mathcal{H}^{N-1}(\{\boldsymbol{x}\in\Omega:\ \mathrm{d}_E(\boldsymbol{x})=r\})\,dr \qquad (27)\\
&\quad\le\sup_{|t|\le\varphi_\varepsilon(b)}\mathcal{H}^{N-1}(\{\boldsymbol{x}\in\Omega:\ \mathrm{d}_E(\boldsymbol{x})=t\})\int_0^{\varphi_\varepsilon(b)}\left(\frac{1}{\varepsilon}W(\varphi_\varepsilon^{-1}(t))+\varepsilon\left|\frac{d\varphi_\varepsilon^{-1}}{dt}(t)\right|^2\right)dt\\
&\quad\le\sup_{|t|\le\varphi_\varepsilon(b)}\mathcal{H}^{N-1}(\{\boldsymbol{x}\in\Omega:\ \mathrm{d}_E(\boldsymbol{x})=t\})\int_0^{\varphi_\varepsilon(b)}\left(\frac{\varepsilon+W(\varphi_\varepsilon^{-1}(t))}{\varepsilon}+\varepsilon\left|\frac{d\varphi_\varepsilon^{-1}}{dt}(t)\right|^2\right)dt\\
&\quad\le\sup_{|t|\le\varphi_\varepsilon(b)}\mathcal{H}^{N-1}(\{\boldsymbol{x}\in\Omega:\ \mathrm{d}_E(\boldsymbol{x})=t\})\int_0^{\varphi_\varepsilon(b)}2\sqrt{\varepsilon+W(\varphi_\varepsilon^{-1}(t))}\frac{d\varphi_\varepsilon^{-1}}{dt}(t)\,dt\\
&\quad=\sup_{|t|\le\varphi_\varepsilon(b)}\mathcal{H}^{N-1}(\{\boldsymbol{x}\in\Omega:\ \mathrm{d}_E(\boldsymbol{x})=t\})\int_a^b 2\sqrt{\varepsilon+W(s)}\,ds.
\end{aligned}$$

In turn,

$$\limsup_{\varepsilon\to 0^+}\int_\Omega \left(\frac{1}{\varepsilon}W(u_\varepsilon)+\varepsilon|\nabla u_\varepsilon|^2\right)\,d\boldsymbol{x}\le \mathcal{H}^{N-1}\left(\Omega\cap\partial E\right)\int_a^b 2\sqrt{W\left(s\right)}\,ds.$$

Finally, it remains to show that $u_\varepsilon\to u$ in $L^1\left(\Omega\right)$. Again by the coarea formula for Lipschitz functions and the fact that $|\nabla\,\mathrm{d}_E\left(\boldsymbol{x}\right)|=1$,

$$\begin{aligned}\int_\Omega |u_\varepsilon\left(\boldsymbol{x}\right)-u\left(\boldsymbol{x}\right)|\,d\boldsymbol{x}&=\int_\Omega |g_\varepsilon\left(\mathrm{d}_E\left(\boldsymbol{x}\right)\right)-g_0\left(\mathrm{d}_E\left(\boldsymbol{x}\right)\right)|\,|\nabla\,\mathrm{d}_E\left(\boldsymbol{x}\right)|\,d\boldsymbol{x}\\&=\int_{-s_\varepsilon}^{\varphi_\varepsilon(b)-s_\varepsilon}|g_\varepsilon\left(r\right)-g_0\left(r\right)|\,\mathcal{H}^{N-1}\left(\left\{\boldsymbol{x}\in\Omega:\,\mathrm{d}_E\left(\boldsymbol{x}\right)=r\right\}\right)\,dr\\&\le C\varphi_\varepsilon\left(b\right)\sup_{|t|\le\varphi_\varepsilon(b)}\mathcal{H}^{N-1}\left(\left\{\boldsymbol{x}\in\Omega:\,\mathrm{d}_E\left(\boldsymbol{x}\right)=t\right\}\right)\\&\le C\left(b-a\right)\varepsilon^{1/2}\sup_{|t|\le\varphi_\varepsilon(b)}\mathcal{H}^{N-1}\left(\left\{\boldsymbol{x}\in\Omega:\,\mathrm{d}_E\left(\boldsymbol{x}\right)=t\right\}\right)\\&\to 0\cdot\mathcal{H}^{N-1}\left(\Omega\cap\partial E\right)=0.\end{aligned}$$

Step 3: To remove the regularity assumption on the set E, by Lemma 1.15 there exists a sequence of open sets E_k with ∂E_k a nonempty compact hypersurface of class C^2 and $\mathcal{H}^{N-1}\left(\partial E_k\cap\partial\Omega\right)=0$ such that $\chi_{E_k}\to\chi_E$ in $L^1\left(\Omega\right)$, $\mathrm{P}\left(E_k,\Omega\right)\to\mathrm{P}\left(E,\Omega\right)$ and $\mathcal{L}^N\left(E_k\right)=\mathcal{L}^N\left(E\right)$ for all k. By Step 2, for each fixed k we can find a sequence $\{u_{n,k}\}\subset W^{1,2}\left(\Omega\right)$ such that $u_{n,k}\to u_k:=b\chi_{E_k}+a\left(1-\chi_{E_k}\right)$ in $L^1\left(\Omega\right)$ and

$$\lim_{n\to\infty}\int_\Omega\left(\frac{1}{\varepsilon_n}W(u_{n,k})+\varepsilon|\nabla u_{n,k}|^2\right)\,d\boldsymbol{x}=c_W\mathcal{H}^{N-1}\left(\Omega\cap\partial E_k\right)\int_a^b 2\sqrt{W\left(s\right)}\,ds.$$

In turn,

$$\begin{aligned}\limsup_{k\to\infty}\lim_{n\to\infty}\int_\Omega\left(\frac{1}{\varepsilon_n}W(u_{n,k})+\varepsilon|\nabla u_{n,k}|^2\right)\,d\boldsymbol{x}&\le c_W\limsup_{k\to\infty}\mathcal{H}^{N-1}\left(\Omega\cap\partial E_k\right)\\&=c_W\,\mathrm{P}\left(E,\Omega\right).\end{aligned}$$

A diagonalization argument (see Proposition 6.9) yields a sequence $\{u_{n,k_n}\}$ such that $u_{n,k_n}\to u$ in $L^1\left(\Omega\right)$ and

$$\limsup_{n\to\infty}\int_\Omega\left(\frac{1}{\varepsilon_n}W(u_{n,k_n})+\varepsilon|\nabla u_{n,k_n}|^2\right)\,d\boldsymbol{x}\le c_W\,\mathrm{P}\left(E,\Omega\right).$$

□

Theorem 1.14 (Coarea Formula for Lipschitz Functions)**.** *Let $\Omega\subseteq\mathbb{R}^N$ be an open set, let $\psi:\Omega\to\mathbb{R}$ be a Lipschitz function and let $h:\mathbb{R}\to\mathbb{R}$ be a Borel function and assume that $h\circ\psi$ is integrable. Then*

$$\int_\Omega h\left(\psi\left(\boldsymbol{x}\right)\right)|\nabla\psi\left(\boldsymbol{x}\right)|\,d\boldsymbol{x}=\int_{\mathbb{R}}h\left(r\right)\mathcal{H}^{N-1}\left(\left\{\boldsymbol{x}\in\Omega:\,\psi\left(\boldsymbol{x}\right)=r\right\}\right)\,dr.$$

Lemma 1.15. *Assume that $\Omega \subset \mathbb{R}^N$ is a bounded open set with Lipschitz boundary and that $E \subset \mathbb{R}^N$ is a set of finite perimeter. Then there exists a sequence of open sets E_n with ∂E_n a nonempty compact hypersurface of class C^2 and $\mathcal{H}^{N-1}(\partial E_n \cap \partial\Omega) = 0$ such that $\chi_{E_n} \to \chi_E$ in $L^1(\Omega)$, $\mathrm{P}(E_n, \Omega) \to \mathrm{P}(E, \Omega)$ and $\mathcal{L}^N(E_n) = \mathcal{L}^N(E)$ for all n.*

Proof. Extend χ_E outside Ω to a function $w \in BV(\mathbb{R}^N)$, with $0 \le w \le 1$, such that $|Dw|(\partial\Omega) = 0$. Let $w_n := w * \varphi_n$, where φ_n are standard mollifiers. Then $w_n \in C^\infty(\mathbb{R}^N)$, $w_n \to w$ in $L^1(\mathbb{R}^N)$ and

$$\int_{\mathbb{R}^N} |\nabla w_n| \, d\boldsymbol{x} \to |Dw|(\mathbb{R}^N)$$
$$\int_{\Omega} |\nabla w_n| \, d\boldsymbol{x} \to |Dw|(\Omega) = \mathrm{P}(E, \Omega).$$

Consider the open sets $E_{n,t} := \{\boldsymbol{x} \in \mathbb{R}^N : \, w_n(\boldsymbol{x}) > t\}$. By Sard's theorem, for all n and all but $\mathcal{L}^1$ a.e. $t \in \mathbb{R}$, we have that $\partial E_{n,t}$ is a C^∞ manifold of dimension $N-1$. Since $\mathcal{H}^{N-1}(\partial\Omega) < \infty$, and for every fixed n the sets $\{\partial E_{n,t}\}_t$ are disjoint, we have that

$$\mathcal{H}^{N-1}(\partial\Omega \cap \partial E_{n,t}) = 0$$

for all but countably many t. Moreover, for $t \in (0,1)$, using the definition of $E_{n,t}$, we have that

$$\begin{aligned}
\int_{\Omega} |w_n - \chi_E| \, d\boldsymbol{x} &\ge \int_{\Omega \cap (E_{n,t} \setminus E)} |w_n - \chi_E| \, d\boldsymbol{x} + \int_{\Omega \cap (E \setminus E_{n,t})} |w_n - \chi_E| \, d\boldsymbol{x} \\
&= \int_{\Omega \cap (E_{n,t} \setminus E)} w_n \, d\boldsymbol{x} + \int_{\Omega \cap (E \setminus E_{n,t})} |1 - w_n| \, d\boldsymbol{x} \\
&\ge t\mathcal{L}^N(\Omega \cap (E_{n,t} \setminus E)) + (1-t)\,\mathcal{L}^N(\Omega \cap (E \setminus E_{n,t})).
\end{aligned}$$

Letting $n \to \infty$ and since $w_n \to \chi_E$ in $L^1(\Omega)$, we conclude that $\mathcal{L}^N(\Omega \cap (E_{n,t} \setminus E)) \to 0$ and $\mathcal{L}^N(\Omega \cap (E \setminus E_{n,t})) \to 0$, so that

$$\chi_{E_{n,t}} \to \chi_E$$

as $n \to \infty$ for every $t \in (0,1)$. In turn, by the lower semicontinuity of the total variation,

$$\liminf_{n\to\infty} \mathrm{P}(E_{n,t}, \Omega) \ge \mathrm{P}(E, \Omega) \tag{28}$$

for every $t \in (0,1)$. On the other hand, by the coarea formula, the fact that $0 \le w \le 1$, and Fatou's lemma.

$$\begin{aligned}
\mathrm{P}(E, \Omega) = |Dw|(\Omega) &= \lim_{n\to\infty} \int_{\Omega} |\nabla w_n| \, d\boldsymbol{x} \\
&= \lim_{n\to\infty} \int_0^1 \mathrm{P}(\{\boldsymbol{x} \in \Omega : \, w_n(\boldsymbol{x}) > t\}, \Omega) \, dt \\
&= \lim_{n\to\infty} \int_0^1 \mathrm{P}(E_{n,t}, \Omega) \, dt \ge \int_0^1 \liminf_{n\to\infty} \mathrm{P}(E_{n,t}, \Omega) \, dt.
\end{aligned}$$

Hence,

$$\int_0^1 \left(\liminf_{n\to\infty} \mathrm{P}\left(E_{n,t},\Omega\right) - \mathrm{P}\left(E,\Omega\right)\right) dt \le 0.$$

It follows by (28), that $\liminf_{n\to\infty} \mathrm{P}\left(E_{n,t},\Omega\right) = \mathrm{P}\left(E,\Omega\right)$ for $\mathcal{L}^1$ a.e. $t \in (0,1)$.

In conclusion, we have shown that for $\mathcal{L}^1$ a.e. $t \in (0,1)$, $\partial E_{n,t}$ is a C^∞ manifold of dimension $N-1$ for all n with

$$\mathcal{H}^{N-1}\left(\partial\Omega \cap \partial E_{n,t}\right) = 0,$$

and $\chi_{E_{n,t}} \to \chi_E$ in $L^1(\Omega)$ and $\liminf_{n\to\infty} \mathrm{P}\left(E_{n,t},\Omega\right) = \mathrm{P}\left(E,\Omega\right)$. We choose one such t and set $E_n := E_{n,t}$. Next, we want to modify E_n in such a way that

$$\mathcal{L}^N\left(E_n\right) = \mathcal{L}^N\left(E\right)$$

for all n. The argument below is due to Ryan Murray.

For a set of finite perimeter, it can be shown that the total variation measure $|D\chi_E|$ coincides with the measure $\mathcal{H}^{N-1}$ of the *essential boundary* $\partial^* E$ of E, which is given by complement of the set of points $\boldsymbol{x} \in \mathbb{R}^N$ such that the limit

$$\lim_{r\to 0} \frac{\mathcal{L}^N\left(E \cap B\left(\boldsymbol{x},r\right)\right)}{\mathcal{L}^N\left(B\left(\boldsymbol{x},r\right)\right)}$$

exists and is either 0 or 1. It can be shown that for $\mathcal{H}^{N-1}$ a.e. $\boldsymbol{x} \in \partial^* E$, there exists

$$\lim_{r\to 0} \frac{\mathcal{L}^N\left(E \cap B\left(\boldsymbol{x},r\right)\right)}{\mathcal{L}^N\left(B\left(\boldsymbol{x},r\right)\right)} = \frac{1}{2}.$$

Let $\boldsymbol{x}_1$ and $\boldsymbol{x}_2 \in \partial^* E$ be two such points and consider the sets

$$D_k := \left(E \cup B\left(\boldsymbol{x}_1, 1/k\right)\right) \setminus B\left(\boldsymbol{x}_2, 1/k\right).$$

Then $\chi_{D_k} \to \chi_E$ in $L^1(\Omega)$ and

$$\begin{aligned} |D\chi_{D_k}|(\Omega) &= \mathcal{H}^{N-1}\left(\Omega \cap \partial^* D_k\right) \\ &\le \mathcal{H}^{N-1}\left(\Omega \cap \partial^* E\right) + \mathcal{H}^{N-1}\left(\partial B\left(\boldsymbol{x}_1, 1/k\right)\right) + \mathcal{H}^{N-1}\left(\partial B\left(\boldsymbol{x}_2, 1/k\right)\right) \\ &\to \mathcal{H}^{N-1}\left(\Omega \cap \partial^* E\right) = |D\chi_D|(\Omega). \end{aligned}$$

Hence, $\lim_{k\to\infty} \mathrm{P}\left(D_{k,},\Omega\right) = \mathrm{P}\left(E,\Omega\right)$. Using the fact that $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are points of density $\frac{1}{2}$ for E, for k large we have that

$$\mathcal{L}^N\left(E \cap B\left(\boldsymbol{x}_1, 1/k\right)\right) > \frac{1}{4}\mathcal{L}^N\left(B\left(\boldsymbol{x}_1, 1/k\right)\right), \tag{29}$$

$$\mathcal{L}^N\left(E \cap B\left(\boldsymbol{x}_2, 1/k\right)\right) < \frac{3}{4}\mathcal{L}^N\left(B\left(\boldsymbol{x}_2, 1/k\right)\right). \tag{30}$$

Now we approximate D_k as before to get smooth sets $D_{k,n}$. For fixed k, for all n large we have that

$$\left|\mathrm{P}\left(D_{k,},\Omega\right) - \mathrm{P}\left(D_{k,n},\Omega\right)\right| \le \frac{1}{k}, \quad \int_\Omega \left|\chi_{D_k} - \chi_{D_{k,n}}\right| d\boldsymbol{x} \le \frac{1}{k},$$

and by properties of mollifiers,

$$B\left(\boldsymbol{x}_1,\left(\frac{4}{5}\right)^N\frac{1}{k}\right)\subset D_{k,n},\quad \Omega\setminus D_{k,n}\supset B\left(\boldsymbol{x}_1,\left(\frac{4}{5}\right)^N\frac{1}{k}\right). \tag{31}$$

Assume that $\mathcal{L}^N(D_{k,n})>\mathcal{L}^N(E)$. Let

$$A_{k,n}:=D_{k,n}\setminus B(\boldsymbol{x}_1,r_{k,n}),$$

where $r_{k,n}$ is chosen so that $\mathcal{L}^N(B(\boldsymbol{x}_1,r_{k,n}))=\mathcal{L}^N(D_{k,n})-\mathcal{L}^N(E)>0$. We claim that $r_{k,n}<\left(\frac{4}{5}\right)^N\frac{1}{k}$. Indeed, by (29),

$$\begin{aligned}\mathcal{L}^N(D_k)&=\mathcal{L}^N(E)+\mathcal{L}^N(B(\boldsymbol{x}_1,1/k)\setminus E)-\mathcal{L}^N(E\cap B(\boldsymbol{x}_2,1/k))\\&=\mathcal{L}^N(E)+\mathcal{L}^N(B(\boldsymbol{x}_1,1/k))-\mathcal{L}^N(B(\boldsymbol{x}_1,1/k)\cap E)\\&-\mathcal{L}^N(E\cap B(\boldsymbol{x}_2,1/k))\\&<\mathcal{L}^N(E)+\mathcal{L}^N(B(\boldsymbol{x}_1,1/k))-\frac{1}{4}\mathcal{L}^N(B(\boldsymbol{x}_1,1/k))\\&=\mathcal{L}^N(E)+\frac{3}{4}\mathcal{L}^N(B(\boldsymbol{x}_1,1/k)).\end{aligned}$$

Hence, for n large enough,

$$\mathcal{L}^N(D_{k,n})-\mathcal{L}^N(E)<\frac{3}{4}\mathcal{L}^N(B(\boldsymbol{x}_1,1/k)).$$

This shows that $r_{k,n}\le\left(\frac{3}{4}\right)^N\frac{1}{k}<\left(\frac{4}{5}\right)^N\frac{1}{k}$. Hence, in view of (31), the set $A_{k,n}$ is still smooth.

In the case $\mathcal{L}^N(D_{k,n})<\mathcal{L}^N(E)$, we will consider instead

$$A_{k,n}:=D_{k,n}\cup B(\boldsymbol{x}_2,r_{k,n})$$

and use (30). □

Corollary 1.16 (Gurtin's Conjectures)**.** *Let $\Omega\subset\mathbb{R}^N$ be an open bounded set with Lipschitz boundary. Assume that the double-well potential W satisfies conditions (H_1) and (H_2). Then Gurtin's conjectures hold.*

Proof. Let $v_\varepsilon\in W^{1,2}(\Omega)$ be a solution of $(\mathcal{P}_\varepsilon)$ (why does this exist? Exercise). Then v_ε is also a minimizer of the functional $\mathcal{F}_\varepsilon$ defined in (19). Fix any $u\in BV(\Omega;\{a,b\})$ with $\int_\Omega u\,d\boldsymbol{x}=m$ and let u_ε be the function defined in (26). By the minimality of v_ε and by (27),

$$\mathcal{F}_\varepsilon(v_\varepsilon)\le\mathcal{F}_\varepsilon(u_\varepsilon)\le M.$$

It follows by the compactness theorem that up to a subsequence v_ε converges in $L^1(\Omega)$ to a function $v\in BV(\Omega;\{a,b\})$ with $\int_\Omega v\,d\boldsymbol{x}=m$. It follows by Exercise 1.2 that v is a minimizer of the functional $\mathcal{F}$ defined in (20), and so, writing $v=b\chi_{E_0}+a(1-\chi_{E_0})$, we have that

$$c_W\,\mathrm{P}(E_0,\Omega)\le c_W\,\mathrm{P}(E,\Omega)$$

for all functions $u \in BV(\Omega; \{a,b\})$ and $\int_\Omega u\, d\boldsymbol{x} = m$, where $E := \{\boldsymbol{x} \in \Omega : u(\boldsymbol{x}) = b\}$. Moreover, again by Exercise 1.2,

$$\lim_{\varepsilon \to 0^+} \mathcal{F}_\varepsilon (v_\varepsilon) = c_W \operatorname{P}(E_0, \Omega)$$

and so

$$\mathcal{G}_\varepsilon (v_\varepsilon) \sim \varepsilon c_W \operatorname{P}(E_0, \Omega),$$

where $\mathcal{G}_\varepsilon$ is the functional defined in (2). □

2. Variants

2.1. **Phase Transitions for Second Order Materials.** Let's consider the functional

$$\int_\Omega \left(\frac{1}{\varepsilon} W(u) - q\varepsilon |\nabla u|^2 + \varepsilon^3 |\nabla^2 u|^2 \right) d\boldsymbol{x}, \tag{32}$$

where $q \in \mathbb{R}$. The case $q = 0$ was studied by Fonseca and Mantegazza [FMa2000], the case $q < 0$ by Hilhorst, Peletier, and Schätzle [HPS2002], while the case $q > 0$ small by Chermisi, Dal Maso, Fonseca, and G. L. [CDMFL2011] and by Cicalese, Spadaro, and Zeppieri [CSZ2011] for $N = 1$.

When $N = 1$, $\varepsilon = 1$, and

$$W(t) = \frac{1}{2} \left(t^2 - 1\right)^2, \tag{33}$$

the functional reduces to

$$\int_I \left(\frac{1}{2} \left(u^2 - 1\right)^2 - q|u'|^2 + |u''|^2 \right) dx.$$

In this case the Euler–Lagrange equation of the functional is given by

$$u^{(iv)} + qu'' + u^3 - u = 0. \tag{34}$$

When $q < 0$ this is the stationary solution of the extended Fisher–Kolmogorov equation

$$\frac{\partial v}{\partial t} = -\gamma \frac{\partial^4 v}{\partial x^4} + \frac{\partial^2 v}{\partial x^2} + v - v^3, \quad \gamma > 0,$$

which was introduced by Coullet, Elphick, and Repaux [CER1987] and by Dee and van Saarloos [DvS1988] to study pattern formation in bistable systems, while for $q > 0$ it is the stationary solution of the Swift–Hohenberg equation

$$\frac{\partial v}{\partial t} = -\left(1 + \frac{\partial^2}{\partial x^2}\right)^2 v + \alpha v - v^3, \quad \alpha > 0,$$

which was introduced in [SH1977] to study Rayleigh-Bénard convection. The equation (34) has been studied by several authors in both cases $q > 0$ and $q < 0$ (see, e.g., [MPT1998], [PT1997], [SvdB2002] and the references therein).

Our motivation in [CDMFL2011] was a nonlocal variational model introduced by Andelman, Kawasaki, Kawakatsu, and Taniguchi [KAKT1993], [TKAK1994], (see also [LA1987], [SA1995]) for the shape deformation of unilamellar membranes undergoing an inplane phase separation. A simplified local version of this model (see [SA1995]) leads to the study of (32).

The model (32) in the one-dimensional case was independently proposed by Coleman, Marcus, and Mizel in [CMM1992] (see also [LM1989]) in connection with the study of periodic or quasiperiodic layered structures.

In the case $q < 0$, compactness follows from what we have done before. The difficult case is $q \geq 0$. Note that if $q > 0$ is very large, then the functional may not be bounded from below. We will study here the case in which q is very small. For simplicity, we take $N = 1$ and W takes the form (33) and refer to [CDMFL2011] for the case $N \geq 1$ and more general wells.

The main result behind compactness is the following nonlinear interpolation result. The proof is taken from [CSZ2011] (see also [CDMFL2011] for an alternative proof). Classical interpolation results are due to Gagliardo [Ga1959] and Nirenberg [Ni1966].

Theorem 2.1. *There exists $q_0 > 0$ such that*

$$q_0 \int_c^d |u'|^2 \, dx \leq \frac{1}{(d-c)^2} \int_c^d |W(u)| \, dx + (d-c)^2 \int_c^d |u''|^2 \, dx$$

for all $c < d$ and all $u \in W^{2,2}(c,d)$.

Proof. By rescaling and translating, we can assume that $(c,d) = (0,1)$. By the mean value theorem there exists $x_0 \in (0,1)$ such that

$$u'(x_0) = \int_0^1 u' \, dx$$

and so, by the fundamental theorem of calculus,

$$u'(x) = u'(x_0) + \int_{x_0}^x u'' \, dt.$$

In turn,

$$|u'(x) - u'(x_0)| \leq \int_0^1 |u''| \, dt.$$

It follows by Hölder's inequality,

$$|u'(x) - u'(x_0)|^2 \leq \int_0^1 |u''|^2 \, dt, \tag{35}$$

and so

$$|u'(x)|^2 \leq 2|u'(x_0)|^2 + 2\int_0^1 |u''|^2 \, dt.$$

Upon integration over x, we get

$$\int_0^1 |u'(x)|^2 \, dx \leq 2|u'(x_0)|^2 + 2\int_0^1 |u''|^2 \, dt.$$

To conclude the proof, it remains to show that there exists a constant $\ell > 0$ such that

$$\ell |u'(x_0)|^2 \leq \int_0^1 |W(u)| \, dx + \int_0^1 |u''|^2 \, dx.$$

There are two cases. If

$$\left|u'\left(x_0\right)\right|^2 \leq 4 \int_0^1 \left|u''\right|^2 \, dx,$$

then there is nothing to prove. Thus, assume that

$$4 \int_0^1 \left|u''\right|^2 \, dx < \left|u'\left(x_0\right)\right|^2 .$$

Then by (35),

$$\left|u'\left(x\right) - u'\left(x_0\right)\right|^2 < \frac{1}{4} \left|u'\left(x_0\right)\right|^2 . \tag{36}$$

This implies that

$$\frac{1}{2} \left|u'\left(x_0\right)\right| < \left|u'\left(x\right)\right| < \frac{3}{2} \left|u'\left(x_0\right)\right| \tag{37}$$

for all $x \in (0,1)$. Therefore u is strictly monotone. Hence, it does not vanish in one of the two intervals $\left(0, \frac{1}{2}\right)$ and $\left(\frac{1}{2}, 1\right)$, say $u > 0$ in $\left(\frac{1}{2}, 1\right)$ (the other cases are analogous). By the classical interpolation inequality applied to the function $u - 1$, we get

$$\int_{\frac{1}{2}}^1 \left|u'\right|^2 \, dx \leq C \int_{\frac{1}{2}}^1 \left(u-1\right)^2 \, dx + C \int_{\frac{1}{2}}^1 \left|u''\right|^2 \, dx.$$

Now, since $u > 0$ in $\left(\frac{1}{2}, 1\right)$, $\left(u-1\right)^2 \left(u+1\right)^2 \geq \left(u-1\right)^2$, and so we obtain

$$\int_{\frac{1}{2}}^1 \left|u'\right|^2 \, dx \leq C \int_0^1 W\left(u\right) \, dx + C \int_0^1 \left|u''\right|^2 \, dx.$$

In turn, from (36),

$$\left|u'\left(x_0\right)\right|^2 \leq C \int_0^1 W\left(u\right) \, dx + C \int_0^1 \left|u''\right|^2 \, dx,$$

which is what we wanted to prove. □

From the previous theorem, we obtain the following result.

Corollary 2.2. *For every open interval I and every $q < q_0$ there exists $\varepsilon_0 = \varepsilon_0\left(I, q\right) > 0$ such that for $0 < \varepsilon < \varepsilon_0$,*

$$q\varepsilon^2 \int_I \left|u'\right|^2 \, dx \leq \int_I W(u) \, dx + \varepsilon^4 \int_I \left|u''\right|^2 \, dx$$

for all $u \in W^{2,2}_{\text{loc}}(I)$.

Proof. Assume $0 < q < q_0$. Consider the function $v\left(y\right) := u\left(\varepsilon y\right)$ for $y \in I/\varepsilon := \{z \in \mathbb{R} : \varepsilon z \in I\}$. Let n_ε be the integer part of $\frac{1}{\varepsilon}\,\text{length}\left(I\right)$ and divide I into n_ε open intervals $I_{k,\varepsilon}$ of length $\frac{1}{\varepsilon n_\varepsilon}\,\text{length}\left(I\right)$ and apply the previous theorem to the function v in each interval $I_{k,\varepsilon}$ to get

$$q_0\varepsilon^2 \int_{I_{k,\varepsilon}} \left|u'\left(\varepsilon y\right)\right|^2 \, dy \leq \frac{\varepsilon^2 n_\varepsilon^2}{\text{length}^2\left(I\right)} \int_{I_{k,\varepsilon}} \left|W\left(u\left(\varepsilon y\right)\right)\right| \, dy + \frac{\text{length}^2\left(I\right)}{\varepsilon^2 n_\varepsilon^2} \varepsilon^4 \int_{I_{k,\varepsilon}} \left|u''\left(\varepsilon y\right)\right|^2 \, dy.$$

Summing over k and changing variables, we get

$$q\varepsilon^2 \int_I |u'(x)|^2 \, dx \le \frac{q}{q_0} \frac{\varepsilon^2 n_\varepsilon^2}{\text{length}^2(I)} \int_I |W(u(x))| \, dx + \frac{q}{q_0} \frac{\text{length}^2(I)}{\varepsilon^2 n_\varepsilon^2} \varepsilon^4 \int_I |u''(x)|^2 \, dx.$$

The result now follows by observing that $\frac{q}{q_0} < 1$ and

$$\frac{\varepsilon n_\varepsilon}{\text{length}(I)} \to 1$$

as $\varepsilon \to 0^+$. □

Using the previous corollary, we can prove compactness for $q < q_0$.

Theorem 2.3 (Compactness)**.** *Let $I \subset \mathbb{R}$ be an interval and let $q < q_0$. Let $\varepsilon_n \to 0^+$ and let $\{u_n\} \subset W^{1,2}(\Omega)$ be such that*

$$\sup_n \int_I \left(\frac{1}{\varepsilon_n} W(u_n) - q\varepsilon_n |u_n'|^2 + \varepsilon_n^3 |u_n''|^2 \right) dx < \infty.$$

Then there exist a subsequence $\{u_{n_k}\}$ of $\{u_n\}$ and $u \in BV(I; \{-1, 1\})$ such that

$$u_{n_k} \to u \text{ for in } L^1(I).$$

Proof. Let $\delta > 0$ be so small that $\frac{q+\delta}{1-\delta} < q_0$ and write

$$\begin{aligned} \frac{1}{\varepsilon_n} W(u_n) - q\varepsilon_n |u_n'|^2 + \varepsilon_n^3 |u_n''|^2 &= (1-\delta) \left(\frac{1}{\varepsilon_n} W(u_n) - \frac{q+\delta}{1-\delta} \varepsilon_n |u_n'|^2 + \varepsilon_n^3 |u_n''|^2 \right) dx \\ &\quad + \delta \left(\frac{1}{\varepsilon_n} W(u_n) + \varepsilon_n |u_n'|^2 + \varepsilon_n^3 |u_n''|^2 \right). \end{aligned}$$

By the previous theorem, for all n large enough

$$\int_I \left(\frac{1}{\varepsilon_n} W(u_n) - \frac{q+\delta}{1-\delta} \varepsilon_n |u_n'|^2 + \varepsilon_n^3 |u_n''|^2 \right) dx \ge 0$$

and so

$$\delta \int_I \left(\frac{1}{\varepsilon_n} W(u_n) + \varepsilon_n |u_n'|^2 \right) dx \le \int_I \left(\frac{1}{\varepsilon_n} W(u_n) - q\varepsilon_n |u_n'|^2 + \varepsilon_n^3 |u_n''|^2 \right) dx$$

for all n sufficiently large. We can now apply Theorem 1.6. □

2.2. **The Vectorial Case $d \ge 1$.** This was studied by Sternberg [St1988] in the case when the wells are two closed curves in $\mathbb{R}^2$, by Fonseca and Tartar [FT1989] in the case of two wells in $\mathbb{R}^d$, by Baldo [Ba1990] in the case of multiple wells, and by Ambrosio [Am1990] who considered the case in which the set of zeros of W is a compact set (see also [St1991]). Let's describe the case of two wells. For $\varepsilon > 0$ consider the functional

$$\mathcal{F}_\varepsilon : W^{1,2}(\Omega; \mathbb{R}^d) \to [0, \infty]$$

defined by

$$\mathcal{F}_\varepsilon(\boldsymbol{u}) := \int_\Omega \left(\frac{1}{\varepsilon} W(\boldsymbol{u}) + \varepsilon |\nabla \boldsymbol{u}|^2 \right) d\boldsymbol{x},$$

where the double well potential $W : \mathbb{R}^d \to [0, \infty)$ satisfies the following hypotheses:

(H_1) W is continuous, $W(\boldsymbol{z}) = 0$ if and only if $\boldsymbol{z} \in \{\boldsymbol{a}, \boldsymbol{b}\}$ for some $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ with $\boldsymbol{a} \neq \boldsymbol{b}$.

(H_2) There exist $L > 0$ and $T > 0$ such that

$$W(\boldsymbol{z}) \geq L |\boldsymbol{z}| .$$

for all $\boldsymbol{z} \in \mathbb{R}^d$ with $|\boldsymbol{z}| \geq T$.

The analogue of Theorem 1.6 is the following compactness theorem. There are several proofs of this result: the one in [FT1989] uses Young measures, while the one [Ba1990] does not. We present here another proof, due to Massimiliano Morini, that makes use of Theorem 1.6.

Theorem 2.4 (Compactness)**.** *Let $\Omega \subset \mathbb{R}^N$ be an open set with finite measure. Assume that the double-well potential W satisfies conditions (H_1) and (H_2). Let $\varepsilon_n \to 0^+$ and let $\{\boldsymbol{u}_n\} \subset W^{1,2}(\Omega; \mathbb{R}^d)$ be such that*

$$M := \sup_n \mathcal{F}_{\varepsilon_n}(\boldsymbol{u}_n) < \infty.$$

Then there exist a subsequence $\{\boldsymbol{u}_{n_k}\}$ of $\{\boldsymbol{u}_n\}$ and $\boldsymbol{u} \in BV(\Omega; \{\boldsymbol{a}, \boldsymbol{b}\})$ such that

$$\boldsymbol{u}_{n_k} \to \boldsymbol{u} \text{ for in } L^1(\Omega; \mathbb{R}^d) .$$

First proof. **Step 1:** Assume that $|\boldsymbol{a}| \neq |\boldsymbol{b}|$. For every $t \geq 0$ define

$$V(t) := \min_{|\boldsymbol{z}|=t} W(\boldsymbol{z}) .$$

Then V is upper semicontinuous, $V(t) > 0$ for $t \neq |\boldsymbol{a}|$, $|\boldsymbol{b}|$, $V(|\boldsymbol{a}|) = V(|\boldsymbol{b}|) = 0$, and $V(t) \geq Lt$ for $t \geq T$. For every $\boldsymbol{u} \in W^{1,2}(\Omega; \mathbb{R}^d)$ define

$$\mathcal{H}_\varepsilon(\boldsymbol{u}) := \int_\Omega \left(\frac{1}{\varepsilon} V(|\boldsymbol{u}|) + \varepsilon |\nabla |\boldsymbol{u}||^2 \right) d\boldsymbol{x} \leq \mathcal{F}_\varepsilon(\boldsymbol{u}) .$$

Then by (9),

$$\sup_n \mathcal{H}_{\varepsilon_n}(\boldsymbol{u}_n) \leq \sup_n \mathcal{F}_{\varepsilon_n}(\boldsymbol{u}_n; \Omega) < \infty,$$

and so by the compactness in the scalar case $d = 1$, there exist a subsequence $\{\boldsymbol{u}_{n_k}\}$ and $w \in BV(\Omega)$ such that

$$w_k := \Phi_2 \circ |\boldsymbol{u}_{n_k}| \to w \text{ in } L^1_{\text{loc}}(\Omega) ,$$

where

$$\Phi_2(t) := \frac{1}{2} \int_0^t \sqrt{V_1(s)}\, ds, \quad t \in \mathbb{R}$$

and

$$V_1(\boldsymbol{z}) := \min \{V(\boldsymbol{z}), K\}, \quad \boldsymbol{z} \in \mathbb{R}.$$

Hence,

$$|\boldsymbol{u}_{n_k}| \to v := \Phi_2^{-1} \circ w \text{ in } L^1(\Omega) .$$

By taking a further subsequence, if necessary, without loss of generality, we may assume that $|\boldsymbol{u}_{n_k}(\boldsymbol{x})| \to v(\boldsymbol{x})$ and that $W(\boldsymbol{u}_{n_k}(\boldsymbol{x})) \to 0$ for $\mathcal{L}^N$ a.e. $\boldsymbol{x} \in \Omega$. This implies that $v \in BV(\Omega; \{\boldsymbol{a}, \boldsymbol{b}\})$. Define

$$\boldsymbol{u}(\boldsymbol{x}) := \begin{cases} \boldsymbol{a} & \text{if } v(\boldsymbol{x}) = |\boldsymbol{a}|, \\ \boldsymbol{b} & \text{if } v(\boldsymbol{x}) = |\boldsymbol{b}|. \end{cases}$$

We claim that

$$\boldsymbol{u}_{n_k} \to \boldsymbol{u} \text{ in } L^1\left(\Omega; \mathbb{R}^d\right).$$

To see this, fix $\boldsymbol{x} \in \Omega$ such that $|\boldsymbol{u}_{n_k}(\boldsymbol{x})| \to v(\boldsymbol{x})$ and $W(\boldsymbol{u}_{n_k}(\boldsymbol{x})) \to 0$. Then, by (H_1), necessarily, $\boldsymbol{u}_{n_k}(\boldsymbol{x}) \to \boldsymbol{u}(\boldsymbol{x})$.

Step 2: If $|\boldsymbol{a}| = |\boldsymbol{b}|$, let $\boldsymbol{e}_i$ be a vector of the canonical basis of $\mathbb{R}^d$ such that $\boldsymbol{a} \cdot \boldsymbol{e}_i \neq \boldsymbol{b} \cdot \boldsymbol{e}_i$. Then $|\boldsymbol{a} + \boldsymbol{e}_i| \neq |\boldsymbol{b} + \boldsymbol{e}_i|$. It suffices to apply the previous step with W replaced by

$$\hat{W}(\boldsymbol{z}) := W(\boldsymbol{z} - \boldsymbol{e}_i), \quad \boldsymbol{z} \in \mathbb{R}^d,$$

and $\boldsymbol{u}_n$ by $\boldsymbol{u}_n + \boldsymbol{e}_i$. □

The second proof is adapted from [Ba1990] and [FT1989].

Second proof. For $K > 0$ define

$$W_1(\boldsymbol{z}) := \min\{W(\boldsymbol{z}), K\}, \quad \boldsymbol{z} \in \mathbb{R}^d \tag{38}$$

and consider the "geodesic distance" in $\mathbb{R}^d$ given by

$$d(\boldsymbol{v}, \boldsymbol{w}) := 2\inf\Big\{\textstyle\int_{-1}^{1} \sqrt{W_1(\boldsymbol{g}(t))}\, |\boldsymbol{g}'(t)|\, dt : \boldsymbol{g} \text{ piecewise } C^1 \text{ curve}, \boldsymbol{g}(-1) = \boldsymbol{v}, \boldsymbol{g}(1) = \boldsymbol{w}\Big\} \tag{39}$$

for $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^d$. We claim that the function

$$f(\boldsymbol{z}) := d(\boldsymbol{a}, \boldsymbol{z}), \quad \boldsymbol{z} \in \mathbb{R}^d,$$

is Lipschitz. Indeed, let $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^d$ and let $\boldsymbol{\gamma}$ be a piecewise C^1 curve joining $\boldsymbol{a}$ with $\boldsymbol{v}$. Then

$$f(\boldsymbol{w}) \leq 2\int_{\boldsymbol{\gamma}} \sqrt{W_1}\, ds + 2\int_{[\boldsymbol{v},\boldsymbol{w}]} \sqrt{W_1}\, ds \leq 2\int_{\boldsymbol{\gamma}} \sqrt{W_1}\, ds + 2\sqrt{K}\, |\boldsymbol{v} - \boldsymbol{w}|,$$

where $[\boldsymbol{v}, \boldsymbol{w}]$ is the segment of endpoints $\boldsymbol{v}, \boldsymbol{w}$. Taking the infimum over all curves $\boldsymbol{\gamma}$, we get

$$f(\boldsymbol{w}) \leq f(\boldsymbol{v}) + 2\sqrt{K}\, |\boldsymbol{v} - \boldsymbol{w}|,$$

which shows that f is Lipschitz continuous.

Next we prove that

$$\int_\Omega |\nabla (f \circ \mathbf{u})(\boldsymbol{x})|\, d\,\boldsymbol{x} \leq 2\int_\Omega \sqrt{W(\mathbf{u}(\boldsymbol{x}))} |\nabla \mathbf{u}(\boldsymbol{x})|\, d\boldsymbol{x} \tag{40}$$

for all $\mathbf{u} \in W^{1,2}\left(\Omega; \mathbb{R}^d\right)$. Assume first that $\mathbf{u} \in C^1\left(\Omega; \mathbb{R}^d\right)$. Then the function $f \circ \mathbf{u}$ is locally Lipschitz, and thus, by Rademacher's theorem, it is differentiable $\mathcal{L}^N$ a.e. Fix a point $\mathbf{x}_0 \in \Omega$ such that $f \circ \mathbf{u}$ is differentiable at $\mathbf{x}_0$. Let $i \in \{1, \dots, N\}$, let $h > 0$ and let $\boldsymbol{\gamma}$ be a piecewise C^1 curve joining $\boldsymbol{a}$ with $\mathbf{u}(\boldsymbol{x}_0)$. Then

$$\begin{aligned} f\left(\mathbf{u}\left(\boldsymbol{x}_0 + h\boldsymbol{e}_i\right)\right) &\leq 2\int_{\boldsymbol{\gamma}} \sqrt{W_1}\, ds + 2\int_{[\mathbf{u}(\boldsymbol{x}_0), \mathbf{u}(\boldsymbol{x}_0 + h\boldsymbol{e}_i)]} \sqrt{W_1}\, ds = 2\int_{\boldsymbol{\gamma}} \sqrt{W_1}\, ds \\ &+ 2\left|\mathbf{u}\left(\boldsymbol{x}_0 + h\boldsymbol{e}_i\right) - \mathbf{u}\left(\boldsymbol{x}_0\right)\right| \int_0^1 \sqrt{W_1\left(s\mathbf{u}\left(\boldsymbol{x}_0\right) + (1-s)\,\mathbf{u}\left(\boldsymbol{x}_0 + h\boldsymbol{e}_i\right)\right)}\, ds. \end{aligned}$$

Taking the infimum over all curves $\boldsymbol{\gamma}$ and applying the mean value theorem to the last integral, we get

$$\begin{aligned} & f\left(\mathbf{u}\left(\boldsymbol{x}_0 + h\boldsymbol{e}_i\right)\right) - f\left(\mathbf{u}\left(\boldsymbol{x}_0\right)\right) \\ & \leq 2\left|\mathbf{u}\left(\boldsymbol{x}_0 + h\boldsymbol{e}_i\right) - \mathbf{u}\left(\boldsymbol{x}_0\right)\right| \sqrt{W_1\left(\theta\mathbf{u}\left(\boldsymbol{x}_0\right) + (1-\theta)\,\mathbf{u}\left(\boldsymbol{x}_0 + h\boldsymbol{e}_i\right)\right)} \end{aligned}$$

for some $\theta \in [0,1]$. Dividing by h and letting $h \to 0^+$ yields

$$\frac{\partial\left(f \circ \mathbf{u}\right)}{\partial x_i}\left(\mathbf{x}_0\right) \leq 2\left|\frac{\partial \mathbf{u}}{\partial x_i}\left(\mathbf{x}_0\right)\right| \sqrt{W_1\left(\mathbf{u}\left(\boldsymbol{x}_0\right)\right)}.$$

By inverting the roles of $\mathbf{u}\left(\boldsymbol{x}_0 + h\boldsymbol{e}_i\right)$ and $\mathbf{u}\left(\boldsymbol{x}_0\right)$ we get

$$\left|\frac{\partial\left(f \circ \mathbf{u}\right)}{\partial x_i}\left(\mathbf{x}_0\right)\right| \leq 2\left|\frac{\partial \mathbf{u}}{\partial x_i}\left(\mathbf{x}_0\right)\right| \sqrt{W_1\left(\mathbf{u}\left(\boldsymbol{x}_0\right)\right)}. \tag{41}$$

In turn,

$$\begin{aligned} \left|\nabla\left(f \circ \mathbf{u}\right)\left(\boldsymbol{x}_0\right)\right| &= \sqrt{\sum_{i=1} \left|\frac{\partial\left(f \circ \mathbf{u}\right)}{\partial x_i}\left(\mathbf{x}_0\right)\right|^2} \\ &\leq \sqrt{\sum_{i=1} \left(2\left|\frac{\partial \mathbf{u}}{\partial x_i}\left(\mathbf{x}_0\right)\right| \sqrt{W_1\left(\mathbf{u}\left(\boldsymbol{x}_0\right)\right)}\right)^2} = 2\sqrt{W_1\left(\mathbf{u}\left(\boldsymbol{x}_0\right)\right)}\left|\nabla\mathbf{u}\left(\boldsymbol{x}_0\right)\right|. \end{aligned}$$

This proves the claim for $\mathbf{u} \in C^1\left(\Omega; \mathbb{R}^d\right)$. In the general case, $\mathbf{u} \in W^{1,2}\left(\Omega; \mathbb{R}^d\right)$, we can use the Meyers–Serrin theorem to approximate $\mathbf{u} \in W^{1,2}\left(\Omega; \mathbb{R}^d\right)$ with a sequence of functions $\mathbf{u}_k \in W^{1,2}\left(\Omega; \mathbb{R}^d\right) \cap C^1\left(\Omega; \mathbb{R}^d\right)$ converging to $\mathbf{u}$ in $W^{1,2}\left(\Omega; \mathbb{R}^d\right)$. By selecting a subsequence, we may assume that $\{\mathbf{u}_k\}$ and $\{\nabla\mathbf{u}_k\}$ converge to $\mathbf{u}$ and $\nabla\mathbf{u}$ pointwise $\mathcal{L}^N$ a.e. and that |

$$\left|\mathbf{u}_k\left(\boldsymbol{x}\right)\right|^2 + \left|\nabla\mathbf{u}_k\left(\boldsymbol{x}\right)\right|^2 \leq h\left(\boldsymbol{x}\right) \tag{42}$$

for $\mathcal{L}^N$ a.e. $\mathbf{x} \in \Omega$ and all k and for some integrable function h. Since f is Lipschitz, it follows that $\{f \circ \mathbf{u}_k\}$ converges to $f \circ \mathbf{u}$ in $L^2(\Omega)$ and pointwise $\mathcal{L}^N$ a.e. Moreover, by (41) applied to $\mathbf{u}_k$ and the fact that W_1 is bounded, we have that $\{f \circ \mathbf{u}_k\}$ is bounded in $W^{1,2}(\Omega)$. Hence, it converges weakly to $f \circ \mathbf{u}$ in $W^{1,2}(\Omega)$. By (40) applied to $\mathbf{u}_k$,

$$\int_\Omega \left|\nabla\left(f \circ \mathbf{u}_k\right)\left(\boldsymbol{x}\right)\right| d\,\boldsymbol{x} \leq 2\int_\Omega \sqrt{W\left(\mathbf{u}_k\left(\boldsymbol{x}\right)\right)}\left|\nabla\mathbf{u}_k\left(\boldsymbol{x}\right)\right| d\boldsymbol{x}$$

for all k. Letting $k \to \infty$, and using the lower semicontinuity of the L^2 norm on the left-hand side and Lebesgue dominated convergence theorem (which can be applied by (40)) on the right-hand side, we conclude that (40) holds for $\mathbf{u}$.

Using (40) in place of (15), we can conclude as in the proof of Theorem 1.6 that if $\{\boldsymbol{u}_n\} \subset W^{1,2}\left(\Omega; \mathbb{R}^d\right)$ is such that $\sup_n \mathcal{F}_{\varepsilon_n}(\boldsymbol{u}_n) < \infty$, then there exist a subsequence, not relabeled, and a function $w \in BV(\Omega)$ such that

$$w_n := f \circ \boldsymbol{u}_n \to w \text{ in } L^1_{\mathrm{loc}}(\Omega).$$

By selecting a further subsequence, we may assume that $\{w_n\}$ converges to w pointwise $\mathcal{L}^N$ a.e. It remains to show that $\{\boldsymbol{u}_n\}$ converges to $\boldsymbol{u}$ pointwise $\mathcal{L}^N$ a.e. Define

$$E := \{x \in \Omega : \, w(\boldsymbol{x}) > 0\}$$

and

$$\boldsymbol{u} := \boldsymbol{b}\chi_E + \boldsymbol{a}\chi_{\Omega \setminus E}.$$

Let $\boldsymbol{x} \in E$ be such that $w_n(\boldsymbol{x}) \to w(\boldsymbol{x})$ and $W(\boldsymbol{u}_n(\boldsymbol{x})) \to 0$. Consider a subsequence $\{\boldsymbol{u}_{n_k}(\boldsymbol{x})\}$. By (H_1) and the fact that $W(\boldsymbol{u}_n(\boldsymbol{x})) \to 0$, there exists a further subsequence $\left\{\boldsymbol{u}_{n_{k_i}}(\boldsymbol{x})\right\}$ of $\{\boldsymbol{u}_{n_k}(\boldsymbol{x})\}$ such that either $\boldsymbol{u}_{n_{k_i}}(\boldsymbol{x}) \to \boldsymbol{a}$ or $\boldsymbol{u}_{n_{k_i}}(\boldsymbol{x}) \to \boldsymbol{b}$. We claim that the case $\boldsymbol{u}_{n_{k_i}}(\boldsymbol{x}) \to \boldsymbol{a}$ cannot happen. Indeed, if this were the case, then by the continuity of f,

$$w_{k_i}(\boldsymbol{x}) = f(\boldsymbol{u}_{n_k}(\boldsymbol{x})) \to f(\boldsymbol{a}) = 0,$$

which contradicts the fact that $w_n(\boldsymbol{x}) \to w(\boldsymbol{x}) > 0$. This shows that $\boldsymbol{u}_{n_{k_i}}(\boldsymbol{x}) \to \boldsymbol{b}$, and by the arbitrariness of the subsequence, that $\boldsymbol{u}_n(\boldsymbol{x}) \to \boldsymbol{b}$. Similarly, we can show that if $w(\boldsymbol{x}) = 0$, $w_n(\boldsymbol{x}) \to w(\boldsymbol{x})$, and $W(\boldsymbol{u}_n(\boldsymbol{x})) \to 0$, then $\boldsymbol{u}_n(\boldsymbol{x}) \to \boldsymbol{a}$. This proves that $\{\boldsymbol{u}_n\}$ converges pointwise to $\boldsymbol{u}$. □

A Γ-convergence result similar to the one given in Theorems 1.12 and 1.13 holds. The main changes are the constant c_W in (21) and the fact that, in addition to (H_1) and (H_2), W is also assumed to be locally Lipschitz and quadratic near the wells, precisely:

(H_3) W is Lipschitz on compact sets and there exist $l > 0$ and $\delta > 0$ such that

$$\frac{1}{l}|\boldsymbol{z} - \boldsymbol{a}|^2 \leq W(\boldsymbol{z}) \leq l|\boldsymbol{z} - \boldsymbol{a}|^2 \quad \text{for all } \boldsymbol{z} \in \mathbb{R}^d \text{ with } |\boldsymbol{z} - \boldsymbol{a}| \leq \delta,$$
$$\frac{1}{l}|\boldsymbol{z} - \boldsymbol{b}|^2 \leq W(\boldsymbol{z}) \leq l|\boldsymbol{z} - \boldsymbol{b}|^2 \quad \text{for all } \boldsymbol{z} \in \mathbb{R}^d \text{ with } |\boldsymbol{z} - \boldsymbol{b}| \leq \delta.$$

As in the scalar case we extend $\mathcal{F}_\varepsilon$ to $L^1\left(\Omega; \mathbb{R}^d\right)$ by setting

$$\mathcal{F}_\varepsilon(\boldsymbol{u}) := \begin{cases} \displaystyle\int_\Omega \left(\frac{1}{\varepsilon}W(\boldsymbol{u}) + \varepsilon|\nabla \boldsymbol{u}|^2\right) d\boldsymbol{x} & \text{if } \boldsymbol{u} \in W^{1,2}\left(\Omega; \mathbb{R}^d\right) \text{ and } \int_\Omega \boldsymbol{u}\, d\boldsymbol{x} = \boldsymbol{m}, \\ \infty & \text{otherwise in } L^1\left(\Omega; \mathbb{R}^d\right). \end{cases}$$

Let $\varepsilon_n \to 0^+$. Under appropriate hypotheses on W and Ω, we will show that the sequence of functionals $\{\mathcal{F}_{\varepsilon_n}\}$ Γ-converges to the functional

$$\mathcal{F}(\boldsymbol{u}) := \begin{cases} c_W \operatorname{P}(E,\Omega) & \text{if } \boldsymbol{u} \in BV(\Omega;\{\boldsymbol{a},\boldsymbol{b}\}) \text{ and } \int_\Omega \boldsymbol{u}\, d\boldsymbol{x} = \boldsymbol{m}, \\ \infty & \text{otherwise in } L^1(\Omega;\mathbb{R}^d), \end{cases}$$

where

$$c_W := 2\inf\Big\{\textstyle\int_{-1}^1 \sqrt{W(\boldsymbol{g}(t))}\,|\boldsymbol{g}'(t)|\,dt : \\ \boldsymbol{g} \text{ piecewise } C^1 \text{ curve}, \boldsymbol{g}(-1)=\boldsymbol{a},\, \boldsymbol{g}(1)=\boldsymbol{b}\Big\},$$

and $E := \{\boldsymbol{x}\in\Omega : \boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{b}\}$.

Theorem 2.5 (Liminf inequality)**.** *Let $\Omega \subset \mathbb{R}^N$ be an open set with finite measure. Assume that the double-well potential W satisfies conditions (H_1), and (H_2), and (H_3). Let $\varepsilon_n \to 0^+$. Let $\varepsilon_n \to 0^+$ and let $\{\boldsymbol{u}_n\} \subset L^1(\Omega;\mathbb{R}^d)$ be such that $\boldsymbol{u}_n \to \boldsymbol{u}$ in $L^1(\Omega)$. Then*

$$\liminf_{n\to\infty} \mathcal{F}_{\varepsilon_n}(\boldsymbol{u}_n) \geq \mathcal{F}(\boldsymbol{u}).$$

Proof. We begin by proving that

$$\boldsymbol{c}_W = f(\boldsymbol{b}) = d(\boldsymbol{a},\boldsymbol{b}) \tag{43}$$

if the constant K in (38) is chosen large enough. To see this, not that since $W_1 \leq W$, we have that $f(\boldsymbol{b}) \leq \boldsymbol{c}_W$. To prove the converse inequality, let

$$h(r) := \min_{\left|\boldsymbol{z}-\frac{\boldsymbol{a}+\boldsymbol{b}}{2}\right|=r} \sqrt{W(\boldsymbol{z})}.$$

By (H_2) there exists $r_1 > r_0 := \left|\frac{\boldsymbol{a}-\boldsymbol{b}}{2}\right|$ such that

$$\int_{r_0}^{r_1} h(r)\,dr > c_W.$$

Take $K := \max_{\left|\boldsymbol{z}-\frac{\boldsymbol{a}+\boldsymbol{b}}{2}\right|\leq r_1} W(\boldsymbol{z})$. Let now $\boldsymbol{g}$ be a piecewise C^1 curve with $\boldsymbol{g}(-1) = \boldsymbol{a}$ and $\boldsymbol{g}(1) = \boldsymbol{b}$. If $\left|\boldsymbol{g}(t) - \frac{\boldsymbol{a}+\boldsymbol{b}}{2}\right| \leq r_1$ for all $t \in [-1,1]$, then $W(\boldsymbol{g}(t)) = W_1(\boldsymbol{g}(t))$ for all $t\in[-1,1]$. Hence,

$$\textstyle\int_{-1}^1 \sqrt{W_1(\boldsymbol{g}(t))}\,|\boldsymbol{g}'(t)|\,dt = \int_{-1}^1 \sqrt{W(\boldsymbol{g}(t))}\,|\boldsymbol{g}'(t)|\,dt \geq c_W.$$

On the other hand, if there exists a $t_0 \in [-1,1]$ such that $\left|\boldsymbol{g}(t_0) - \frac{\boldsymbol{a}+\boldsymbol{b}}{2}\right| > r_1$, then since $\left|\boldsymbol{g}(-1) - \frac{\boldsymbol{a}+\boldsymbol{b}}{2}\right| = r_0 < r_1$, we have that there exists a first time t_1 such that $\left|\boldsymbol{g}(t_1) - \frac{\boldsymbol{a}+\boldsymbol{b}}{2}\right| = r_1$. In turn, $W(\boldsymbol{g}(t)) = W_1(\boldsymbol{g}(t))$ for all $t \in [-1, t_1]$, and so by the area formula for Lipschitz functions

$$\begin{aligned}\textstyle\int_{-1}^{1} \sqrt{W_1(\boldsymbol{g}(t))}\,|\boldsymbol{g}'(t)|\,dt &\geq \textstyle\int_{-1}^{t_1} \sqrt{W_1(\boldsymbol{g}(t))}\,|\boldsymbol{g}'(t)|\,dt = \int_{-1}^{t_1} \sqrt{W(\boldsymbol{g}(t))}\,|\boldsymbol{g}'(t)|\,dt \\ &= \int_{\mathbb{R}^d} \sum_{t\in[-1,t_1]\cap \boldsymbol{g}^{-1}(\{\boldsymbol{y}\})} \sqrt{W(\boldsymbol{g}(t))}\,d\boldsymbol{y} \\ &= \int_{\mathbb{R}^d} \sum_{t\in[-1,t_1]\cap \boldsymbol{g}^{-1}(\{\boldsymbol{y}\})} \sqrt{W(\boldsymbol{y})}\,d\boldsymbol{y} \geq \textstyle\int_{r_0}^{r_1} h(r)\,dr > c_W.\end{aligned}$$

This shows that $f(\boldsymbol{b}) \geq c_W$.

The remaining of the proof is very similar to the one of Theorem 1.12. □

It can be shown that

$$\begin{aligned} c_W = \inf\Big\{ &\int_{-\infty}^{\infty} \left(W(\boldsymbol{g}(t)) + |\boldsymbol{g}'(t)|^2 \right) dt : \\ &\boldsymbol{g} \text{ piecewise } C^1 \text{ curve}, \boldsymbol{g}(-\infty) = \boldsymbol{a},\, \boldsymbol{g}(\infty) = \boldsymbol{b} \Big\}, \\ = \inf\Big\{ &\int_{-R}^{R} \left(W(\boldsymbol{g}(t)) + |\boldsymbol{g}'(t)|^2 \right) dt : \\ &R > 0,\, \boldsymbol{g} \text{ piecewise } C^1 \text{ curve}, \boldsymbol{g}(-R) = \boldsymbol{a},\, \boldsymbol{g}(R) = \boldsymbol{b} \Big\}. \end{aligned}$$

2.3. **The Anisotropic Case.** In the scalar case $d = 1$ the family of functionals (7) was generalized to

$$\int_\Omega \frac{1}{\varepsilon} f(\boldsymbol{x}, u, \varepsilon \nabla u)\, d\boldsymbol{x}$$

by Bouchitté [Bo1990], Owen [Ow1988], and Owen and Sternberg [OS1991], while the vectorial case $d \geq 1$ was considered by Barroso and Fonseca [BF1994], who studied the functional

$$\int_\Omega \left(\frac{1}{\varepsilon} W(\boldsymbol{u}) + \varepsilon \Phi^2(\boldsymbol{x}, \nabla \boldsymbol{u}) \right) d\boldsymbol{x}, \quad \boldsymbol{u} \in W^{1,2}\left(\Omega; \mathbb{R}^d\right).$$

Of particular importance in the case $d = 1$ is the functional

$$\mathcal{H}_\varepsilon(u) := \int_{\mathbb{R}^N} \left(\frac{1}{\varepsilon} W(u) + \varepsilon \Phi^2(\nabla u) \right) d\boldsymbol{x}$$

defined for $u \in W^{1,2}\left(\mathbb{R}^N\right)$ and $\int_{\mathbb{R}^N} u(\boldsymbol{x})\,d\boldsymbol{x} = m$, where $\Phi : \mathbb{R}^N \to \mathbb{R}$ is convex and positively homogeneous of degree one. We extend H_ε to $L^1\left(\mathbb{R}^N\right)$ by setting $H_\varepsilon(u) := +\infty$ if $u \notin W^{1,2}\left(\mathbb{R}^N\right)$ or if the constraint $\int_{\mathbb{R}^N} u(\boldsymbol{x})\,d\boldsymbol{x} = m$ is not satisfied.

By adapting the arguments developed in [BF1994], [Bo1990], [OS1991], it can be shown the Γ-limit of $\{H_\varepsilon\}$, is given by

$$\mathcal{H}_0(u) := c_W \, \mathrm{P}_\Phi(E) \tag{44}$$

if $u = b\chi_E + a\chi_{\Omega\setminus E}$, with E a set of finite perimeter satisfying (1), while $\mathcal{H}_0(u) := \infty$ otherwise. Here, c_W is the constant given in 21 and P_Φ is the Φ-*perimeter*, defined for every $E \subset \mathbb{R}^N$ with finite perimeter by

$$\mathrm{P}_\Phi(E) := \int_{\partial^* E} \Phi(\boldsymbol{\nu}_E)\, d\mathcal{H}^{N-1}, \tag{45}$$

where $\partial^* E$ is the reduced boundary of E, $\boldsymbol{\nu}_E$ is the measure theoretic outer unit normal of E.

It was established by Fonseca [Fo1991] and Fonseca and Müller [FMu1991] (see also the work of Taylor [Ta1971], [Ta1975], [Ta1978]) that the minimum of the problem

$$\min\left\{\mathrm{P}_\Phi(E) : E \text{ set of finite perimeter, } \mathcal{L}^N(E) = \frac{m-a}{b-a}\right\} \tag{46}$$

is uniquely attained (up to a translation) by an appropriate rescaling E_0 of the Wulff set

$$B_{\Phi^\circ} := \{x \in \mathbb{R}^n : \Phi^\circ(x) \le 1\}, \tag{47}$$

where Φ° is the polar function of Φ. More precisely $E_0 := rB_{\Phi^\circ}$ (up to a translation), where $r > 0$ is chosen so that $\mathcal{L}^N(E_\Phi) = \frac{m-a}{b-a}$. A key ingredient in the proof is the Brunn-Minkowski inequality (see [Gar2002])

$$(\mathcal{L}^N(A))^{1/N} + (\mathcal{L}^N(B))^{1/N} \le (\mathcal{L}^N(A+B))^{1/N}, \tag{48}$$

which holds for all Lebesgue measurable sets $A, B \subset \mathbb{R}^N$ such that $A + B$ is also Lebesgue measurable.

Functionals of the type (44) describe the surface energy of crystals and, since the fundamental work of Herring [He1951], they play a central role in many fields of physics, chemistry and materials science. If the dimension of the crystals is sufficiently small, then the leading morphological mechanism is driven by the minimization of surface energy.

2.4. **Solid-Solid Phase Transitions.** The corresponding problem for gradient vector fields, where in place of $\mathcal{F}_\varepsilon$ we introduce

$$\mathcal{I}_\varepsilon(\boldsymbol{u}) := \int_\Omega \left(\frac{1}{\varepsilon} W(\nabla \boldsymbol{u}) + \varepsilon |\nabla^2 \boldsymbol{u}|^2\right) d\boldsymbol{x}, \quad \boldsymbol{u} \in W^{2,2}\left(\Omega; \mathbb{R}^d\right), \tag{49}$$

arises naturally in the study of elastic solid-to-solid phase transitions [BJ1987] and [KM1994]. Here $\boldsymbol{u} : \Omega \to \mathbb{R}^d$ stands for the deformation. One of the main differences with the Modica–Mortola functional is that in the case of gradients, some geometrical compatibility conditions must exist between the wells. Indeed, the following Hadamard's compatibility condition holds.

Theorem 2.6. *Let $\Omega \subset \mathbb{R}^N$ be an open connected set, let $\mathbf{u} \in W^{1,\infty}\left(\Omega; \mathbb{R}^d\right)$ satisfy*

$$\nabla \mathbf{u}(x) = \chi_E(\mathbf{x})\boldsymbol{A} + (1 - \chi_E(\mathbf{x}))\boldsymbol{B}$$

for $\mathcal{L}^N$ a.e. $\mathbf{x} \in \Omega$, where $E \subset \Omega$ is a measurable set with $0 < \mathcal{L}^N(E) < \mathcal{L}^N(\Omega)$, and $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{d\times N}$. Then there exist $\boldsymbol{\nu} \in S^{N-1}$, $\boldsymbol{a}, \mathbf{u}_0 \in \mathbb{R}^d$, with $\boldsymbol{a} \cdot \mathbf{u}_0 = 0$, $\theta \in W^{1,\infty}(\Omega)$ with $\nabla\theta(\mathbf{x}) = \chi_E(\mathbf{x})\boldsymbol{\nu}$ for $\mathcal{L}^N$ a.e. $\mathbf{x} \in \Omega$, such that

$$\boldsymbol{A} - \boldsymbol{B} = \boldsymbol{a} \otimes \boldsymbol{\nu} \tag{50}$$

and

$$\mathbf{u}(\mathbf{x}) = \mathbf{u}_0 + \boldsymbol{B}\mathbf{x} + \theta(\mathbf{x})\,\boldsymbol{a} \tag{51}$$

for $\mathcal{L}^N$ *a.e.* $\mathbf{x} \in \Omega$.

Proof. Let $\boldsymbol{z}(\mathbf{x}) := \mathbf{u}(\mathbf{x}) - \boldsymbol{B}\mathbf{x}$, and let $\boldsymbol{C} := \boldsymbol{A} - \boldsymbol{B}$. Then $\nabla \boldsymbol{z}(\mathbf{x}) = \chi_E(\mathbf{x})\,\boldsymbol{C}$ for $\mathcal{L}^N$ a.e. $\mathbf{x} \in \Omega$, and since χ_E is not constant because $|E| > 0$, we may find $\varphi \in C_c^\infty(\Omega)$ such that

$$\int_E \nabla\varphi\, d\mathbf{x} \neq 0.$$

(Exercise). Define

$$\boldsymbol{\nu} := \frac{\int_E \nabla\varphi\, d\mathbf{x}}{\left|\int_E \nabla\varphi\, d\mathbf{x}\right|}.$$

We have

$$\begin{aligned} 0 &= \frac{1}{\left|\int_E \nabla\varphi\, d\mathbf{x}\right|} \int_\Omega \left(\frac{\partial z_i}{\partial x_j}\frac{\partial \varphi}{\partial x_k} - \frac{\partial z_i}{\partial x_k}\frac{\partial \varphi}{\partial x_j} \right) d\mathbf{x} \\ &= C_{ij}\nu_k - C_{ik}\nu_j, \end{aligned}$$

and so the vectors $C_{i\cdot} := (C_{i1}, \ldots, C_{iN})$ are paralell to $\boldsymbol{\nu}$, i.e., there exists $\boldsymbol{a} = (a_1, \ldots, a_d) \in \mathbb{R}^d$ such that $C_{i\cdot} = a_i \boldsymbol{\nu}$, and this proves (50).

Note that if $\boldsymbol{b} \in \mathbb{R}^d$ is orthogonal to $\boldsymbol{a}$, then

$$\nabla(\boldsymbol{z}(\mathbf{x}) \cdot \boldsymbol{b}) = \chi_E(\mathbf{x})(\boldsymbol{a} \cdot \boldsymbol{b})\,\boldsymbol{\nu} = 0,$$

therefore for any fixed $\mathbf{x}_0 \in \Omega$ the function $\boldsymbol{z} - \boldsymbol{z}(\mathbf{x}_0)$ is parallel to $\boldsymbol{a}$, hence there exists $\theta_1 \in W^{1,\infty}(\Omega)$ such that

$$\boldsymbol{z}(\mathbf{x}) - \boldsymbol{z}(\mathbf{x}_0) = \theta_1(\mathbf{x})\,\boldsymbol{a}$$

for $\mathcal{L}^N$ a.e. $\mathbf{x} \in \Omega$. Therefore,

$$\begin{aligned} \mathbf{u}(\mathbf{x}) &= \boldsymbol{z}(\mathbf{x}_0) + \boldsymbol{B}\mathbf{x} + \theta_1(\mathbf{x})\,\boldsymbol{a} \\ &= \mathbf{u}_0 + \boldsymbol{B}\mathbf{x} + \theta(\mathbf{x})\,\boldsymbol{a}, \end{aligned} \tag{52}$$

where

$$\mathbf{u}_0 := \boldsymbol{z}(\mathbf{x}_0) - \frac{(\boldsymbol{z}(\mathbf{x}_0) \cdot \boldsymbol{a})}{|\boldsymbol{a}|^2}\boldsymbol{a}, \quad \theta(\mathbf{x}) := \theta_1(\mathbf{x}) + \frac{(\boldsymbol{z}(\mathbf{x}_0) \cdot \boldsymbol{a})}{|\boldsymbol{a}|^2}.$$

Moreover, in view of (50) we have

$$\nabla\mathbf{u}(\mathbf{x}) = \boldsymbol{B} + \chi_E(\mathbf{x})\,\boldsymbol{a} \otimes \boldsymbol{\nu},$$

and in turn by (52)

$$\nabla\mathbf{u}(\mathbf{x}) = \boldsymbol{B} + \boldsymbol{a} \otimes \nabla\theta(\mathbf{x}).$$

We conclude that $\nabla\theta(\mathbf{x}) = \chi_E(\mathbf{x})\,\boldsymbol{\nu}$ for $\mathcal{L}^N$ a.e. $\mathbf{x} \in \Omega$. ∎

In the case of elastic solid-to-solid phase transitions, $\boldsymbol{A}, \boldsymbol{B}$ represent two variants of martensite. The Γ-convergence for the family $\{\mathcal{I}_\varepsilon\}$ in the case of two wells satisfying (50) was studied by Conti, Fonseca, and G.L. in [CFL2002]. This result was extended by Conti and Schweizer [CS2006] in dimension $N = 2$, who took into account frame-indifference, and assumed that $\{W = 0\} = SO(N)\boldsymbol{A} \cup SO(N)\boldsymbol{B}$, where $SO(N)$ is the set of rotations in $\mathbb{R}^N$, and by Chermisi and Conti [CC2010], who considered the case of multiple wells in any dimension and under frame-indifference.

More recently, Zwicknagl [Z2013] has studied several variants of a classical model of Kohn and Müller [KM1994] for the fine scale structure of twinning near an austenite-twinned-martensite interface.

2.5. **Nonlocal Functionals.** In [ABII1998] (see also [ABI1998]) Alberti and Bellettini studied the Γ-convergence of the family of nonlocal functionals defined by

$$\mathcal{J}_\varepsilon(u) := \frac{1}{\varepsilon}\int_\Omega W(u(\boldsymbol{x}))\,d\boldsymbol{x} + \frac{1}{4\varepsilon}\int_\Omega\int_\Omega J_\varepsilon(\boldsymbol{x}-\boldsymbol{y})\,(u(\boldsymbol{x})-u(\boldsymbol{y}))^2\,d\boldsymbol{x}d\boldsymbol{y} \tag{53}$$

for $u \in L^1(\Omega)$, where

$$J_\varepsilon(\boldsymbol{x}-\boldsymbol{y}) := \frac{1}{\varepsilon^N} J\left(\frac{\boldsymbol{x}}{\varepsilon}\right),$$

with $J : \mathbb{R}^N \to [0,\infty)$ an even integrable function satisfying $\int_{\mathbb{R}^N} J(\boldsymbol{x})\,|\boldsymbol{x}|\,d\boldsymbol{x} < \infty$, and W satisfies (H_1) and (H_2) with $a = 1$ and $b = -1$.

In equilibrium statistical mechanics functionals of the form (53) arise as free energies of continuum limits of Ising spin systems on lattices. In this context, u plays the role of a macroscopic magnetization density and J is a ferromagnetic Kac potential (see [ABCP1996]). It was proved in [ABII1998] that a compactness result similar to that of Theorem 1.6 holds, and that the family $\{\mathcal{J}_\varepsilon\}$ Γ-converges in $L^1(\Omega)$ to the functionalis given by

$$\mathcal{J}_0(u) := \int_{\partial^* E} \Phi(\boldsymbol{\nu}_E)\,d\mathcal{H}^{N-1},$$

if $u = \chi_E - \chi_{\Omega\setminus E}$, with E a set of finite perimeter, while $\mathcal{J}_0(u) := \infty$ otherwise, for an appropriate function Φ. We recall that $\partial^* E$ is the reduced boundary of E, $\boldsymbol{\nu}_E$ is the measure theoretic outer unit normal of E.

Note that the kernel J is assumed to be integrable and thus it excludes the classical seminorm for fractional Sobolev spaces $W^{s,p}$, $0 < s < 1$, introduced by Gagliardo [Ga1957] to characterize traces of functions in the Sobolev space $W^{1,p}(\Omega)$, $p > 1$, (see also [DNPV2012] and [Le2009]). This case was considered for $s = \frac{1}{2}$ and $N = 1$ by Alberti, Bouchitté, and Seppecher in [ABS1994] (see also [ABS1998]), who studied the functional

$$\mathcal{K}_\varepsilon(u) := \lambda_\varepsilon \int_I W(u)\,dx + \varepsilon \int_I\int_I \left|\frac{u(x)-u(y)}{x-y}\right|^2 dxdy,$$

where $\varepsilon\lambda_\varepsilon \to \ell \in (0,\infty)$ and $I \subset \mathbb{R}$ is a bounded interval. Under hypotheses (H_1) and (H_2) (with quadratic growth at infinity in place of linear growth), they proved that the family of functionals $\mathcal{K}_\varepsilon$ (extended to $L^1(I)$ by setting $+\infty$) Γ-converges to $2\ell\,(b-a)^2\,\mathrm{P}(E, I)$, where $E := \{x \in I : u(x) = b\}$, if $u \in BV(I; \{a, b\})$ and to ∞ otherwise. The case $s = \frac{1}{2}$ and $N \geq 1$ is contained in [ABS1998], althought it is not stated explicitly. Garroni and Palatucci [GP2006] extended the results in [ABS1994] to the functional

$$\frac{1}{\varepsilon}\int_I W(u)\,dx+\varepsilon^{p-2}\int_I\int_I\left|\frac{u(x)-u(y)}{x-y}\right|^p dxdy,$$

$p>2$. More recently Savin and Valdinoci [SV2012] considered the functional

$$\begin{aligned}\mathcal{X}_\varepsilon(u):=&\int_\Omega W(u)\,dx\\&+\frac{1}{2}\varepsilon^{2s}\int_\Omega\int_\Omega\frac{|u(x)-u(y)|^2}{|x-y|^{N+2s}}\,dxdy+\varepsilon^{2s}\int_{\mathbb{R}^N\setminus\Omega}\int_\Omega\left|\frac{u(x)-u(y)}{x-y}\right|^p dxdy,\end{aligned}$$

where $s\in(0,1)$ and dimension $N\geq 2$ and found the Γ-limit of the rescaled functionals

$$\begin{array}{ll}\varepsilon^{-2s}\mathcal{X}_\varepsilon(u) & \text{if } 0<s<\frac{1}{2},\\ |\varepsilon\log\varepsilon|^{-1}\mathcal{X}_\varepsilon(u) & \text{if } s=\frac{1}{2},\\ \varepsilon^{-1}\mathcal{X}_\varepsilon(u) & \text{if } \frac{1}{2}<s<1.\end{array}$$

2.6. **Higher Order Γ-Convergence.** A sequence of functionals $\mathcal{F}_\varepsilon:Y\to(-\infty,\infty]$, defined on an arbitrary metric space Y, has the *asymptotic development of order* k and we write

$$\mathcal{F}_\varepsilon\overset{\Gamma}{=}\mathcal{F}^{(0)}+\varepsilon\mathcal{F}^{(1)}+\cdots+\varepsilon^k\mathcal{F}^{(k)}+o\left(\varepsilon^k\right)$$

if $\mathcal{F}_\varepsilon\overset{\Gamma}{\to}\mathcal{F}^{(0)}$ and

$$\mathcal{F}_\varepsilon^{(i)}:=\frac{\mathcal{F}_\varepsilon^{(i-1)}-\inf_X\mathcal{F}^{(i-1)}}{\varepsilon}\overset{\Gamma}{\to}\mathcal{F}^{(i)}\tag{54}$$

for $i=1,\ldots,k$, where $\mathcal{F}_\varepsilon^{(0)}:=\mathcal{F}_\varepsilon$ (see [AB1993], [ABO1996]). The second order asymptotic development for the Modica–Mortola functional 7 is still an open problem. In [AB1993] Anzellotti and Baldo considered the case $N=1$ under the assumption that $\{t\in\mathbb{R}:\,W(t)=0\}=[a,c]\cup[d,b]$ and the constraint $\int_\Omega u(x)\,dx=m$ is replaced by boundary condition $u=g$ on $\partial\Omega$, while in the N-dimensional setting Anzellotti, Baldo, and Orlandi [ABO1996] studied the second order asymptotic development for 7 in the case in which W is the one-well potential $W(t)=t^2$ and again with the boundary condition $u=g$ on $\partial\Omega$. More recently Dal Maso, Fonseca, and G. L. [DMFL2013] have proved that the second order asymptotic development of 7 with the constraint $\int_\Omega u(x)\,dx=m$ is zero when W is of class C^1 but not C^2 near the wells and under the additional assumption that $u=a$ on $\partial\Omega$, which forces the phase $\{\boldsymbol{x}\in\Omega:\,u(\boldsymbol{x})=b\}$ of minimizers u to stay away from the boundary of Ω.

3. Related Problems and Other Applications of Γ-Convergence

3.1. **Allen–Cahn Equation.**

3.2. **Cahn–Hiliard Equation.**

3.3. **Dimension Reduction.**

3.4. **Discrete to Continuum.**

3.5. **Dislocations.**

3.6. **Ginsburgh Landau.**

3.7. **Homogenization.**

4. The Distance Function

The material in this section is taken from [DZ2001], [Fe1959], [Fo1984], [KP1981].

Given a nonempty set $E \subseteq \mathbb{R}^N$, let

$$f(\mathbf{x}) := \operatorname{dist}(\mathbf{x}, E), \quad \mathbf{x} \in \mathbb{R}^N,$$

where

$$\operatorname{dist}(\mathbf{x}, E) := \inf\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{y} \in E\}. \tag{55}$$

Proposition 4.1. *Let $E \subseteq \mathbb{R}^N$ be a nonempty set. Then f is Lipschitz continuous with Lipschitz constant less than or equal to one.*

Proof. If $\mathbf{y} \in E$, by (55) we have

$$\operatorname{dist}(\mathbf{x}, E) \leq \|\mathbf{x} - \mathbf{y}\| = \|\mathbf{x} - \mathbf{z} + \mathbf{z} - \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{z}\| + \|\mathbf{z} - \mathbf{y}\|.$$

Hence,

$$\operatorname{dist}(\mathbf{x}, E) - \|\mathbf{x} - \mathbf{z}\| \leq \|\mathbf{z} - \mathbf{y}\|$$

for all $\mathbf{y} \in E$, which shows that $\operatorname{dist}(\mathbf{x}, E) - \|\mathbf{x} - \mathbf{z}\|$ is a lower bound for the set $\{\|\mathbf{z} - \mathbf{y}\| : \mathbf{y} \in E\}$. Since the infimum is the greatest of all lower bounds, it follows that

$$\operatorname{dist}(\mathbf{x}, E) - \|\mathbf{x} - \mathbf{z}\| \leq \inf\{\|\mathbf{z} - \mathbf{y}\| : \mathbf{y} \in E\} = f(\mathbf{z}).$$

Hence,

$$f(\mathbf{x}) \leq f(\mathbf{z}) + \|\mathbf{z} - \mathbf{x}\|. \tag{56}$$

By interchanging the role of $\mathbf{x}$, $\mathbf{z}$, we have that

$$f(\mathbf{z}) \leq f(\mathbf{x}) + \|\mathbf{z} - \mathbf{x}\| = f(\mathbf{x}) + \|\mathbf{x} - \mathbf{z}\|,$$

or, equivalently,

$$-\|\mathbf{x} - \mathbf{z}\| \leq f(\mathbf{x}) - f(\mathbf{z}),$$

which together with (56) gives

$$-\|\mathbf{x} - \mathbf{z}\| \leq f(\mathbf{x}) - f(\mathbf{z}) \leq \|\mathbf{x} - \mathbf{z}\|,$$

or, equivalently,

$$|f(\mathbf{x}) - f(\mathbf{z})| \leq \|\mathbf{x} - \mathbf{z}\|.$$

□

Since f is Lipschitz continuous, it follows by Rademacher's theorem (see, e.g., Theorem 11.49 in [Le2009]) that f is differentiable for $\mathcal{L}^N$ a.e. $\mathbf{x} \in \mathbb{R}^N$.

For $\mathbf{x} \in \mathbb{R}^N$ define the set

$$\Pi_E(\mathbf{x}) := \left\{ \mathbf{y} \in \overline{E} : \, f(\mathbf{x}) = \|\mathbf{x} - \mathbf{y}\| \right\}.$$

Proposition 4.2. *Let $E \subseteq \mathbb{R}^N$ be a nonempty set. Then for every $\mathbf{x} \in \mathbb{R}^N$ the set $\Pi_E(\mathbf{x})$ is nonempty and compact. Moroever, if $\mathbf{x} \in \mathbb{R}^N \setminus \overline{E}$, then $\Pi_E(\mathbf{x}) \subseteq \partial E$, while if $\mathbf{x} \in \overline{E}$, then $\Pi_E(\mathbf{x}) = \{\mathbf{x}\}$.*

Proof. Let's prove that $\Pi_E(\mathbf{x})$ is nonempty. By (55) for every $n \in \mathbb{N}$ there exists $\mathbf{y}_n \in E$ such that

$$\operatorname{dist}(\mathbf{x}, E) \leq \|\mathbf{x} - \mathbf{y}_n\| < \operatorname{dist}(\mathbf{x}, E) + \frac{1}{n}. \tag{57}$$

Since E is nonempty, $\operatorname{dist}(\mathbf{x}, E) < \infty$, and so the sequence $\{\mathbf{y}_n\}$ is bounded. It follows by the Weierstrass theorem that there exist a subsequence of $\{\mathbf{y}_n\}$, not relabeled, such that $\mathbf{y}_n \to \mathbf{y}$ for some $\mathbf{y} \in \mathbb{R}^N$. Since $\mathbf{y}_n \in E$, we have that $\mathbf{y} \in \overline{E}$ and, by letting $n \to \infty$ in (57), we obtain that $\operatorname{dist}(\mathbf{x}, E) = \|\mathbf{x} - \mathbf{y}\|$. This shows that $\mathbf{y} \in \Pi_E(\mathbf{x})$. The remaining of the proof is left as an exercise. □

Exercise 4.3. *Let $E \subseteq \mathbb{R}^N$ be a nonempty set. Prove that for every $\mathbf{x} \in \mathbb{R}^N$,*

$$\operatorname{dist}(\mathbf{x}, E) = \operatorname{dist}\left(\mathbf{x}, \overline{E}\right).$$

Proposition 4.4. *Let $E \subseteq \mathbb{R}^N$ be a nonempty set. Let $\boldsymbol{v} \in \partial B(\mathbf{0}, 1)$ be a direction and let $\mathbf{x} \in \mathbb{R}^N$. Then there exists*

$$\frac{\partial^+ f^2(\mathbf{x})}{\partial \mathbf{v}} := \lim_{t \to 0^+} \frac{f^2(\mathbf{x} + t\mathbf{v}) - f^2(\mathbf{x})}{t} = 2 \min_{\mathbf{y} \in \Pi_E(\mathbf{x})} \left[(\mathbf{x} - \mathbf{y}) \cdot \mathbf{v} \right].$$

Proof. Let $\mathbf{y} \in \Pi_E(\mathbf{x})$ and $\mathbf{y}_t \in \Pi_E(\mathbf{x} + t\mathbf{v})$. Then, by (55),

$$\|\mathbf{x} + t\mathbf{v} - \mathbf{y}_t\| = f(\mathbf{x} + t\mathbf{v}) \leq \|\mathbf{x} + t\mathbf{v} - \mathbf{y}\|$$

and so

$$\begin{aligned} \frac{f^2(\mathbf{x} + t\mathbf{v}) - f^2(\mathbf{x})}{t} &= \frac{\|\mathbf{x} + t\mathbf{v} - \mathbf{y}_t\|^2 - \|\mathbf{x} - \mathbf{y}\|^2}{t} \\ &\leq \frac{\|\mathbf{x} + t\mathbf{v} - \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2}{t} = t + 2(\mathbf{x} - \mathbf{y}) \cdot \mathbf{v}. \end{aligned}$$

It follows that

$$\limsup_{t \to 0^+} \frac{f^2(\mathbf{x} + t\mathbf{v}) - f^2(\mathbf{x})}{t} \leq 2(\mathbf{x} - \mathbf{y}) \cdot \mathbf{v}$$

for all $\mathbf{p} \in \Pi_E(\mathbf{x})$, which implies that

$$\limsup_{t \to 0^+} \frac{f^2(\mathbf{x} + t\mathbf{v}) - f^2(\mathbf{x})}{t} \leq 2 \min_{\mathbf{y} \in \Pi_E(\mathbf{x})} \left[(\mathbf{x} - \mathbf{y}) \cdot \mathbf{v} \right].$$

On the other hand, let $t_n \to 0^+$ be such that

$$\liminf_{t\to 0^+} \frac{f^2(\mathbf{x}+t\mathbf{v}) - f^2(\mathbf{x})}{t} = \lim_{n\to\infty} \frac{f^2(\mathbf{x}+t_n\mathbf{v}) - f^2(\mathbf{x})}{t_n}$$

and let $\mathbf{y}_n \in \Pi_E(\mathbf{x}+t_n\mathbf{v})$. Then by Proposition 4.1,

$$\begin{aligned}\|\mathbf{x}-\mathbf{y}_n\| &\le \|\mathbf{x}+t_n\mathbf{v}-\mathbf{y}_n\| + \|t_n\mathbf{v}\| \\ &= f(\mathbf{x}+t_n\mathbf{v}) + t_n \le f(\mathbf{x}) + 2t_n.\end{aligned}$$

It follows that the sequence $\{\mathbf{y}_n\}$ is bounded, and so, up to a subsequence, not relabeled, $\mathbf{y}_n \to \mathbf{y}_0 \in \overline{E}$. In turn, since f is continuous,

$$f(\mathbf{x}) = \lim_{n\to\infty} f(\mathbf{x}+t_n\mathbf{v}) = \lim_{n\to\infty} \|\mathbf{x}+t_n\mathbf{v}-\mathbf{y}_n\| = \|\mathbf{x}-\mathbf{y}_0\|.$$

This implies that $\mathbf{y}_0 \in \Pi_E(\mathbf{x})$. Hence, by (55),

$$\|\mathbf{x}-\mathbf{y}_0\| = f(\mathbf{x}) \le \|\mathbf{x}-\mathbf{y}_n\|,$$

and so

$$\begin{aligned}\frac{f^2(\mathbf{x}+t_n\mathbf{v}) - f^2(\mathbf{x})}{t_n} &= \frac{\|\mathbf{x}+t_n\mathbf{v}-\mathbf{y}_n\|^2 - \|\mathbf{x}-\mathbf{y}_0\|^2}{t_n} \\ &\ge \frac{\|\mathbf{x}+t_n\mathbf{v}-\mathbf{y}_n\|^2 - \|\mathbf{x}-\mathbf{y}_n\|^2}{t_n} \\ &= t_n + 2(\mathbf{x}-\mathbf{y}_n)\cdot\mathbf{v}.\end{aligned}$$

Letting $n \to \infty$ we get

$$\liminf_{t\to 0^+} \frac{f^2(\mathbf{x}+t\mathbf{v}) - f^2(\mathbf{x})}{t} \ge 2(\mathbf{x}-\mathbf{y}_0)\cdot\mathbf{v} \ge 2 \min_{\mathbf{y}\in\Pi_E(\mathbf{x})} [(\mathbf{x}-\mathbf{y})\cdot\mathbf{v}],$$

which concludes the proof. □

Let F be the set of points $\mathbf{x} \in \mathbb{R}^N$ for which $\Pi_E(\mathbf{x})$ is a singleton and define the function and define

$$\boldsymbol{p}(\mathbf{x}) := \mathbf{y},$$

where $\Pi_E(\mathbf{x}) = \{\mathbf{y}\}$.

Proposition 4.5. *Let $E \subseteq \mathbb{R}^N$ be a nonempty set. Then the function $\boldsymbol{p} : F \to \mathbb{R}^N$ is continuous.*

Proof. Let $\mathbf{x} \in F$. If P is not continuous at $\mathbf{x}$, then we can find $\varepsilon > 0$ and a sequence $\{\mathbf{x}_n\} \subset F$ converging to $\mathbf{x}$ such that

$$\|\boldsymbol{p}(\mathbf{x}) - \boldsymbol{p}(\mathbf{x}_n)\| \ge \varepsilon \tag{58}$$

for all n. Then

$$\begin{aligned}\|\mathbf{x}-\boldsymbol{p}(\mathbf{x}_n)\| &\le \|\mathbf{x}_n-\boldsymbol{p}(\mathbf{x}_n)\| + \|\mathbf{x}_n-\mathbf{x}\| \\ &= f(\mathbf{x}_n) + \|\mathbf{x}_n-\mathbf{x}\| \le f(\mathbf{x}) + 2\|\mathbf{x}_n-\mathbf{x}\|.\end{aligned}$$

It follows that the sequence $\{\boldsymbol{p}(\mathbf{x}_n)\}$ is bounded, and so, up to a subsequence, not relabeled, $\boldsymbol{p}(\mathbf{x}_n) \to \mathbf{y} \in \overline{E}$. In turn, since f is continuous,

$$f(\mathbf{x}) = \lim_{n\to\infty} f(\mathbf{x}_n) = \lim_{n\to\infty} \|\mathbf{x}_n - \boldsymbol{p}(\mathbf{x}_n)\| = \|\mathbf{x} - \mathbf{y}\|.$$

This implies that $\mathbf{y} = \boldsymbol{p}(\mathbf{x})$. This contradicts (58) and completes the proof. □

Exercise 4.6. *Let $g : \mathbb{R}^N \to \mathbb{R}$ be Lipschitz on compact sets and let $\mathbf{x} \in \mathbb{R}^N$. Assume that there exist $\frac{\partial^+ g(\mathbf{x})}{\partial \mathbf{v}}$ for every direction $\boldsymbol{v}$ and $\boldsymbol{b} \in \mathbb{R}^N$ such that*

$$\frac{\partial^+ g(\mathbf{x})}{\partial \mathbf{v}} = \boldsymbol{b} \cdot \boldsymbol{v}$$

for all $\boldsymbol{v}$. Prove that g is differentiable at $\mathbf{x}$.

Proposition 4.7. *Let $E \subseteq \mathbb{R}^N$ be a nonempty set. Then f^2 is differentiable at $\mathbf{x} \in \mathbb{R}^N$ if and only if $\mathbf{x}$ belongs to F, with*

(59) $$\nabla f^2(\mathbf{x}) = 2(\boldsymbol{p}(\mathbf{x}) - \mathbf{x}).$$

Moreover, the partial derivatives of f^2 are continuous in F.

Proof. Assume $\mathbf{x}$ belongs to F. Then by Proposition 4.4 and the fact that $\Pi_E(\mathbf{x}) = \{\boldsymbol{p}(\mathbf{x})\}$,

$$\frac{\partial^+ f^2(\mathbf{x})}{\partial \mathbf{v}} = 2(\mathbf{x} - \boldsymbol{p}(\mathbf{x})) \cdot \mathbf{v}$$

for every direction $\boldsymbol{v}$. It follows by Exercise 4.6 that f^2 is differentiable at $\mathbf{x}$, with

$$\nabla f^2(\mathbf{x}) = 2(\boldsymbol{p}(\mathbf{x}) - \mathbf{x}).$$

Moreover, its partial derivatives are continuous by Proposition 4.5.

Conversely, assume that f^2 is differentiable at $\mathbf{x}$. Then by standard properties of differentiation and by Proposition 4.4,

$$\nabla f^2(\mathbf{x}) \cdot \mathbf{v} = \frac{\partial^+ f^2(\mathbf{x})}{\partial \mathbf{v}} = 2 \min_{\mathbf{y} \in \Pi_E(\mathbf{x})} (\mathbf{x} - \mathbf{y}) \cdot \mathbf{v}$$

for all $\mathbf{v}$. We claim that this implies that $\Pi_E(\mathbf{x})$ is a singleton. If $\mathbf{x} \in \overline{E}$, then $\Pi_E(\mathbf{x}) = \{\mathbf{x}\}$ and so there is nothing to prove. Thus, assume that $\mathbf{x} \in \mathbb{R}^N \setminus \overline{E}$ and let $\mathbf{y}_0 \in \Pi_E(\mathbf{x})$. Then $\mathbf{x} - \mathbf{y}_0 \neq \mathbf{0}$. Let $\mathbf{v}_0 := -\frac{\mathbf{x}-\mathbf{y}_0}{\|\mathbf{x}-\mathbf{y}_0\|}$. Then for every $\mathbf{y} \in \Pi_E(\mathbf{x})$, with $\mathbf{y} \neq \mathbf{y}_0$,

$$-\|\mathbf{x} - \mathbf{y}_0\| = (\mathbf{x} - \mathbf{y}_0) \cdot \mathbf{v}_0 < (\mathbf{x} - \mathbf{y}) \cdot \mathbf{v}_0.$$

Hence,

$$\nabla f^2(\mathbf{x}) \cdot \mathbf{v}_0 = 2 \min_{\mathbf{y} \in \Pi_E(\mathbf{x})} (\mathbf{x} - \mathbf{y}) \cdot \mathbf{v}_0 = -\|\mathbf{x} - \mathbf{y}_0\| = -f(\mathbf{x}).$$

Since $f(\mathbf{x}) > 0$, we have that $f = \sqrt{f^2}$ is differentiable at $\mathbf{x}$ with

(60) $$\nabla f(\mathbf{x}) = \frac{1}{f(\mathbf{x})} \nabla f^2(\mathbf{x}).$$

Hence,

$$\nabla f(\mathbf{x}) \cdot \mathbf{v}_0 = \frac{1}{f(\mathbf{x})} \nabla f^2(\mathbf{x}) \cdot \mathbf{v}_0 = -1$$

On the other hand, since f is Lipschitz continuous with Lipschitz constant at most 1, we have that

$$\|\nabla f(\mathbf{x})\| \leq 1.$$

It follows that $\|\nabla f(\mathbf{x})\| = 1$ and that $\nabla f(\mathbf{x}) = -\mathbf{v}_0$. In turn, $\Pi_E(\mathbf{x}) = \{\mathbf{y}_0\}$.

The last part of the statement follows from (59) and Proposition 4.5. ☐

Corollary 4.8. *Let $E \subseteq \mathbb{R}^N$ be a nonempty set. Let $\mathbf{x} \in \mathbb{R}^N \setminus \overline{E}$. Then f is differentiable at $\mathbf{x}$ if and only if $\mathbf{x}$ belongs to F, with*

$$\nabla f(\mathbf{x}) = \frac{\boldsymbol{p}(\mathbf{x}) - \mathbf{x}}{f(\mathbf{x})} = \frac{\boldsymbol{p}(\mathbf{x}) - \mathbf{x}}{\|\boldsymbol{p}(\mathbf{x}) - \mathbf{x}\|}. \tag{61}$$

On the other hand, $\nabla f \equiv \mathbf{0}$ in E°, while f is differentiable at $\mathcal{L}^N$ a.e. $\mathbf{x} \in \partial E$, with $\nabla f(\mathbf{x}) = \mathbf{0}$. Moreover, the partial derivatives of f are continuous in $F \setminus \partial E$.

Proof. Let $\mathbf{x} \in \mathbb{R}^N \setminus \overline{E}$. If f is differentiable at $\mathbf{x}$, then so is f^2 and so $\mathbf{x}$ belongs to F by Proposition 4.7. On the other hand, if $\mathbf{x} \in F$, then f^2 is differentiable at $\mathbf{x}$ again by Proposition 4.7. Since $f(\mathbf{x}) > 0$, it follows that f is differentiable at $\mathbf{x}$ and formula (61) holds by (59) and (60).

Since f is Lipschitz, by the Rademacher's theorem, f is differentiable at $\mathcal{L}^N$ a.e. $\mathbf{x} \in \mathbb{R}^N$, and so, in particular, f is differentiable at $\mathcal{L}^N$ a.e. $\mathbf{x} \in \partial E$. Let $\mathbf{x} \in \partial E$ be such that f is differentiable at $\mathbf{x}$. Since $f(\mathbf{x}) = 0$, at any direction $\mathbf{v}$ we have that

$$\nabla f(\mathbf{x}) \cdot \mathbf{v} = \frac{\partial^+ f(\mathbf{x})}{\partial \mathbf{v}} = \lim_{t \to 0^+} \frac{f(\mathbf{x} + t\mathbf{v}) - f(\mathbf{x})}{t} = \lim_{t \to 0^+} \frac{f(\mathbf{x} + t\mathbf{v})}{t} \geq 0.$$

This implies that $\nabla f(\mathbf{x}) = \mathbf{0}$, since otherwise, we could take $\mathbf{v} := -\nabla f(\mathbf{x}) / \|\nabla f(\mathbf{x})\|$ and obtain a contradiction. ☐

Note that in general we cannot expect f to be of class C^1 across ∂E, since $\|\nabla f(\mathbf{x})\| = 1$ in $\left(\mathbb{R}^N \setminus \overline{E}\right) \cap E$, while $f \equiv 0$ in $\overline{E}$.

5. The Signed Distance Function

Given a set $E \subset \mathbb{R}^N$ with nonempty boundary, the *signed distance function* of E is defined by

$$\mathrm{d}_E(\boldsymbol{x}) := \begin{cases} \operatorname{dist}(\boldsymbol{x}, \partial E) & \text{if } \boldsymbol{x} \in E, \\ -\operatorname{dist}(\boldsymbol{x}, \partial E) & \text{if } \boldsymbol{x} \in \mathbb{R}^N \setminus E. \end{cases}$$

Exercise 5.1. *Given a set $E \subset \mathbb{R}^N$ with nonempty boundary, prove that*

$$\mathrm{d}_E(\boldsymbol{x}) = \begin{cases} \operatorname{dist}\left(\boldsymbol{x}, \overline{\mathbb{R}^N \setminus E}\right) & \text{if } \boldsymbol{x} \in E, \\ -\operatorname{dist}\left(\boldsymbol{x}, \overline{E}\right) & \text{if } \boldsymbol{x} \in \mathbb{R}^N \setminus E. \end{cases}$$

Proposition 5.2. *Let $E \subset \mathbb{R}^N$ be a set with nonempty boundary. Then d_E is Lipschitz continuous with Lipschitz constant less than or equal to one.*

Proof. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$. If $\mathbf{x}, \mathbf{y} \in \overline{E}$ or $\mathbf{x}, \mathbf{y} \in \overline{\mathbb{R}^N \setminus E}$, then

$$|\mathrm{d}_E(\boldsymbol{x}) - \mathrm{d}_E(\boldsymbol{y})| = |\mathrm{dist}(\boldsymbol{x}, \partial E) - \mathrm{dist}(\boldsymbol{y}, \partial E)| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$$

by Proposition 4.1. If $\mathbf{x} \in E^\circ$ and $\mathbf{y} \in \overline{\mathbb{R}^N \setminus E}$, then $\mathrm{dist}(\boldsymbol{x}, \partial E) > 0$ and $B(\boldsymbol{x}, \mathrm{dist}(\boldsymbol{x}, \partial E)) \subset E^\circ$. Hence the point

$$\boldsymbol{z} := \mathbf{x} + \mathrm{dist}(\boldsymbol{x}, \partial E) \frac{\mathbf{y} - \mathbf{x}}{\|\boldsymbol{y} - \boldsymbol{x}\|}$$

belongs to $\overline{E}$. Since $\mathbf{y} \in \overline{\mathbb{R}^N \setminus E}$ It follows that

$$\begin{aligned}\mathrm{dist}(\boldsymbol{y}, \partial E) = \mathrm{dist}\left(\boldsymbol{y}, \overline{E}\right) \leq \|\boldsymbol{y} - \boldsymbol{z}\| &= \left\|(\boldsymbol{y} - \mathbf{x})\left(1 - \frac{\mathrm{dist}(\boldsymbol{x}, \partial E)}{\|\boldsymbol{y} - \boldsymbol{x}\|}\right)\right\| \\ &= \|\boldsymbol{y} - \mathbf{x}\| - \mathrm{dist}(\boldsymbol{x}, \partial E),\end{aligned}$$

and so

$$|\mathrm{d}_E(\boldsymbol{x}) - \mathrm{d}_E(\boldsymbol{y})| = \mathrm{dist}(\boldsymbol{y}, \partial E) + \mathrm{dist}(\boldsymbol{x}, \partial E) \leq \|\boldsymbol{y} - \mathbf{x}\|.$$

The case $\mathbf{x} \in \overline{E}$ and $\mathbf{y} \in \left(\mathbb{R}^N \setminus E\right)^\circ$ is similar. □

Using the notation of the previous section, we recall that for $\mathbf{x} \in \mathbb{R}^N$,

$$\Pi_{\partial E}(\mathbf{x}) := \{\mathbf{y} \in \partial E : \mathrm{dist}(\mathbf{x}, \partial E) = \|\mathbf{x} - \mathbf{y}\|\}$$

and we set

$$F := \left\{\mathbf{x} \in \mathbb{R}^N : \Pi_{\partial E}(\mathbf{x}) \text{ is a singleton}\right\},$$

and for $\mathbf{x} \in F$ we write $\Pi_{\partial E}(\mathbf{x}) = \{\boldsymbol{p}(\mathbf{x})\}$.

Proposition 5.3. *Let $E \subset \mathbb{R}^N$ be a set with nonempty boundary. Then d_E^2 is differentiable at $\mathbf{x} \in \mathbb{R}^N$ if and only if $\mathbf{x}$ belongs to F, with*

$$\nabla \mathrm{d}_E^2(\mathbf{x}) = 2(\boldsymbol{p}(\mathbf{x}) - \mathbf{x}).$$

Moreover, the partial derivatives of d_E^2 are continuous in F.

Proof. Since $\mathrm{d}_E^2 = \mathrm{dist}(\cdot, \partial E)$, the result follows from Proposition 4.7. □

Proposition 5.4. *Let $E \subset \mathbb{R}^N$ be a set with nonempty boundary. If $\mathbf{x} \in \mathbb{R}^N \setminus \partial E$, then d_E is differentiable at $\mathbf{x}$ if and only if $\mathbf{x}$ belongs to F, with*

$$\nabla \mathrm{d}_E(\mathbf{x}) = \frac{(\boldsymbol{p}(\mathbf{x}) - \mathbf{x})}{\mathrm{d}_E(\mathbf{x})} = \pm\frac{(\boldsymbol{p}(\mathbf{x}) - \mathbf{x})}{\|\boldsymbol{p}(\mathbf{x}) - \mathbf{x}\|}. \tag{62}$$

On the other hand, d_E is differentiable at $\mathcal{L}^N$ a.e. $\mathbf{x} \in \partial E$ with $\nabla \mathrm{d}_E(\mathbf{x}) = \mathbf{0}$. Moreover, the partial derivatives of d_E are continuous in $F \setminus \partial E$.

Proof. If $\mathbf{x} \in E^\circ$, then in a neighborhood of $\mathbf{x}$, $\mathrm{d}_E = \mathrm{dist}\left(\cdot, \mathbb{R}^N \setminus E\right)$, and so we can apply Corollary 4.8 to conclude that d_E is differentiable at $\mathbf{x}$ if and only if $\mathbf{x}$ belongs to F, with

$$\nabla \mathrm{d}_E(\mathbf{x}) = \frac{\boldsymbol{p}(\mathbf{x}) - \mathbf{x}}{\mathrm{d}_E(\mathbf{x})} = \frac{\boldsymbol{p}(\mathbf{x}) - \mathbf{x}}{\|\boldsymbol{p}(\mathbf{x}) - \mathbf{x}\|}.$$

On the other hand, if $\mathbf{x} \in \mathbb{R}^N \setminus \overline{E}$, then in a neighborhood of $\mathbf{x}$, $\mathrm{d}_E = -\operatorname{dist}(\cdot, E)$, and so we can apply Corollary 4.8 to conclude that d_E is differentiable at $\mathbf{x}$ if and only if $\mathbf{x}$ belongs to F, with

$$\begin{aligned}\nabla \mathrm{d}_E(\mathbf{x}) = -\nabla \operatorname{dist}(\mathbf{x}, E) &= -\frac{\boldsymbol{p}(\mathbf{x}) - \mathbf{x}}{\operatorname{dist}(\mathbf{x}, E)} \\ &= \frac{\boldsymbol{p}(\mathbf{x}) - \mathbf{x}}{\mathrm{d}_E(\mathbf{x})} = -\frac{\boldsymbol{p}(\mathbf{x}) - \mathbf{x}}{\|\boldsymbol{p}(\mathbf{x}) - \mathbf{x}\|}.\end{aligned}$$

Since d_E is Lipschitz, by the Rademacher's theorem, d_E is differentiable at $\mathcal{L}^N$ a.e. $\mathbf{x} \in \mathbb{R}^N$, and so, in particular, d_E is differentiable at $\mathcal{L}^N$ a.e. $\mathbf{x} \in \partial E$. Let $\mathbf{x} \in \partial E$ be such that d_E and $\operatorname{dist}(\cdot, \partial E)$ are differentiable at $\mathbf{x}$. Since $\mathrm{d}_E(\mathbf{x}) = \operatorname{dist}(\mathbf{x}, \partial E) = 0$, we have that

$$\begin{aligned}\left|\frac{\mathrm{d}_E(\mathbf{x} + t\mathbf{e}_i) - \mathrm{d}_E(\mathbf{x})}{t}\right| &= \left|\frac{\mathrm{d}_E(\mathbf{x} + t\mathbf{e}_i)}{t}\right| \\ &= \left|\frac{\operatorname{dist}(\mathbf{x} + t\mathbf{e}_i, \partial E)}{t}\right| = \left|\frac{\operatorname{dist}(\mathbf{x} + t\mathbf{e}_i, \partial E) - \operatorname{dist}(\mathbf{x}, \partial E)}{t}\right| \to 0\end{aligned}$$

as $t \to 0^+$ by Corollary 4.8. This implies that $\frac{\partial \mathrm{d}_E}{\partial x_i}(\mathbf{x}) = 0$. ∎

Theorem 5.5. *Assume that $V \subset \mathbb{R}^N$ is an open with compact boundary of class C^k, $k \geq 2$. Then there exists an open set U containing ∂V such d_V is of class $C^k(U \setminus \partial V)$.*

Proof. **Step 1:** For every $\boldsymbol{x} \in \partial V$ there exist a ball $B(\boldsymbol{x}, r_{\boldsymbol{x}})$, local coordinates $\boldsymbol{y} = (\boldsymbol{y}', y_N) \in \mathbb{R}^{N-1} \times \mathbb{R}$ such that $\boldsymbol{x}$ corresponds to $\boldsymbol{y} = \mathbf{0}$ and a function g of class C^k such that $g(\mathbf{0}) = 0$, $\frac{\partial g}{\partial y_i}(\mathbf{0}) = 0$ for all $i = 1, \ldots, N-1$, and

$$\begin{aligned}V \cap B(\boldsymbol{x}, r_{\boldsymbol{x}}) &= \{\boldsymbol{y} \in B(\mathbf{0}, r_{\boldsymbol{x}}) : \, y_N < g(\boldsymbol{y}')\}, \\ \partial V \cap B(\boldsymbol{x}, r_{\boldsymbol{x}}) &= \{\boldsymbol{y} \in B(\mathbf{0}, r_{\boldsymbol{x}}) : \, y_N = g(\boldsymbol{y})\}.\end{aligned}$$

In what follows we use local coordinates and we set

$$\nabla' := \left(\frac{\partial}{\partial y_1}, \ldots, \frac{\partial}{\partial y_{N-1}}\right).$$

Let $0 < s_{\boldsymbol{x}} < r_{\boldsymbol{x}}$ be so small that

$$\|\nabla' g(\boldsymbol{y}') - \nabla' g(\boldsymbol{z}')\|_{N-1} \leq L_{\boldsymbol{x}} \|\boldsymbol{y}' - \boldsymbol{z}'\|_{N-1} \tag{63}$$

for all $\boldsymbol{y}', \boldsymbol{z}' \in B_{N-1}(\mathbf{0}, s_{\boldsymbol{x}})$ and for some constant $L_{\boldsymbol{x}} > 0$. If $\boldsymbol{z} \in B\left(\mathbf{0}, \frac{1}{2}s_{\boldsymbol{x}}\right)$, then the points of closest distance in ∂V will be in $\partial V \cap B(\mathbf{0}, s_{\boldsymbol{x}})$, and will be found by minimizing the function

$$h(\boldsymbol{y}') = \sum_{i=1}^{N-1} (y_i - z_i)^2 + (g(\boldsymbol{y}') - z_N)^2.$$

Hence,

$$0 = \frac{\partial h}{\partial y_i}(\boldsymbol{y}') = 2(y_i - z_i) + 2(g(\boldsymbol{y}') - z_N)\frac{\partial g}{\partial y_i}(\boldsymbol{y}') \tag{64}$$

for all $i = 1, \ldots, N-1$. It follows that

$$\begin{aligned}
\mathrm{d}_V^2(\boldsymbol{z}) &= \sum_{i=1}^{N-1} (y_i - z_i)^2 + (g(\boldsymbol{y}') - z_N)^2 \\
&= \sum_{i=1}^{N-1} \left(-(g(\boldsymbol{y}') - z_N) \frac{\partial g}{\partial y_i}(\boldsymbol{y}') \right)^2 + (g(\boldsymbol{y}') - z_N)^2 \\
&= (g(\boldsymbol{y}') - z_N)^2 \left(1 + \|\nabla' g(\boldsymbol{y}')\|_{N-1}^2 \right).
\end{aligned}$$

Note that if $\boldsymbol{z} \in V$, then $\mathrm{d}_V(\boldsymbol{z}) > 0$ and $g(\boldsymbol{y}') - z_N > 0$, while if $\boldsymbol{z} \in \mathbb{R}^N \setminus V$, then $\mathrm{d}_V(\boldsymbol{z}) < 0$, while $g(\boldsymbol{y}') - z_N < 0$. Thus,

$$\mathrm{d}_V(\boldsymbol{z}) = (g(\boldsymbol{y}') - z_N) \sqrt{1 + \|\nabla' g(\boldsymbol{y}')\|_{N-1}^2}.$$

Hence,

$$z_N = g(\boldsymbol{y}') + \mathrm{d}_V(\boldsymbol{z}) \frac{1}{\sqrt{1 + \|\nabla' g(\boldsymbol{y}')\|_{N-1}^2}}. \tag{65}$$

In turn, from (64),

$$z_i = y_i + \mathrm{d}_V(\boldsymbol{z}) \frac{\frac{\partial g}{\partial y_i}(\boldsymbol{y}')}{\sqrt{1 + \|\nabla' g(\boldsymbol{y}')\|_{N-1}^2}} \tag{66}$$

for all $i = 1, \ldots, N-1$. We now show that there can only be one such point $\boldsymbol{y}'$. By (63), for all $\boldsymbol{y}', \boldsymbol{w}' \in \overline{B_{N-1}(\boldsymbol{0}, s_{\boldsymbol{x}})}$,

$$\begin{aligned}
&\left\| \frac{\nabla' g(\boldsymbol{y}')}{\sqrt{1 + \|\nabla' g(\boldsymbol{y}')\|_{N-1}^2}} - \frac{\nabla' g(\boldsymbol{w}')}{\sqrt{1 + \|\nabla' g(\boldsymbol{y}')\|_{N-1}^2}} \right\|_{N-1} \\
&= \left\| \frac{\nabla' g(\boldsymbol{y}') - \nabla' g(\boldsymbol{w}')}{\sqrt{1 + \|\nabla' g(\boldsymbol{y}')\|_{N-1}^2}} + \nabla' g(\boldsymbol{w}') \left(\frac{1}{\sqrt{1 + \|\nabla' g(\boldsymbol{y}')\|_{N-1}^2}} - \frac{1}{\sqrt{1 + \|\nabla' g(\boldsymbol{w}')\|_{N-1}^2}} \right) \right\|_{N-1} \\
&\leq \left(\|\nabla' g(\boldsymbol{y}') - \nabla' g(\boldsymbol{w}')\|_{N-1} + \left| \sqrt{1 + \|\nabla' g(\boldsymbol{y}')\|_{N-1}^2} - \sqrt{1 + \|\nabla' g(\boldsymbol{w}')\|_{N-1}^2} \right| \right) \\
&\leq 2 \|\nabla' g(\boldsymbol{y}') - \nabla' g(\boldsymbol{w}')\|_{N-1} \leq 2L_{\boldsymbol{x}} \|\boldsymbol{y}' - \boldsymbol{w}'\|_{N-1}.
\end{aligned}$$

Consider the open set

$$U_{\boldsymbol{x}} := \left\{ \boldsymbol{z} \in B\left(\boldsymbol{0}, \frac{1}{2} s_{\boldsymbol{x}} \right) : |\mathrm{d}_V(\mathbf{z})| < \frac{1}{2L_{\boldsymbol{x}}} \right\}. \tag{67}$$

Assume that for $\boldsymbol{z} \in U_{\boldsymbol{x}}$ there exist two points $\boldsymbol{y}', \boldsymbol{w}' \in \overline{B_{N-1}(\mathbf{0}, s_{\boldsymbol{x}})}$ satisfying (65) and (66). Then by (66),

$$\|\boldsymbol{y}' - \boldsymbol{w}'\|_{N-1} \leq |\mathrm{d}_V(\mathbf{z})| \left\| \frac{\nabla' g(\boldsymbol{y}')}{\sqrt{1 + \|\nabla' g(\boldsymbol{y}')\|_{N-1}^2}} - \frac{\nabla' g(\boldsymbol{w}')}{\sqrt{1 + \|\nabla' g(\boldsymbol{y}')\|_{N-1}^2}} \right\|_{N-1} \leq \frac{1}{2} \|\boldsymbol{y}' - \boldsymbol{w}'\|_{N-1},$$

which implies that $\boldsymbol{y}' = \boldsymbol{w}'$. This shows that for every $\boldsymbol{z} \in U_{\boldsymbol{x}}$ there exists only one $\boldsymbol{y}' \in \overline{B_{N-1}(\mathbf{0}, s_{\boldsymbol{x}})}$ satisfying (65) and (66). Define $\boldsymbol{p}(\mathbf{z}) := (\boldsymbol{y}', g(\boldsymbol{y}'))$. Then, $\Pi_{\partial V}(\mathbf{z}) = \{\boldsymbol{p}(\mathbf{z})\}$. Thus, by Proposition 5.4, the function d_V is of class $C^1(U_{\boldsymbol{x}} \setminus \partial V)$.

Since the family of open sets $\{U_{\boldsymbol{x}}\}_{\boldsymbol{x} \in \partial V}$ covers the compact set ∂V, there exists a finite number of points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_\ell \in \partial V$ such that

$$\partial V \subset \bigcup_{i=1}^{\ell} U_{\boldsymbol{x}_i} := U.$$

Step 2: To conclude the proof, it remains to show that d_V is of class $C^k(U \setminus \partial V)$. We use the implicit function theorem. Consider the function

$$\boldsymbol{k}(\boldsymbol{y}, \boldsymbol{z}) := \boldsymbol{y} - \boldsymbol{z} + \mathrm{d}_V(\mathbf{z}) \frac{(\nabla' g(\boldsymbol{y}'), 1)}{\|(\nabla' g(\boldsymbol{y}'), 1)\|}$$

defined for $\boldsymbol{z} \in U_{\boldsymbol{x}}$ and $\boldsymbol{y} \in B(\mathbf{0}, s_{\boldsymbol{x}})$. Then $\boldsymbol{k}(\boldsymbol{p}(\mathbf{z}), \boldsymbol{z}) = \mathbf{0}$, while

$$\frac{\partial \boldsymbol{k}}{\partial \boldsymbol{y}}(\boldsymbol{y}, \boldsymbol{z}) = I_N - \mathrm{d}_V(\mathbf{z}) \frac{\partial}{\partial \boldsymbol{y}} \left(\frac{(\nabla' g(\boldsymbol{y}'), 1)}{\|(\nabla' g(\boldsymbol{y}'), 1)\|} \right).$$

It follows from (63) that

$$\frac{\partial}{\partial \boldsymbol{y}} \left(\frac{(\nabla' g(\boldsymbol{y}'), 1)}{\|(\nabla' g(\boldsymbol{y}'), 1)\|} \right) \leq M_{\boldsymbol{x}}$$

for all $\boldsymbol{y}' \in B_{N-1}(\mathbf{0}, s_{\boldsymbol{x}})$, where the constant $M_{\boldsymbol{x}}$ depends only on N, $s_{\boldsymbol{x}}$, and $L_{\boldsymbol{x}} > 0$. Hence, by replacing $\frac{1}{2L_{\boldsymbol{x}}}$ with a smaller constant in the set $U_{\boldsymbol{x}}$ defined in (67), we have that $\frac{\partial \boldsymbol{k}}{\partial \boldsymbol{y}}$ is invertible. Since $\boldsymbol{k}$ is of class C^1 it follows by the implicit function theorem that the function $\boldsymbol{p}$ is also of class C^1. In turn, by (62), $\nabla \mathrm{d}_V$ is of class C^1, and so d_V is of class C^2. A bootstrap argument and again the implicit function theorem shows that d_V is of class C^k. □

Remark 5.6. *Note that the first step continues to hold for open sets V with compact boundary of class $C^{1,1}$. However, it fails for open sets with compact boundary of class C^1.*

Exercise 5.7. *Let $\varepsilon > 0$ and consider the open set $V \subset \mathbb{R}^2$ bounded by the curve*

$$M = \left\{ \left(t, |t|^{2-\varepsilon}\right) : t \in [-1, 1] \right\} \cup \gamma,$$

where γ is any curve joining the points $(-1, 1)$ and $(1, 1)$ and range contained in the halfspace $y \geq 1$. Prove that d_V is not differentiable at points $(0, y)$ for $y > 0$ small.

Lemma 5.8. *Assume that $\Omega \subset \mathbb{R}^N$ is an open set with Lipschitz boundary, that $E \subset \mathbb{R}^N$ is an open set with ∂E a nonempty compact hypersurface of class C^2 with $\mathcal{H}^{N-1}(\partial E \cap \partial\Omega) = 0$. Then*

$$\lim_{r\to 0} \mathcal{H}^{N-1}\left(\{\boldsymbol{x} \in \Omega : \mathrm{d}_E(\boldsymbol{x}) = r\}\right) = \mathcal{H}^{N-1}(\Omega \cap \partial E).$$

Sketch of the proof. **Step 1:** Assume first that $\Omega = \mathbb{R}^N$. We claim that

$$\lim_{r\to 0} \mathcal{H}^{N-1}\left(\left\{\boldsymbol{x} \in \mathbb{R}^N : \mathrm{d}_E(\boldsymbol{x}) = r\right\}\right) = \mathcal{H}^{N-1}(\partial E).$$

Fix a point $\boldsymbol{x} \in \partial V$. Reasoning as in the previous proof we have that for every $\boldsymbol{z} \in U_{\boldsymbol{x}}$ there exists only one $\boldsymbol{y}' \in \overline{B_{N-1}(\mathbf{0}, s_{\boldsymbol{x}})}$ satisfying (65) and (66). If $\mathrm{d}_E(\boldsymbol{z}) = r$, then by (65) and (66),

$$z_N = g(\boldsymbol{y}') + r\frac{1}{\sqrt{1 + \|\nabla' g(\boldsymbol{y}')\|_{N-1}^2}} \tag{68}$$

and

$$z_i = y_i + r\frac{\frac{\partial g}{\partial y_i}(\boldsymbol{y}')}{\sqrt{1 + \|\nabla' g(\boldsymbol{y}')\|_{N-1}^2}} \tag{69}$$

for all $i = 1, \ldots, N-1$. This shows that locally the set $\{\boldsymbol{x} \in \Omega : \mathrm{d}_E(\boldsymbol{x}) = r\}$ is given by an $N-1$ dimensional manifold parametrized by the chart $\boldsymbol{\varphi}^r : B_{N-1}(0, s_{\boldsymbol{x}}) \to \mathbb{R}^N$ given by (68) and (69). Hence, locally the surface measure of $\{\boldsymbol{x} \in \Omega : \mathrm{d}_E(\boldsymbol{x}) = r\}$ is given by the surface integral

$$\int_{B_{N-1}(0,R)} \sqrt{\sum_{\boldsymbol{\alpha}\in\Lambda} \left[\det \frac{\partial\left(\varphi^r_{\alpha_1}, \ldots, \varphi^r_{\alpha_{N-1}}\right)}{\partial(y_1, \ldots, y_{N-1})}(\mathbf{y})\right]^2}\, d\mathbf{y},$$

where

$$\Lambda := \left\{\boldsymbol{\alpha} \in \mathbb{N}^{N-1} : 1 \le \alpha_1 < \alpha_2 < \cdots < \alpha_{N-1} \le N\right\}.$$

It follows from (68) and (69) that when $r \to 0$, $\boldsymbol{\varphi}^r(\boldsymbol{y}')$ converges to $\boldsymbol{\varphi}(\boldsymbol{y}') := (\boldsymbol{y}', g(\boldsymbol{y}'))$ (pointwise and uniformly), and in turn

$$\int_{B_{N-1}(0,R)} \sqrt{\sum_{\boldsymbol{\alpha}\in\Lambda} \left[\det \frac{\partial\left(\varphi^r_{\alpha_1}, \ldots, \varphi^r_{\alpha_{N-1}}\right)}{\partial(y_1, \ldots, y_{N-1})}(\mathbf{y})\right]^2}\, d\mathbf{y}$$
$$\to \int_{B_{N-1}(0,R)} \sqrt{\sum_{\boldsymbol{\alpha}\in\Lambda} \left[\det \frac{\partial\left(\varphi_{\alpha_1}, \ldots, \varphi_{\alpha_{N-1}}\right)}{\partial(y_1, \ldots, y_{N-1})}(\mathbf{y})\right]^2}\, d\mathbf{y},$$

which is locally the surface measure of ∂E. The general case follows using partitions of unity. We omit the details.

Step 2: For $r > 0$, let

$$V_r := \{\boldsymbol{x} \in E : 0 < \mathrm{d}_E(\boldsymbol{x}) < r\} = \{\boldsymbol{x} \in E : 0 < \operatorname{dist}(\boldsymbol{x}, \partial E) < r\}.$$

Note that

$$\partial (E \setminus V_r) = \{\boldsymbol{x} \in E : \, \mathrm{d}_E(\boldsymbol{x}) = r\}$$

and for every $\boldsymbol{x} \in \Omega$,

$$\chi_E(\boldsymbol{x}) - \chi_{E\setminus V_r}(\boldsymbol{x}) = \chi_{V_r}(\boldsymbol{x}) \to 0.$$

Hence, by the Lebesgue dominated convergence theorem, $\chi_{E\setminus V_r} \to \chi_E$ in $L^1(\Omega)$. It follows that

$$\begin{aligned}\mathcal{H}^{N-1}(\Omega \cap \partial E) = \mathrm{P}(E, \Omega) &\le \liminf_{r\to 0^+} \mathrm{P}(E \setminus V_r, \Omega) \\ &= \liminf_{r\to 0^+} \mathcal{H}^{N-1}(\Omega \cap \partial (E \setminus V_r)) \\ &= \liminf_{r\to 0^+} \mathcal{H}^{N-1}(\{\boldsymbol{x} \in \Omega : \, \mathrm{d}_E(\boldsymbol{x}) = r\}).\end{aligned}$$

To prove the opposite inequality, observe that

$$\begin{aligned}&\mathcal{H}^{N-1}(\{\boldsymbol{x} \in \Omega : \, \mathrm{d}_E(\boldsymbol{x}) = r\}) \\ &\le \mathcal{H}^{N-1}\left(\left\{\boldsymbol{x} \in \mathbb{R}^N : \, \mathrm{d}_E(\boldsymbol{x}) = r\right\}\right) - \mathcal{H}^{N-1}\left(\left\{\boldsymbol{x} \in \mathbb{R}^N \setminus \overline{\Omega} : \, \mathrm{d}_E(\boldsymbol{x}) = r\right\}\right).\end{aligned} \tag{70}$$

Reasoning as before, with Ω replaced by $\mathbb{R}^N \setminus \overline{\Omega}$ (note that that the set V_r is bounded ∂E is compact, so the Lebesgue dominated convergence theorem continues to hold), we have that

$$\mathcal{H}^{N-1}\left(\left(\mathbb{R}^N \setminus \overline{\Omega}\right) \cap \partial E\right) \le \liminf_{r\to 0^+} \mathcal{H}^{N-1}\left(\left\{\boldsymbol{x} \in \mathbb{R}^N \setminus \overline{\Omega} : \, \mathrm{d}_E(\boldsymbol{x}) = r\right\}\right). \tag{71}$$

Hence, from (70) and (71) and the fact that $\mathcal{H}^{N-1}(\partial E \cap \partial\Omega) = 0$, and Step 1,

$$\begin{aligned}\limsup_{r\to 0^+} \mathcal{H}^{N-1}(\{\boldsymbol{x} \in \Omega : \, \mathrm{d}_E(\boldsymbol{x}) = r\}) &\le \limsup_{r\to 0^+} \mathcal{H}^{N-1}\left(\left\{\boldsymbol{x} \in \mathbb{R}^N : \, \mathrm{d}_E(\boldsymbol{x}) = r\right\}\right) \\ &\quad - \liminf_{r\to 0^+} \mathcal{H}^{N-1}\left(\left\{\boldsymbol{x} \in \mathbb{R}^N \setminus \overline{\Omega} : \, \mathrm{d}_E(\boldsymbol{x}) = r\right\}\right) \\ &\le \mathcal{H}^{N-1}(\partial E) - \mathcal{H}^{N-1}\left(\left(\mathbb{R}^N \setminus \overline{\Omega}\right) \cap \partial E\right) \\ &= \mathcal{H}^{N-1}(\Omega \cap \partial E).\end{aligned}$$

This concludes the proof. □

6. Appendix

In this section we study different modes of convergence and their relation to one another.

Definition 6.1. *Let $(X, \mathfrak{M}, \mu)$ be a measure space and let $u_n, u : X \to \mathbb{R}$ be measurable functions.*

(i) *$\{u_n\}$ is said to* converge to u pointwise μ-a.e. *if there exists a set $E \in \mathfrak{M}$ such that $\mu(E) = 0$ and*

$$\lim_{n\to\infty} u_n(x) = u(x)$$

for all $x \in X \setminus E$.

(ii) *$\{u_n\}$ is said to* converge to u in measure *if for every $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mu(\{x \in X : |u_n(x) - u(x)| > \varepsilon\}) = 0.$$

The next theorem relates the types of convergence introduced in Definition 6.1 to convergence in $L^p(X)$.

Theorem 6.2. *Let $(X, \mathfrak{M}, \mu)$ be a measure space and let u_n, $u : X \to \mathbb{R}$ be measurable functions.*

(i) *If $\{u_n\}$ converges to u in measure, then there exists a subsequence $\{u_{n_k}\}$ such that $\{u_{n_k}\}$ converges to u almost pointwise μ-a.e.*

(ii) *If $\{u_n\}$ converges to u in $L^p(X)$, $1 \le p < \infty$, then it converges to u in measure and there exist a subsequence $\{u_{n_k}\}$ and an integrable function v such that $\{u_{n_k}\}$ converges to u pointwise μ-a.e. and $|u_{n_k}(x)|^p \le v(x)$ for μ-a.e. $x \in X$ and for all $k \in \mathbb{N}$.*

Theorem 6.3 (Egoroff). *Let $(X, \mathfrak{M}, \mu)$ be a measure space with μ finite and let $u_n : X \to \mathbb{R}$ be measurable functions converging pointwise μ-a.e. Then $\{u_n\}$ converges in measure.*

Theorem 6.4 (Vitali's convergence theorem). *Let $(X, \mathfrak{M}, \mu)$ be a measure space and let $u_n, u \in L^1(X)$. Then $\{u_n\}$ converges to u in $L^1(X)$ if and only if the following conditions hold:*

(i) $\{u_n\}$ converges to u in measure, *that is, for every $\varepsilon > 0$,*
$$\lim_{n\to\infty} \mu(\{x \in X : |u_n(x) - u(x)| > \varepsilon\}) = 0.$$

(ii) $\{u_n\}$ *is* equi-integrable, *that is, for every $\varepsilon > 0$ there exists $\delta > 0$ such that*
$$\int_E |u_n| \, d\mu \le \varepsilon$$
for all n and for every measurable set $E \subset X$ with $\mu(E) \le \delta$..

(iii) *For every $\varepsilon > 0$ there exists $E \subset X$ with $E \in \mathfrak{M}$ such that $\mu(E) < \infty$ and*
$$\int_{X \setminus E} |u_n| \, d\mu \le \varepsilon$$
for all n.

Remark 6.5. *Note that condition (iii) is automatically satisfied when X has finite measure.*

Theorem 6.6 (Rellich–Kondrachov). *Let $\Omega \subset \mathbb{R}^N$ be a bounded Lipschitz domain and let $\{u_n\} \subset W^{1,1}(\Omega)$ be a bounded sequence. Then there exist a subsequence $\{u_{n_k}\}$ of $\{u_n\}$ and a function $u \in BV(\Omega)$ such that $u_{n_k} \to u$ in $L^1(\Omega)$.*

Remark 6.7. *Using a diagonal argument, we can conclude that if $\Omega \subset \mathbb{R}^N$ is an open set and $\{u_n\} \subset W^{1,1}(\Omega)$ a bounded sequence, then there exist a subsequence $\{u_{n_k}\}$ of $\{u_n\}$ and a function $u \in BV(\Omega)$ such that $u_{n_k} \to u$ in $L^1_{\mathrm{loc}}(\Omega)$ and pointwise $\mathcal{L}^N$ a.e. in Ω. To see this, write*
$$\Omega = \bigcup_{i=1}^{\infty} \Omega_i,$$
where Ω_i is an increasing sequence of bounded Lipschitz domain. Apply the Rellich–Kondrachov theorem in Ω_1, to find a subsequence $\{u_{n,1}\}$ of $\{u_n\}$ and a function $w_1 \in BV(\Omega_1)$ such that $u_{n,1} \to w_1$ in $L^1(\Omega_1)$ and (by Theorem 6.2) pointwise $\mathcal{L}^N$ a.e. in Ω_1. Inductively, for each i, it follows by the Rellich–Kondrachov theorem in Ω_i, that there are a subsequence $\{u_{n,i}\}$ of $\{u_{n,i-1}\}$

and a function $w_i \in BV(\Omega_i)$ such that $u_{n,i} \to w_i$ in $L^1(\Omega_i)$ and (by Theorem 6.2) pointwise $\mathcal{L}^N$ a.e. in Ω_i. Note that $w_{i+1} = w_i$ in Ω_i (by the uniqueness of L^1 limits in $L^1(\Omega_i)$). Hence, if $\boldsymbol{x} \in \Omega$, letting i be such that $\boldsymbol{x} \in \Omega_i$, we may define $u(\boldsymbol{x}) := w_i(\boldsymbol{x})$. It follows by construction that the diagonal subsequence $\{u_{i,i}\}$ of $\{u_n\}$ converges to u pointwise $\mathcal{L}^N$ a.e. in Ω and in $L^1(K)$ for every compact set $K \subset \Omega$. We leave as an exercise to verify that u belongs to $BV(\Omega)$.

Theorem 6.8 (Chain Rule)**.** *Let $\Omega \subseteq \mathbb{R}^N$ be an open set, let $1 \leq p < \infty$, and let $g : \mathbb{R} \to \mathbb{R}$ be a Borel function. The $g \circ u$ belongs to $W^{1,p}(\Omega)$ for every $u \in W^{1,p}(\Omega)$ if and only if g is Lipschitz continuous, and, if $\mathcal{L}^N(\Omega) = \infty$, $g(0) = 0$. Moreover, in this case*

$$\frac{\partial}{\partial x_i}(g \circ u)(x) = \begin{cases} g'(u(x))\frac{\partial u}{\partial x_i}(x) & \text{for } \mathcal{L}^N \text{ a.e } x \in u^{-1}(\Omega \setminus \Sigma), \\ 0 & \text{for } \mathcal{L}^N \text{ a.e } x \in u^{-1}(\Sigma), \end{cases}$$

where $\Sigma := \{t \in \mathbb{R} : g \text{ is not differentiable at } t\}$.

Note that by a classical result due to Rademacher, the set Σ has Lebesgue measure zero. See [Le2009] for a proof of Theorems 6.6 and 6.8.

Proposition 6.9. *Let $\{a_{k,n}\}$ and $\{b_{k,n}\}$ be double-indexed sequences of real numbers and let $L, M \in \mathbb{R}$ be such that*

$$\lim_{k\to\infty}\lim_{n\to\infty} a_{k,n} = L, \quad \limsup_{k\to\infty}\limsup_{n\to\infty} b_{k,n} = M.$$

Then there exists a sequence $k_n \to \infty$ such that

$$\lim_{n\to\infty} a_{k_n,n} = L, \quad \limsup_{n\to\infty} b_{k_n,n} \leq M.$$

Proof. Define

$$\begin{aligned}\bar{k}_1 := \min\{k \in \mathbb{N} :\ &\text{there is } n \in \mathbb{N} \text{ such that for all } m \geq n, \\ &|a_{k,m} - L| < 1,\ b_{k,n} < M + 1\}.\end{aligned}$$

We claim that the minimum exists. If not, then for every $k \in \mathbb{N}$ there exists a sequence $\{n^{(k)}\}$ such that either $\left|a_{k,n^{(k)}} - L\right| \geq 1$ or $b_{k,n^{(k)}} \geq M + 1$. Therefore it is possible to extract a further subsequence (not relabelled) such that either

$$\liminf_{n^{(k)}\to\infty}\left|a_{k,n^{(k)}} - L\right| \geq 1$$

or

$$\liminf_{n^{(k)}\to\infty} b_{k,n^{(k)}} \geq M + 1.$$

Hence, there exists a subsequence $\{k_j\}$ such that for every $j \in \mathbb{N}$ either

$$\liminf_{n^{(k_j)}\to\infty}\left|a_{k_j,n^{(k_j)}} - L\right| \geq 1$$

or

$$\liminf_{n^{(k_j)}\to\infty} b_{k_j,n^{(k_j)}} \geq M + 1.$$

In the first case we obtain

$$0=\lim_{k\to\infty}\lim_{n\to\infty}|a_{k,n}-L|=\liminf_{j\to\infty}\liminf_{n^{\left(k_j\right)}\to\infty}\left|a_{k_j,n^{\left(k_j\right)}}-L\right|\geq 1,$$

which is a contradiction. In the second case we have

$$M=\limsup_{k\to\infty}\limsup_{n\to\infty}b_{k,n}\geq\limsup_{j\to\infty}\liminf_{n^{\left(k_j\right)}\to\infty}b_{k_j,n^{\left(k_j\right)}}\geq M+1,$$

which is again a contradiction. This proves the claim and we define

$$n_1:=\min\left\{n\in\mathbb{N}:\text{ for all } m\geq n, \left|a_{\bar{k}_1,m}-L\right|<1,\, b_{\bar{k}_1,m}<M+1\right\}.$$

Recursively, for $p\geq 2$ we define

$$\begin{aligned}\bar{k}_p:=\min\Big\{&k>\bar{k}_{p-1}:\text{ there is } n>n_{p-1} \text{ such that for all } m\geq n,\\ &|a_{k,m}-L|<\frac{1}{p},\, b_{k,n}<M+\frac{1}{p}\Big\}\end{aligned}$$

and

$$n_p:=\min\left\{n>n_{p-1}:\text{ for all } m\geq n, \left|a_{\bar{k}_p,m}-L\right|<1,\, b_{\bar{k}_p,m}<M+1\right\}.$$

For every $p\in\mathbb{N}$, $p\geq 2$, and for every $n\in\{n_{p-1},\dots,n_p\}$ define

$$k_n:=\bar{k}_{p-1}$$

and note that $|a_{k_n,n}-L|<\frac{1}{p}$ and $b_{k_n,n}<M+\frac{1}{p}$. □

References

[ABI1998] G. Alberti and G. Bellettini, A nonlocal anisotropic model for phase transitions. I. The optimal profile problem. Math. Ann. 310 (1998) 527–560.

[ABII1998] G. Alberti and G. Bellettini, A non-local anisotropic model for phase transitions: asymptotic behaviour of rescaled energies. European J. Appl. Math. 9 (1998) 261–284.

[ABCP1996] G. Alberti, G. Bellettini, M. Cassandro, and E. Presutti, Surface tension in Ising systems with Kac potentials, J. Stat. Phys. 82 (1996) 743–796.

[ABS1994] G. Alberti, G. Bouchitté, and P. Seppecher, Un résultat de perturbations singulières avec la norme $H^{1/2}$. C. R. Acad. Sci. Paris Sér. I Math. 319 (1994) 333–338.

[ABS1998] G. Alberti, G. Bouchitté, and P. Seppecher, Phase transition with the line-tension effect. Arch. Rat. Mech. Anal. 144 (1998) 1–46.

[Am1990] L. Ambrosio, Metric space valued functions of bounded variation. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 17 (1990) 439–478.

[AFP2000] L. Ambrosio, N. Fusco, and D. Pallara, Functions of bounded variation and free discontinuity problems, Oxford Mathematical Monographs. Oxford: Clarendon Press, 2000.

[AB1993] G. Anzellotti and S. Baldo, Asymptotic development by Γ-convergence. Appl. Math. Optim. 27 (1993) 105–123.

[ABO1996] G. Anzellotti, S. Baldo, and G. Orlandi, Γ-asymptotic developments, the Cahn-Hilliard functional, and curvatures. J. Math. Anal. Appl. 197 (1996) 908–924.

[Ba1990] S. Baldo, Minimal interface criterion for phase transitions in mixtures of Cahn-Hilliard fluids. Ann. Inst. H. Poincaré Anal. Non Linéaire 7 (1990) 67–90.

[BJ1987] J. M. Ball and R.D. James, Fine phase mixtures as minimizers of the energy. Arch. Rat. Mech. Anal. 100 (1987) 13–52.

[BF1994] A.C. Barroso and I. Fonseca, Anisotropic singular perturbations - the vectorial case. Proc. Roy. Soc. Edinburgh Sect. A 124 (1994) 527–571.

[Br2002] A. Braides, Γ-convergence for beginners. Oxford Lecture Series in Mathematics and its Applications, 22. Oxford: Oxford University Press, 2002.

[Bo1990] G. Bouchitté, Singular perturbations of variational problems arising from a two-phase transition model. Appl. Math. Optim. 21 (1990), 289–314.

[CH1958] J. W. Cahn and J.E. Hilliard, Free energy of a nonuniform system. I. Interfacial free energy, J. Chem. Phys 28 (1958) 258–267.

[CGS1984] J. Carr, M.E. Gurtin, and M. Slemrod, Structured phase transitions on a finite interval. Arch. Rat. Mech. Anal. 86 (1984) 317–351.

[CC2010] M. Chermisi and S. Conti, Multiwell rigidity in nonlinear elasticity. SIAM J. Math. Anal. 42 (2010) 1986–2012.

[CDMFL2011] M. Chermisi, G. Dal Maso, I. Fonseca, and G. Leoni, Singular perturbation models in phase transitions for second-order materials. Indiana Univ. Math. J. 60 (2011) 367–409.

[CSZ2011] M. Cicalese, E. Spadaro, C.I. Zeppieri, Asymptotic analysis of a second-order singular perturbation model for phase transitions. Calc. Var. Partial Differential Equations 41 (2011) 127–150.

[CMM1992] B. D. Coleman, M. Marcus, and V. J. Mizel, On the thermodynamics of periodic phases, Arch. Ration. Mech. Anal. 117 (1992) 321–347.

[CFL2002] S. Conti, I. Fonseca, and G. Leoni, A Γ-convergence result for the two-gradient theory of phase transitions. Comm. Pure Appl. Math. 55 (2002) 857–936.

[CS2006] S. Conti and B. Schweizer, Rigidity and gamma convergence for solid-solid phase transitions with $SO(2)$ invariance. Comm. Pure Appl. Math. 59 (2006) 830–868.

[CER1987] P. Coullet, C. Elphick, and D. Repaux, Nature of spatial chaos, Phys. Rev. Lett., 58 (1987) 431–434.

[DM1993] G. Dal Maso, An introduction to Γ-convergence, Progress in Nonlinear Differential Equations and their Applications. 8. Basel: Birkhäuser, 1993.

[DMFL2013] G. Dal Maso, I. Fonseca, and G. Leoni, Second order asymptotic development for the Cahn-Hilliard functional, in preparation.

[DvS1988] G. T. Dee and W. van Saarloos, Bistable systems with propagating fronts leading to pattern formation, Phys. Rev. Lett., 60 (1988) 2641–2644.

[DG1975] E. De Giorgi, Sulla convergenza di alcune successioni d'integrali del tipo dell'area. Rend. Mat. (IV) 8 (1975) 277–294.

[DZ2001] M.C. Delfour and J.P. Zolésio, *Shapes and geometries. Analysis, differential calculus, and optimization,* Advances in Design and Control, 4. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001.

[DNPV2012] E. Di Nezza, G. Palatucci, and E. Valdinoci, Hitchhiker's guide to the fractional Sobolev spaces. Bull. Sci. Math. 136 (2012) 521–573.

[Fe1959] H. Federer, Curvature measures, Transactions of the American Mathematical Society, 93 (1959) 418-491.

[Fo1991] I. Fonseca, The Wulff theorem revisited. Proc. Roy. Soc. London Ser. A 432 (1991) 125–145.

[FMa2000] I. Fonseca and C. Mantegazza, Second order singular perturbation models for phase transitions. SIAM J. Math. Anal. 31 (2000), no. 5, 1121–1143.

[FMu1991] I. Fonseca, S. Müller, A uniqueness proof for the Wulff theorem. Proc. Roy. Soc. Edinburgh Sect. A 119 (1991) 125–136.

[FT1989] I. Fonseca and L. Tartar, The gradient theory of phase transitions for systems with two potential wells. Proc. Roy. Soc. Edinburgh Sect. A 111 (1989) 89–102.

[Fo1984] R. Foote, Regularity of the distance function. Proc. Amer. Math. Soc. 92 (1984)153–155.

[Ga1957] E. Gagliardo, Caratterizzazioni delle tracce sulla frontiera relative ad alcune classi di funzioni in n variabili, Rend. Sem. Mat. Univ. Padova 27 (1957) 284–305.

[Ga1959] E. Gagliardo, Ulteriori proprietà di alcune funzioni di piú variabili, Ricerche Mat. 8 (1959) 24–51.

[Gar2002] R.J. Gardner, The Brunn-Minkowski inequality, Bull. Amer. Math. Soc. 39 (2002) 355–405.

[GP2006] A. Garroni and G. Palatucci, A singular perturbation result with a fractional norm. Variational problems in materials science, 111–126, Progr. Nonlinear Differential Equations Appl., 68, Birkhäuser, Basel, 2006.

[Gu1985] M.E. Gurtin, Some results and conjectures in the gradient theory of phase transitions. IMA, preprint 156 (1985).

[He1951] C. Herring, Some theorems on the free energies of crystal surfaces. Phys. Rev. 82 (1951) 87–93.

[HPS2002] D. Hilhorst, L. A. Peletier, and R. Schätzle, Γ-limit for the extended Fisher-Kolmogorov equation, Proc. Roy. Soc. Edinburgh Sect. A 132 (2002) 141–162.

[KAKT1993] T. Kawakatsu, D. Andelman, K. Kawasaki, and T. Taniguchi, Phase-transitions and shapes of two-component membranes and vesicles I: strong segregation limit, Journal de Physique II 3 (1993) 971–997.

[KM1994] R.V. Kohn and S. Müller, Surface energy and microstructure in coherent phase transitions. Comm. Pure Appl. Math. 47 (1994), 405–435.

[KP1981] S. Krantz and H. Parks, Distance to Ck hypersurfaces, J. Differential Equations 40 (1981) 116-120.

[LA1987] S. Leibler and D. Andelman, Ordered and curved meso-structures in membranes and amphiphilic films, J. Phys. (France) 48 (1987) 2013–2018.

[LM1989] A. Leizarowitz and V. J. Mizel, One-dimensional infinite-horizon variational problems arising in continuum mechanics, Arch. Ration. Mech. Anal. 106 (1989) 161–193.

[Le2009] G. Leoni, A first course in Sobolev spaces, Graduate Studies in Mathematics 105. Providence, RI: American Mathematical Society (AMS), 2009.

[MPT1998] V. J. Mizel, L. A. Peletier, and W. C. Troy, Periodic phases in second-order materials, Arch. Ration. Mech. Anal. 145 (1998) 343–382.

[MM1977] L. Modica and S. Mortola, Un esempio di Γ-convergenza. (Italian) Boll. Un. Mat. Ital. B (5) 14 (1977) 285–299.

[Mo1987] L. Modica, The gradient theory of phase transitions and the minimal interface criterion. Arch. Rational Mech. Anal. 98 (1987) 123–142.

[Ni1966] L. Nirenberg, An extended interpolation inequality, Ann. Sc. Norm. Super. Pisa, Sci. Fis. Mat., III. Ser. 20 (1966), 733–737.

[Ow1988] N.C. Owen, Nonconvex variational problems with general singular perturbations. Trans. Amer. Math. Soc. 310 (1988) 393–404.

[OS1991] N.C. Owen and P. Sternberg, Nonconvex variational problems with anisotropic perturbations. Nonlinear Anal. 16 (1991) 705–719.

[PT1997] L. A. Peletier and W. C. Troy, Spatial patterns described by the extended Fisher-Kolmogorov equation: Periodic solutions, SIAM J. Math. Anal. 28 (1997) 1317–1353.

[Ro1979] J.S. Rowlinson, Translation of J.D. Van der Waals: The thermodynamic theory of capillarity under the hypothesis of a continuous variation of density, J. Stat. Phys. 20 (1979) 200–244.

[SV2012] O. Savin and E. Valdinoci, Γ-convergence for nonlocal phase transitions. Ann. I. H. Poincaré 29 (2012) 479–500.

[SA1995] M. Seul and D. Andelman, Domain shapes and patterns - the phenomenology of modulated phases, Science 267 (1995) 476–483.

[SvdB2002] D. Smets and J. B. van den Berg, Homoclinic solutions for Swift-Hohenberg and suspension bridge type equations, J. Differential Equations 184 (2002) 78–96.

[SH1977] J. B. Swift and P. C. Hohenberg, Hydrodynamic Fluctuations at the convective instability, Phys. Rev. A 15 (1977) 319–328.

[St1988] P. Sternberg, The effect of a singular perturbation on nonconvex variational problems. Arch. Rational Mech. Anal. 101 (1988) 209–260.

[St1991] P. Sternberg, Vector-valued local minimizers of nonconvex variational problems. Rocky Mountain J. Math. 21 (1991), 799–807.

[Ta1971] J.E. Taylor, Existence and structure of solutions to a class of nonelliptic variational problems. Symposia Mathematica, Vol. XIV (Convegno di Teoria Geometrica dell'Integrazione e Varietà Minimali, INDAM, Roma, Maggio 1973), 499–508. Academic Press, London, 1974.

[Ta1975] J.E. Taylor, Unique structure of solutions to a class of nonelliptic variational problems. (Proc. Sympos. Pure. Math., Vol. XXVII, Stanford Univ., Stanford, Calif., 1973), Part 1, 419–427. Amer. Math. Soc., Providence, R.I., 1975.

[Ta1978] J.E. Taylor, Crystalline variational problems. Bull. Amer. Math. Soc. 84 (1978), no. 4, 568–588.

[TKAK1994] T. Taniguchi, K. Kawasaki, D. Andelman, and T. Kawakatsu, Phase-transitions and shapes of two-component membranes and vesicles II: weak segregation limit, Journal de Physique II 4 (1994) 1333–1362.

[VdW1893] J.D. van der Waals, The thermodynamic theory of capillarity under the hypothesis of a continuous variation of density, Zeitschrift für Physikalische Chemie 13 (1894) 657–725.

[Z2013] B. Zwicknagl, Microstructures in low-hysteresis shape memory alloys: scaling regimes and optimal needle shapes. Preprint 2013.