Chapter 11

Relaxation in Sobolev, BV, and Young measures spaces

11.1 ▪ Relaxation in abstract metrizable spaces

This section is devoted to the description of the relaxation principle in a general metrizable space, or more generally, in a first countable topological space X. Roughly speaking, given an extended real-valued function $F : X \longrightarrow \mathbf{R} \cup \{+\infty\}$, we wish to apply the direct method in the calculus of variations to the lower semicontinuous envelope $\mathrm{cl}(F)$ of the function F so that $\inf_X F = \min_X \mathrm{cl}(F)$. Such a procedure is very important in various applications and leads to the concept of generalized solutions for the optimization problem $\inf_X F$. We begin by giving some complements on the sequential version, denoted by $\overline{F}$, of the general notion of lower semicontinuous envelope $\mathrm{cl}(F)$ introduced in Chapter 3.

Proposition 11.1.1. *Let $F : X \longrightarrow \mathbf{R} \cup \{+\infty\}$ be a proper extended real-valued function defined on a metrizable space (X, d) or, more generally, on a first countable topological space, and let us define the extended real-valued function $\overline{F} : X \longrightarrow \overline{\mathbf{R}}$ by*

$$\overline{F}(x) := \inf\Big\{ \liminf_{n\to+\infty} F(x_n) \ :\ (x_n)_{n\in\mathbf{N}},\ x = \lim_{n\to+\infty} x_n \Big\}. \tag{11.1}$$

Then the function $\overline{F}$ is characterized for every x in X by the two following assertions:

(i) $\forall (x_n)_{n\in\mathbf{N}}$ *such that* $x_n \to x$, $\overline{F}(x) \le \liminf_{n\to+\infty} F(x_n)$;

(ii) *there exists a sequence* $(y_n)_{n\in\mathbf{N}}$ *in* X *such that* $y_n \to x$ *and* $\overline{F}(x) \ge \limsup_{n\to+\infty} F(y_n)$.

PROOF. Note that trivially the system of assertions (i) and (ii) is equivalent to (i) and (ii′):

(i) $\forall (x_n)_{n\in\mathbf{N}}$ such that $x_n \to x$, $\overline{F}(x) \le \liminf_{n\to+\infty} F(x_n)$;

(ii′) there exists a sequence $(y_n)_{n\in\mathbf{N}}$ in X such that $y_n \to x$ and $\overline{F}(x) = \lim_{n\to+\infty} F(y_n)$;

and each function $\overline{F}$ satisfying (i) and (ii′) automatically satisfies

$$\overline{F}(x) := \inf\Big\{ \liminf_{n\to+\infty} F(x_n) : (x_n)_{n\in\mathbf{N}},\ x = \lim_{n\to+\infty} x_n \Big\}.$$

We are reduced to establishing that the function $\overline{F}$ defined by formula (11.1) satisfies (i) and (ii′). We only establish the nontrivial assertion (ii′). Its proof is based on the following diagonalization lemma.

Lemma 11.1.1. *Let $(a_{m,n})_{(m,n)\in\mathbf{N}\times\mathbf{N}}$ be a sequence in a first countable topological space X such that*

(i) $\lim_{n\to+\infty} a_{m,n} = a_m$;

(ii) $\lim_{m\to+\infty} a_m = a$.

Then there exists a nondecreasing map $n \mapsto m(n)$ from $\mathbf{N}$ into $\mathbf{N}$ such that

$$\lim_{n\to+\infty} a_{m(n),n} = a.$$

For a proof and other diagonalization results, see [37]. Note that under the same conditions, we have the following more classical diagonalization procedure: there exists an increasing map $m \mapsto n(m)$ such that $\lim_{m\to+\infty} a_{m,m(n)} = a$. This second result could be applied for proving Proposition 11.1.1 but we prefer using Lemma 11.1.1, which will turn out to be the fitting tool for establishing a similar proposition in the context of the Γ-convergence (see the next chapter).

Let us go back to the proof of Proposition 11.1.1. By definition of the infima, for all $m \in \mathbf{N}^*$ there exists a sequence $(x_{m,n})_{n\in\mathbf{N}}$ in X satisfying $\lim_{n\to+\infty} x_{m,n} = x$ and such that

$$\begin{aligned}
&\text{if } \overline{F}(x) \neq -\infty, \quad \overline{F}(x) \le \liminf_{n\to+\infty} F(x_{m,n}) \le \overline{F}(x) + \frac{1}{m},\\
&\text{if } \overline{F}(x) = -\infty, \quad \liminf_{n\to+\infty} F(x_{m,n}) \le -m.
\end{aligned}$$

Therefore $\lim_{m\to+\infty}\lim_{n\to+\infty} F(x_{m,\sigma_m(n)}) = \overline{F}(x)$, where $\sigma_m : \mathbf{N} \longrightarrow \mathbf{N}$ is an increasing map, possibly depending on m, such that

$$\liminf_{n\to+\infty} F(x_{m,n}) = \lim_{n\to+\infty} F(x_{m,\sigma_m(n)}).$$

We end the proof by applying Lemma 11.1.1 to the sequence $(x_{m,\sigma_m(n)}, F(x_{m,\sigma_m(n)}))_{(m,n)\in\mathbf{N}^2}$ in the metrizable space $X \times \overline{\mathbf{R}}$: there exists $n \mapsto m(n)$ mapping $\mathbf{N}$ into $\mathbf{N}$ such that

$$\begin{cases}
\lim_{n\to+\infty} F(x_{m(n),\sigma_{m(n)}(n)}) = \lim_{m\to+\infty}\lim_{n\to+\infty} F(x_{m,\sigma_m(n)}) = \overline{F}(x),\\
\lim_{n\to+\infty} x_{m(n),\sigma_{m(n)}(n)} = \lim_{m\to+\infty}\lim_{n\to+\infty} x_{m,\sigma_m(n)} = x.
\end{cases}$$

The sequence $(y_n)_{n\in\mathbf{N}}$ defined by $y_n = x_{m(n),\sigma_{m(n)}(n)}$ fulfills assertion (ii′). □

Theorem 11.1.1. *The function $\overline{F}$ defined in Proposition 11.1.1 is the lower semicontinuous (lsc) envelope $cl(F)$ of the function F, i.e., the greatest lsc function less than F.*

PROOF. We must establish

$$\begin{cases}
\overline{F} \le F;\\
\overline{F} \text{ lsc};\\
G : X \longrightarrow \overline{\mathbf{R}},\ G \text{ lsc, and } G \le F \Longrightarrow G \le \overline{F}.
\end{cases}$$

For the first assertion, take the constant sequence $(x_n)_{n\in\mathbf{N}} = (x)_{n\in\mathbf{N}}$ in formula (11.1).

Let us prove the second assertion. Let $(y_m)_{m\in\mathbf{N}}$ be a sequence in X converging to $y \in X$ and consider a subsequence $(y_{\sigma(m)})_{m\in\mathbf{N}}$ satisfying $\lim_{m\to+\infty} \overline{F}(y_{\sigma(m)}) = \liminf_{m\to+\infty} \overline{F}(y_m)$. According to Proposition 11.1.1, there exists a sequence $(y_{\sigma(m),n})_{n\in\mathbf{N}}$ in X satisfying

$$\lim_{n\to+\infty} y_{\sigma(m),n} = y_{\sigma(m)}$$

and such that

$$\begin{aligned}\liminf_{m\to+\infty} \overline{F}(y_m) &= \lim_{m\to+\infty} \overline{F}(y_{\sigma(m)})\\ &= \lim_{m\to+\infty} \lim_{n\to+\infty} F(y_{\sigma(m),n}).\end{aligned}$$

On the other hand, we have $\lim_{m\to+\infty} \lim_{n\to+\infty} y_{\sigma(m),n} = y$. Applying the diagonalization lemma, Lemma 11.1.1, to the sequence $(y_{\sigma(m),n}, F(y_{\sigma(m),n}))_{(m,n)\in\mathbf{N}^2}$ in the metrizable space $X \times \overline{\mathbf{R}}$, there exists $n \mapsto m(n)$ mapping $\mathbf{N}$ to $\mathbf{N}$ such that

$$\begin{cases} \lim_{n\to+\infty} F(y_{\sigma(m(n)),n}) = \liminf_{m\to+\infty} \overline{F}(y_m), \\ \lim_{n\to+\infty} y_{\sigma(m(n)),n} = y. \end{cases}$$

Hence

$$\begin{aligned}\liminf_{m\to+\infty} \overline{F}(y_m) &= \lim_{n\to+\infty} F(y_{\sigma(m(n)),n})\\ &\geq \liminf_{n\to+\infty} F(y_{\sigma(m(n)),n})\\ &\geq \inf\Big\{\liminf_{n\to+\infty} F(x_n) \ : \ (x_n)_{n\in\mathbf{N}},\ y = \lim_{n\to+\infty} x_n\Big\} = \overline{F}(y).\end{aligned}$$

We establish now the third assertion. Let $G \leq F$ be a lsc function mapping X into $\overline{\mathbf{R}}$ and, for every x in X, consider any sequence $(x_n)_{n\in\mathbf{N}}$ in X converging to x. We have

$$\begin{aligned}G(x) &\leq \liminf_{n\to+\infty} G(x_n)\\ &\leq \liminf_{n\to+\infty} F(x_n).\end{aligned}$$

Taking the infimum over all the sequences $(x_n)_{n\in\mathbf{N}}$ converging to x, we finally obtain the inequality $G(x) \leq \overline{F}(x)$. □

From now on, we write indifferently $\overline{F}$ or $\mathrm{cl}(F)$ when X is metrizable or first countable. But the two notions differ when X is a general topological space. In the following theorem, we state the abstract relaxation principle in countable topological spaces.

Theorem 11.1.2. *Let $F : X \longrightarrow \mathbf{R} \cup \{+\infty\}$ be a proper extended real-valued function defined on a metrizable space (X,d) or, more generally, on a first countable topological space, and assume that there exists a minimizing sequence $(x_n)_{n\in\mathbf{N}}$ (i.e., $\lim_{n\to+\infty} F(x_n) = \inf_X F$) such that $S = \{x_n \ : \ n \in \mathbf{N}\}$ is relatively compact in X. Then*

(i) $\inf_X F = \min_X cl(F)$;

(ii) *every cluster point $\overline{x}$ of S is a solution of* $\min_X cl(F)$, *i.e.*, $cl(F)(\overline{x}) = \min_X cl(F)$.

PROOF. Let $\overline{x}$ be any cluster point of S and $(x_{\sigma(n)})_{n\in\mathbf{N}}$ a subsequence of $(x_n)_{n\in\mathbf{N}}$ converging to $\overline{x}$. According to Proposition 11.1.1(i), we have

$$\mathrm{cl}(F)(\overline{x}) \le \liminf_{n\to+\infty} F(x_{\sigma(n)}) = \inf_X F. \tag{11.2}$$

Let now x be any element of X. From Proposition 11.1.1(ii), there exists a sequence $(y_n)_{n\in\mathbf{N}}$ converging to x and satisfying

$$\mathrm{cl}(F)(x) \ge \limsup_{n\to+\infty} F(y_n). \tag{11.3}$$

Combining (11.2) and (11.3), we obtain

$$\mathrm{cl}(F)(\overline{x}) \le \liminf_{n\to+\infty} F(x_{\sigma(n)}) = \inf_X F \le \limsup_{n\to+\infty} F(y_n) \le \mathrm{cl}(F)(x). \tag{11.4}$$

This proves that $\mathrm{cl}(F)(\overline{x}) = \min_X \mathrm{cl}(F)$. Taking now $x = \overline{x}$ in (11.4), we obtain

$$\inf_X F = \min_X \mathrm{cl}(F)$$

and the proof is complete. □

In the terminology of the relaxation theory, the problem

$$(\overline{\mathscr{P}}): \qquad \min \mathrm{cl}(F)$$

is called the *relaxed problem* of the optimization problem

$$(\mathscr{P}): \qquad \inf_X F.$$

A solution of $(\overline{\mathscr{P}})$ is sometimes called a *generalized solution* of the initial problem $(\mathscr{P})$. The relaxation procedure consists in making explicit the lsc envelope of the functional F for a suitable topology on the space X, in order to obtain a well-posed problem $(\overline{\mathscr{P}})$ in the sense of Theorem 11.1.2 (i.e., the existence of an optimal solution holds for $(\overline{\mathscr{P}})$).

11.2 ▪ Relaxation of integral functionals with domain $W^{1,p}(\Omega,\mathbf{R}^m)$, $p > 1$

One of the fundamentals hypotheses in elasticity theory is that the total free energy F associated with many materials is of local nature. From a mathematical point of view, the functional F can be represented as the integral over the reference configuration $\Omega \subset \mathbf{R}^N$ ($N = 3$), of a density associated with the possible deformation gradients of the body, which account for the local deformations. The other basic principle is that equilibrium configurations correspond to minimizers of F under prescribed conditions in a Sobolev space $W^{1,p}(\Omega,\mathbf{R}^m)$ ($m = 3$). The functional F may fail to be lower semicontinuous. Indeed, in order to model the various solid/solid phase transformations in the microstructure, the density energy possesses in general a multiwell structure, and corresponding optimization problems have no solutions.

According to Section 11.1, a classical procedure is to replace F by its lower semicontinuous envelope with respect to the weak topology of $W^{1,p}(\Omega,\mathbf{R}^m)$. The relaxed problem possesses now at least a solution giving the same initial energy.

The case $p=1$ will be treated in the next section. We will see that contrary to the case $p>1$, the domain of the lower semicontinuous envelope $\mathrm{cl}(F)$ of the functional F strictly contains the domain $W^{1,1}(\Omega,\mathbf{R}^m)$ of F. This domain is indeed the space $BV(\Omega,\mathbf{R}^m)$ of functions of bounded variation introduced in Chapter 10. This phenomena is due to the lack of reflexivity of the Sobolev space $W^{1,1}(\Omega,\mathbf{R}^m)$.

This approach, which consists in relaxing the functional F with respect to the weak topology of $W^{1,p}(\Omega,\mathbf{R}^m)$ (or to the strong topology of $L^p(\Omega,\mathbf{R}^m)$), is not the only one. Another idea consists in"enlarging" the space $W^{1,p}(\Omega,\mathbf{R}^m)$ of admissible functions and treating the problem in the space $\mathscr{Y}(\Omega;\mathbf{M}^{m\times N})$ of Young measures introduced in Section 4.3. Actually this is the same general procedure: the integral functionals are considered as living on the space $X=\mathscr{Y}(\Omega;\mathbf{M}^{m\times N})$ equipped with a metrizable topology for which compactness of minimizing sequences also holds. One computes the lsc envelope of F relative to this new space. The two relaxed problems are different but there are some important connections between them. This procedure will be described in detail in Section 11.4.

Let us now make precise the structure of the functional F. We consider a bounded open subset Ω of $\mathbf{R}^N$ sufficiently regular in order that the trace theory, the Rellich–Kondrakov theorem, Theorem 5.4.2, and density arguments apply (take, for example, Ω of class $\mathbf{C}^1$). We denote the space of $m\times N$ matrices with entries in $\mathbf{R}$ by $\mathbf{M}^{m\times N}$ and consider a function

$$f:\mathbf{M}^{m\times N}\longrightarrow\mathbf{R}$$

such that there exist three positive constants α, β, L satisfying

$$\forall a\in\mathbf{M}^{m\times N}\quad \alpha|a|^p\leq f(a)\leq\beta(1+|a|^p) \tag{11.5}$$

$$\forall a,\ b\in\mathbf{M}^{m\times N}\quad |f(a)-f(b)|\leq L|b-a|(1+|a|^{p-1}+|b|^{p-1}). \tag{11.6}$$

Let $W^{1,p}(\Omega,\mathbf{R}^m)$ be the space (isomorphic to $W^{1,p}(\Omega)^m$; see Chapter 5), made up of all functions $u:\Omega\longrightarrow\mathbf{R}^m$ whose distributional gradient $\nabla u=(\frac{\partial u_i}{\partial x_j})_{i=1\dots m,\ j=1\dots N}$ belongs to $L^p(\Omega,\mathbf{M}^{m\times N})$. We define the functional $F:L^p(\Omega,\mathbf{R}^m)\longrightarrow\mathbf{R}^+\cup\{+\infty\}$ by

$$F(u)=\begin{cases}\displaystyle\int_\Omega f(\nabla u)\,dx & \text{if } u\in W^{1,p}(\Omega,\mathbf{R}^m),\\ +\infty & \text{otherwise,}\end{cases} \tag{11.7}$$

and we intend to compute its lsc envelope in the space $X=L^p(\Omega,\mathbf{R}^m)$ equipped with its strong topology. Actually, by classical arguments (the compactness of the embedding of $W^{1,p}(\Omega,\mathbf{R}^m)$ into $L^p(\Omega,\mathbf{R}^m)$ and the lower bound in (11.5)), one can easily establish that the lsc envelope of F considered as living on $W^{1,p}(\Omega,\mathbf{R}^m)$ equipped with its weak topology coincides with the restriction to $W^{1,p}(\Omega,\mathbf{R}^m)$ of the lsc envelope of F considered here. Proposition 11.2.1 below states that the domain of F is not relaxed.

Proposition 11.2.1. *The domain of the functional cl(F) is the space $W^{1,p}(\Omega,\mathbf{R}^m)$.*

PROOF. From the inequality $\mathrm{cl}(F)\leq F$, we obviously obtain $W^{1,p}(\Omega,\mathbf{R}^m)\subset\mathrm{dom}(\mathrm{cl}(F))$. For the converse inclusion, let $u\in L^p(\Omega,\mathbf{R}^m)$ such that $\mathrm{cl}(F)(u)<+\infty$ and consider a sequence $(u_n)_{n\in\mathbf{N}}$ strongly converging to u in $L^p(\Omega,\mathbf{R}^m)$ and satisfying $\mathrm{cl}(F)(u)=\lim_{n\to+\infty}F(u_n)$. Such a sequence exists from Proposition 11.1.1. According to the lower bound (11.5) and to the equality $\mathrm{cl}(F)(u)=\lim_{n\to+\infty}F(u_n)<+\infty$ one obtains

$$\sup_{n\in\mathbf{N}}\int_\Omega|\nabla u_n|^p\,dx<+\infty,$$

so that $(u_n)_{n\in\mathbf{N}}$ is bounded in $W^{1,p}(\Omega,\mathbf{R}^m)$. According to the Rellich–Kondrakov theorem, Theorem 5.4.2, there exists a subsequence (not relabeled) and $v \in W^{1,p}(\Omega,\mathbf{R}^m)$ such that

$$u_n \rightharpoonup v \quad \text{weakly in } W^{1,p}(\Omega,\mathbf{R}^m),$$
$$u_n \to v \quad \text{strongly in } L^p(\Omega,\mathbf{R}^m).$$

Consequently $u = v \in W^{1,p}(\Omega,\mathbf{R}^m)$. $\square$

In Theorem 11.2.1, we establish that the functional $\mathrm{cl}(F)$ possesses an integral representation. In the following proposition, also valid for $p=1$, we characterize the density of this integral functional. For every bounded Borel set A of $\mathbf{R}^N$, we will sometimes denote its N-dimensional Lebesgue measure by $|A|$ rather than $\mathscr{L}^N(A)$.

Proposition 11.2.2 (quasi-convex envelope of f). *Let us consider a function $f : \mathbf{M}^{m\times N} \longrightarrow \mathbf{R}^+$ satisfying for $p \geq 1$ and all $a \in \mathbf{M}^{m\times N}$ the upper growth condition $0 \leq f(a) \leq \beta(1+|a|^p)$ and the continuity assumption* (11.6). *Then for each fixed a in $\mathbf{M}^{m\times N}$, the map*

$$D \mapsto I_D := \inf\left\{ \frac{1}{|D|}\int_D f(a+\nabla\phi(x))\,dx \;:\; \phi \in W_0^{1,p}(D,\mathbf{R}^m)\right\}$$

is constant on the family of all open bounded subsets of $\mathbf{R}^N$ whose boundary satisfies $|\partial D| = 0$; we denote it by $Qf(a)$. If f satisfies (11.5), *the function $Qf : \mathbf{M}^{m\times N} \to \mathbf{R}^+$, defined for all $a \in \mathbf{M}^{m\times N}$ by*

$$Qf(a) = \inf\left\{ \frac{1}{|D|}\int_D f(a+\nabla\phi(x))\,dx \;:\; \phi \in W_0^{1,p}(D,\mathbf{R}^m)\right\},$$

satisfies the same condition (11.5) *and* (11.6) *with a new constant L' depending only on α, β, and p. Moreover, Qf is $W^{1,p}$-quasi-convex in the sense of Morrey (quasi-convex for short), namely, it satisfies the so-called quasi-convexity inequality: for all open bounded subset D of $\mathbf{R}^N$ with $|\partial D| = 0$,*

$$\forall a \in \mathbf{M}^{m\times N},\ \forall \phi \in W_0^{1,p}(D,\mathbf{R}^m) \quad Qf(a) \leq \frac{1}{|D|}\int_D Qf(a+\nabla\phi(x))\,dx. \tag{11.8}$$

Furthermore, the function Qf is the greatest quasi-convex function less than or equal to f, also called the quasi-convexification or quasi-convex envelope of f.

PROOF. (a) Let D and D' be two open bounded subsets of $\mathbf{R}^N$ with $|\partial D| = |\partial D'| = 0$. For proving the first assertion, it suffices to establish $I_D \leq I_{D'}$ and to invert the roles of D and D'.

For every $\varepsilon > 0$ there exists a finite family $(x_i + \varepsilon_i D')_{i\in I_\varepsilon}$ of pairwise disjoint sets $x_i + \varepsilon_i D' \subset D$, $\varepsilon_i > 0$, satisfying

$$\left| D \setminus \bigcup_{i\in I_\varepsilon}(x_i + \varepsilon_i D')\right| < \varepsilon. \tag{11.9}$$

First, we claim that the map I verifies the following subadditivity property: if A and B are two disjoint bounded open subsets of $\mathbf{R}^N$, then

$$|A\cup B| I_{A\cup B} \leq |A| I_A + |B| I_B. \tag{11.10}$$

Indeed, let ϕ_A and ϕ_B be two η-minimizers of $|A|I_A$ and $|B|I_B$ in $\mathscr{D}(A,\mathbf{R}^m)$ and $\mathscr{D}(B,\mathbf{R}^m)$, respectively, extended by 0 on $\mathbf{R}^N\setminus A$ and $\mathbf{R}^N\setminus B$. We have

$$|A|I_A \geq \int_A f(a+\nabla\phi_A)\,dx-\eta,$$

$$|B|I_B \geq \int_B f(a+\nabla\phi_B)\,dx-\eta,$$

Such η-minimizers exist thanks to (11.6) and by a density argument. The function ϕ which coincides with ϕ_A and ϕ_B on A and B, respectively, belongs to $W_0^{1,p}(A\cup B,\mathbf{R}^m)$ so that

$$\begin{aligned}|A\cup B|I_{A\cup B} &\leq \int_{A\cup B} f(a+\nabla\phi)\,dx\\ &= \int_A f(a+\nabla\phi_A)\,dx+\int_B f(a+\nabla\phi_B)\,dx\\ &\leq |A|I_A+|B|I_B+2\eta.\end{aligned}$$

The thesis is obtained after making $\eta\to 0$. Using quite similar arguments, one also obtains

$$|A|I_A \leq |A\setminus B|I_{A\setminus\overline{B}}+|B|I_B \tag{11.11}$$

whenever A and B are two open bounded subsets of $\mathbf{R}^N$ with $B\subset A$ and $|\partial B|=0$.

Applying, respectively, (11.10) and (11.11) to the finite union $\bigcup_{i\in I_\varepsilon}(x_i+\varepsilon_i D')$ and to $A=D$, $B=\bigcup_{i\in I_\varepsilon}(x_i+\varepsilon_i D')$, according to (11.9) and to the growth condition (11.5), we obtain

$$\left|\bigcup_{i\in I_\varepsilon}(x_i+\varepsilon_i D')\right| I_{\bigcup_{i\in I_\varepsilon}(x_i+\varepsilon_i D')} \leq \sum_{i\in I_\varepsilon}|\varepsilon_i D'|I_{x_i+\varepsilon_i D'}, \tag{11.12}$$

$$|D|I_D \leq \beta(1+|a|^p)\varepsilon+\left|\bigcup_{i\in I_\varepsilon}(x_i+\varepsilon_i D')\right| I_{\bigcup_{i\in I_\varepsilon}(x_i+\varepsilon_i D')}\,. \tag{11.13}$$

A change of scale easily gives $I_{x_i+\varepsilon_i D'}=I_{D'}$. Combining now (11.12) and (11.13), one obtains

$$I_D \leq \frac{\beta(1+|a|^p)\varepsilon}{|D|}+I_{D'}.$$

Since ε is arbitrary we have indeed established $I_D\leq I_{D'}$. It is worth noticing that the above subadditivity argument is a particular case of a more general result related to subadditive ergodic processes (see Krengel [262], Dal Maso and Modica [185], or Licht and Michaille [274]). Such a general argument will be used in Chapter 12.

(b) We assume that f satisfies (11.5) and show that Qf satisfies the same conditions. Taking $D=Y=(0,1)^N$ in the definition of Qf, from (11.5), we have

$$\begin{aligned}Qf(a) &\geq \alpha\inf\left\{\int_Y|a+\nabla\phi|^p\,dx\ :\ \phi\in W_0^{1,p}(Y,\mathbf{R}^m)\right\}\\ &\geq \alpha\inf\left\{\left|\int_Y(a+\nabla\phi)\,dx\right|^p:\phi\in W_0^{1,p}(Y,\mathbf{R}^m)\right\}\\ &=\alpha|a|^p.\end{aligned}$$

We have used Jensen's inequality satisfied by the convex function $a \mapsto |a|^p$ in the second inequality above. The upper bound of (11.5) is trivially obtained by taking $\phi = 0$ as an admissible function in the expression of the infima and by using the upper bound condition satisfied by f.

Let us now establish (11.6). Given arbitrary $\eta > 0$, let $\phi_\eta \in W_0^{1,p}(Y, \mathbf{R}^m)$ be such that

$$Qf(b) \geq \int_Y f(b + \nabla\phi_\eta)\, dx - \eta.$$

From (11.6) and Hölder's inequality, we obtain

$$\begin{aligned}
&Qf(a) - Qf(b)\\
&\leq \int_Y f(a+\nabla\phi_\eta)\, dx - \int_Y f(b+\nabla\phi_\eta)\, dx + \eta\\
&\leq \int_Y | f(a+\nabla\phi_\eta) - f(b+\nabla\phi_\eta) |\, dx + \eta\\
&\leq L\,|\,a-b\,|\left(\int_Y (1 + |\,a+\nabla\phi_\eta(x)\,|^{p-1} + |\,b+\nabla\phi_\eta(x)\,|^{p-1})^{\frac{p}{p-1}} dx\right)^{\frac{p-1}{p}} + \eta\\
&\leq CL\,|\,a-b\,|\left(\int_Y (1 + |\,a\,|^p + |\,b\,|^p + |\,b+\nabla\phi_\eta(x)\,|^p) dx\right)^{\frac{p-1}{p}} + \eta, \qquad (11.14)
\end{aligned}$$

where C is a constant depending only on p. On the other hand, from (11.5),

$$\begin{aligned}
\int_Y |\,b+\nabla\phi_\eta(x)\,|^p\, dx &\leq \frac{1}{\alpha}\int_Y f(b+\nabla\phi_\eta)\, dx\\
&\leq \frac{1}{\alpha}(Qf(b)+\eta)\\
&\leq \frac{\beta}{\alpha}(1+|\,b\,|^p) + \frac{\eta}{\alpha}. \qquad (11.15)
\end{aligned}$$

Combining (11.14) and (11.15) and letting $\eta \to 0$, we obtain

$$Qf(a) - Qf(b) \leq L'\,|\,b-a\,|\,(1+|\,a\,|^{p-1} + |\,b\,|^{p-1}),$$

where L' is a constant which depends only on p, $\alpha\ \beta$. We end the proof by interchanging the roles of a and b.

(c) We establish the quasi-convex inequality. Let us set

$$\mathrm{Aff}_0(D, \mathbf{R}^m) := \left\{\phi \in W_0^{1,p}(D, \mathbf{R}^m) \,:\, \phi \text{ piecewise affine}\right\}.$$

Note that according to the density of $\mathrm{Aff}_0(D, \mathbf{R}^m)$ in $W_0^{1,p}(D, \mathbf{R}^m)$ equipped with its strong topology, and to (11.6), we have

$$\forall a \in \mathbf{M}^{m\times N}, \quad Qf(a) = \inf\left\{\frac{1}{|D|}\int_D f(a+\nabla\phi)\, dx \,:\, \phi \in \mathrm{Aff}_0(D, \mathbf{R}^m)\right\}.$$

It is now easily seen that the quasi-convex inequality is satisfied with every $\phi \in \mathrm{Aff}_0(D, \mathbf{R}^m)$. Indeed, since $\phi \in \mathrm{Aff}_0(D, \mathbf{R}^m)$, there exist some open bounded and pairwise disjoint sets

$D_i \subset D$, $i=1,\dots,r$, with $\bar{D}=\cup_{i=1}^r \bar{D}_i$, $|\partial D_i|=0$, and $a_i \in \mathbf{M}^{m\times N}$, $i=1,\dots,r$, such that $\nabla\phi \equiv a_i$ on D_i, which clearly implies

$$\int_D Qf(a+\nabla\phi)\,dx=\sum_{i=1}^r Qf(a+a_i)|D_i|. \tag{11.16}$$

On the other hand, for $\eta>0$ and $i=1,\dots,r$, there exists $\phi_{i,\eta}\in \mathrm{Aff}_0(D_i,\mathbf{R}^m)$ such that

$$Qf(a+a_i)\geq \frac{1}{|D_i|}\int_{D_i} f(a+a_i+\nabla\phi_{i,\eta})\,dx-\eta. \tag{11.17}$$

Let us consider the function $\tilde{\phi}$ defined on D by

$$\tilde{\phi}(x)=\phi(x)+\phi_{i,\eta}(x) \text{ when } x\in D_i.$$

Clearly $\tilde{\phi}$ belongs to $W_0^{1,p}(D,\mathbf{R}^m)$ (actually to $\mathrm{Aff}_0(D,\mathbf{R}^m)$). Summing inequalities (11.17) for $i=1,\dots,r$, equality (11.16) yields

$$\begin{aligned}\int_D Qf(a+\nabla\phi)\,dx &\geq \int_D f(a+\nabla\tilde{\phi})\,dx-\eta|D|\\ &\geq Qf(a)|D|-\eta|D|.\end{aligned}$$

Letting $\eta\to 0$, we obtain the quasi-convex inequality for every $\phi\in\mathrm{Aff}_0(D,\mathbf{R}^m)$. Thanks to the density of $\mathrm{Aff}_0(D,\mathbf{R}^m)$ in $W_0^{1,p}(D,\mathbf{R}^m)$ and to (11.6) satisfied by Qf, the quasi-convex inequality is now satisfied for any $\phi\in W_0^{1,p}(D,\mathbf{R}^m)$.

It remains to establish that Qf is the greatest quasi-convex function less than or equal to f. First notice that $g\leq f$ yields $Qg\leq Qf$. On the other hand, if g is quasi-convex, $Qg=g$. Indeed $Qg\leq g$ by definition and inequality (11.8) satisfied by g gives the converse inequality. We then obtain from $g\leq f$ and the quasi-convexity of g that $g=Qg\leq Qf$. □

Remark 11.2.1. One can prove that Qf is the quasi-convex envelope of f under less restrictive assumptions on f, as, for example, without continuity condition (see Dacorogna [182, Theorem 1.1, Chapter 5] and [82]). For the relationship between the notions of convexity, polyconvexity, rank-one convexity, and various examples, consult Dacorogna [182] and Sverak [340], [341], [342].

One can establish that Qf is convex on each interval $[a,b]$ in $\mathbf{M}^{m\times N}$ satisfying rank$(a-b)=1$, i.e., is rank-one convex. Consequently, when $m=1$ or $N=1$, Qf is a convex function and actually $Qf=f^{**}$. For a proof, consult Step 3′ in the proof of Theorem 1.1, Chapter 5 in Dacorogna [182]. Nevertheless, when $m>1$ or $N>1$, Qf is not in general the convexification of f. Consequently the optimization problems related to integral functionals F with a convex density f are not well-posed. One need to uses a relaxation procedure in the sense of the relaxation Theorem 11.1.2.

Let $f:\mathbf{M}^{m\times N}\longrightarrow\mathbf{R}$ satisfying (11.5) and (11.6). In Theorem 13.2.1, we will establish, in a more general situation, that the quasi-convex inequality (11.8) is a necessary and sufficient condition to ensure the lower semicontinuity of the integral functional $u\mapsto\int_\Omega f(\nabla u)\,dx$ defined on $W^{1,p}(\Omega,\mathbf{R}^m)$ equipped with its weak topology.

We now state the main result of this section.

Theorem 11.2.1. *Let us consider a function $f : \mathbf{R}^N \longrightarrow \mathbf{R}$ satisfying* (11.5) *and* (11.6) *with $p > 1$, and F the associated integral functional* (11.7) *defined in $L^p(\Omega, \mathbf{R}^m)$ equipped with its strong topology. Then the lsc envelope of F is given, for every u in $L^p(\Omega, \mathbf{R}^m)$, by*

$$cl(F)(u) = \begin{cases} \displaystyle\int_\Omega Qf(\nabla u)\,dx & \text{if } u \in W^{1,p}(\Omega, \mathbf{R}^m), \\ +\infty & \text{otherwise.} \end{cases}$$

The proof of Theorem 11.2.1 is the straightforward consequence of Propositions 11.2.3 and 11.2.4 below. We denote the functional

$$QF(u) = \begin{cases} \displaystyle\int_\Omega Qf(\nabla u)\,dx & \text{if } u \in W^{1,p}(\Omega, \mathbf{R}^m), \\ +\infty & \text{otherwise} \end{cases}$$

by QF. The proofs given here have the advantage of being easily adapted to the theory of homogenization (see the next chapter).

Proposition 11.2.3. *For every u in $L^p(\Omega, \mathbf{R}^m)$, $p > 1$, and every sequence $(u_n)_{n\in\mathbf{N}}$ strongly converging to u in $L^p(\Omega, \mathbf{R}^m)$, one has*

$$QF(u) \le \liminf_{n\to+\infty} F(u_n). \tag{11.18}$$

Assume that f satisfies (11.6) *and only the upper growth condition $0 \le f(a) \le \beta(1 + |a|^p)$ for all $a \in \mathbf{M}^{m\times N}$. Then, for every u in $W^{1,p}(\Omega, \mathbf{R}^m)$ and every sequence $(u_n)_{n\in\mathbf{N}}$ weakly converging to u in $W^{1,p}(\Omega, \mathbf{R}^m)$, one has*

$$QF(u) \le \liminf_{n\to+\infty} F(u_n).$$

PROOF. We assume that f satisfies (11.5) and (11.6) and we establish the first assertion. The second assertion will be obtained at the end of the proof as a straightforward consequence. Obviously, one can assume $\liminf_{n\to+\infty} F(u_n) < +\infty$ so that u belongs to $W^{1,p}(\Omega, \mathbf{R}^m)$. For a nonrelabeled subsequence, consider the nonnegative Borel measure $\mu_n := f(\nabla u_n(.))\mathcal{L}^N\lfloor\Omega$. We have

$$\sup_{n\in\mathbf{N}} \mu_n(\Omega) < +\infty.$$

Consequently there exists a further subsequence (not relabeled) and a nonnegative Borel measure $\mu \in \mathbf{M}(\Omega)$ such that

$$\mu_n \rightharpoonup \mu \quad \text{weakly in } \mathbf{M}(\Omega).$$

Let $\mu = g\mathcal{L}^N\lfloor\Omega + \mu_s$ be the Lebesgue–Nikodým decomposition of μ where μ_s is a nonnegative Borel measure, singular with respect to the Lebesgue measure $\mathcal{L}^N\lfloor\Omega$. For establishing (11.18) it is enough to prove that for a.e. $x \in \Omega$,

$$g(x) \ge Qf(\nabla u(x)).$$

Indeed, according to Alexandrov's theorem, Theorem 4.2.3, we will obtain

$$\begin{aligned} \liminf_{n\to+\infty} F(u_n) = \liminf_{n\to+\infty} \mu_n(\Omega) \ge \mu(\Omega) &= \int_\Omega g(x)\,dx + \mu_s(\Omega) \\ &\ge \int_\Omega g(x)\,dx \\ &\ge \int_\Omega Qf(\nabla u(x))\,dx. \end{aligned}$$

Let $\rho > 0$ intended to tend to 0 and denote the open ball of radius ρ centered at x_0 by $B_\rho(x_0)$. According to the theory of differentiation of measures (see Theorem 4.2.1), there exists a negligible set N for the measure $\mathscr{L}^N\lfloor\Omega$ such that for all $x_0 \in \Omega \setminus N$,

$$g(x_0) = \lim_{\rho\to 0} \frac{\mu(B_\rho(x_0))}{|B_\rho(x_0)|}.$$

Applying Lemma 4.2.1, for all but countably many $\rho > 0$, one may assume $\mu(\partial B_\rho(x_0)) = 0$. From Alexandrov's theorem, Theorem 4.2.3, we then obtain $\mu(B_\rho(x_0)) = \lim_{n\to+\infty} \mu_n(B_\rho(x_0))$ and we finally are reduced to establishing

$$\lim_{\rho\to 0}\lim_{n\to+\infty} \frac{\mu_n(B_\rho(x_0))}{|B_\rho(x_0)|} \geq Qf(\nabla u(x_0)) \quad \text{for } x_0 \in \Omega \setminus N. \tag{11.19}$$

Let us assume for the moment that the trace of u_n on $\partial B_\rho(x_0)$ is equal to the affine function u_0 defined by $u_0(x) := u(x_0) + \langle \nabla u(x_0), x - x_0\rangle$. It follows that

$$\begin{aligned}
\lim_{n\to+\infty} \frac{\mu_n(B_\rho(x_0))}{|B_\rho(x_0)|} &= \lim_{n\to+\infty} \frac{1}{|B_\rho(x_0)|} \int_{B_\rho(x_0)} f(\nabla u(x_0) + \nabla(u_n - u_0))\, dx \\
&\geq \inf\left\{ \frac{1}{|B_\rho(x_0)|} \int_{B_\rho(x_0)} f(\nabla u(x_0) + \nabla\phi)\, dx : \phi \in W_0^{1,p}(B_\rho(x_0), \mathbf{R}^m) \right\} \\
&= Qf(\nabla u(x_0))
\end{aligned}$$

and the proof would be complete. Note that the small size of the radius ρ of $B_\rho(x_0)$ does not contribute to the estimate above. The idea now consists in modifying u_n by a function of $W^{1,p}(B_\rho(x_0), \mathbf{R}^m)$ which coincides with u_0 on $\partial B_\rho(x_0)$ in the trace sense, to follow the previous procedure and to control additional terms, when ρ goes to zero, thanks to the following classical estimate.

Lemma 11.2.1. *For every u in $W^{1,p}(\Omega, \mathbf{R}^m)$, $p \geq 1$, there exists a negligible set N' such that for all x_0 in $\Omega \setminus N'$ we have*

$$\left[\frac{1}{|B_\rho(x_0)|} \int_{B_\rho(x_0)} |u(x) - \big(u(x_0) + \nabla u(x_0)(x - x_0)\big)|^p\, dx \right]^{1/p} = o(\rho). \tag{11.20}$$

PROOF. For the proof we refer to Ziemer [366, Theorem 3.4.2]. □

From now on, we fix x_0 in the set $\Omega \setminus N \cup N'$. In order to suitably modify u_n on the boundary of $B_\rho(x_0)$, we use a well-known method due to De Giorgi. A neighborhood of the boundary of $B_\rho(x_0)$ is sliced as follows: let $\nu \in \mathbf{N}^*$, $\rho_0 := \lambda\rho$, where $0 < \lambda < 1$, and set

$$B_i := B_{\rho_0 + i\frac{\rho-\rho_0}{\nu}}(x_0) \quad \text{for } i = 0, \dots, \nu.$$

On the other hand, consider for $i = 1, \dots, \nu$,

$$\varphi_i \in \mathbf{C}_0^\infty(B_i),\ 0 \leq \varphi_i \leq 1,\ \varphi_i = 1 \text{ in } B_{i-1}, \quad |\mathrm{grad}\varphi_i|_{L^\infty(B_i)} \leq \frac{\nu}{\rho - \rho_0} = \frac{\nu}{\rho(1-\lambda)},$$

and define $u_{n,i} \in W^{1,p}(B_\rho(x_0), \mathbf{R}^m)$ by

$$u_{n,i} := u_0 + \varphi_i(u_n - u_0).$$

For $i = 1, \ldots, \nu$, we have

$$\begin{aligned}
&Qf(\nabla u(x_0)) \\
&= \inf\left\{ \frac{1}{|B_\rho(x_0)|} \int_{B_\rho(x_0)} f(\nabla u(x_0) + \nabla \phi)\, dx : \phi \in W_0^{1,p}(B_\rho(x_0), \mathbf{R}^m) \right\} \\
&\leq \frac{1}{|B_\rho(x_0)|} \int_{B_\rho(x_0)} f(\nabla u_{n,i})\, dx \\
&= \frac{1}{|B_\rho(x_0)|} \int_{B_{i-1}} f(\nabla u_n)\, dx + \frac{1}{|B_\rho(x_0)|} \int_{B_i \setminus B_{i-1}} f(\nabla u_{n,i})\, dx \\
&\qquad + \frac{1}{|B_\rho(x_0)|} \int_{B_\rho(x_0) \setminus B_i} f(\nabla u(x_0))\, dx \\
&\leq \frac{1}{|B_\rho(x_0)|} \int_{B_\rho(x_0)} f(\nabla u_n)\, dx + \frac{1}{|B_\rho(x_0)|} \int_{B_i \setminus B_{i-1}} f(\nabla u_{n,i})\, dx \\
&\qquad + \beta\big(1 + |\nabla u(x_0)|^p\big)(1-\lambda)^N.
\end{aligned} \tag{11.21}$$

Let us estimate the second term of the right-hand side of (11.21). From the growth condition (11.5) we obtain

$$\begin{aligned}
&\frac{1}{|B_\rho(x_0)|} \int_{B_i \setminus B_{i-1}} f(\nabla u_{n,i}) dx \leq C\big(1 + |\nabla u(x_0)|^p\big)(1-\lambda)^N \\
&+ \frac{C}{|B_\rho(x_0)|} \int_{B_i \setminus B_{i-1}} |\nabla(u_n - u_0)|^p\, dx + \frac{C}{|B_\rho(x_0)|} \frac{\nu^p}{\rho^p(1-\lambda)^p} \int_{B_i \setminus B_{i-1}} |u_n - u_0|^p\, dx,
\end{aligned}$$

where C is a positive constant depending only on p and β. Then, averaging inequalities (11.21), we obtain

$$\begin{aligned}
Qf(\nabla u(x_0)) &= \frac{1}{\nu} \sum_{i=1}^{\nu} Qf(\nabla u(x_0)) \\
&\leq \frac{1}{|B_\rho(x_0)|} \int_{B_\rho(x_0)} f(\nabla u_n)\, dx + \frac{C\nu^{p-1}}{|B_\rho(x_0)|\rho^p(1-\lambda)^p} \int_{B_\rho(x_0)} |u_n - u_0|^p\, dx \\
&\quad + C\big(1 + |\nabla u(x_0)|^p\big)(1-\lambda)^N + \frac{1}{\nu} \frac{C}{|B_\rho(x_0)|} \int_{B_\rho(x_0)} |\nabla u_n|^p\, dx \\
&\leq \frac{\frac{C}{\alpha\nu} + 1}{|B_\rho(x_0)|} \int_{B_\rho(x_0)} f(\nabla u_n)\, dx + \frac{C\nu^{p-1}}{|B_\rho(x_0)|\rho^p(1-\lambda)^p} \int_{B_\rho(x_0)} |u_n - u_0|^p\, dx \\
&\quad + C\big(1 + |\nabla u(x_0)|^p\big)(1-\lambda)^N,
\end{aligned} \tag{11.22}$$

where we have used the lower bound (11.5) in the last inequality. Letting $n \to +\infty$ and $\rho \to 0$, from Lemma 11.2.1 we obtain

$$Qf(\nabla u(x_0)) \leq \left(\frac{C}{\alpha\nu} + 1\right) \lim_{\rho\to 0} \lim_{n\to+\infty} \frac{\mu_n(B_\rho(x_0))}{|B_\rho(x_0)|} + C\big(1 + |\nabla u(x_0)|^p\big)(1-\lambda)^N$$

and (11.19) follows after letting $\lambda\to 1$ and $\nu\to+\infty$. It is worth noticing that we have used the slicing method in order to control the term $\frac{1}{|B_\rho(x_0)|}\int_{B_\rho(x_0)}|\nabla u_n|^p\,dx$ by letting the slices become increasingly thin (i.e., $\nu\to+\infty$).

If now f does not verify the lower bound in (11.5), we first note that the weak convergence of u_n to u in $W^{1,p}(\Omega,\mathbf{R}^m)$ yields the boundedness of

$$\sup_{n\in\mathbf{N}}\int_{B_\rho(x_0)}|\nabla u_n|^p\,dx<+\infty.$$

Moreover, according to the Rellich–Kondrakov compact embedding theorem, Theorem 5.4.2, $u_n\to u$ strongly in $L^p(\Omega,\mathbf{R}^m)$. Then, to conclude, it suffices going to the limit respectively on n and ρ in (11.22) and to let $\lambda\to 1$, $\nu\to+\infty$ as above. □

Proposition 11.2.4. *For every u in $L^p(\Omega,\mathbf{R}^m)$, $p>1$, there exists a sequence $(u_n)_{n\in\mathbf{N}}$ strongly converging to u in $L^p(\Omega,\mathbf{R}^m)$ such that*

$$QF(u)\geq\limsup_{n\to+\infty}F(u_n).$$

PROOF. One can assume $QF(u)<+\infty$. Therefore, taking $D=Y:=(0,1)^N$ in the definition of Qf, and $\eta=1/k$, $k\in\mathbf{N}^*$,

$$\begin{aligned}QF(u)&=\int_\Omega Qf(\nabla u)\,dx\\&=\int_\Omega\inf\left\{\int_Y f(\nabla u(x)+\nabla_y\phi(y))\,dy\ :\ \phi\in W_0^{1,p}(Y,\mathbf{R}^m)\right\}dx\\&\geq\int_{\Omega\times Y}f(\nabla u(x)+\nabla_y\phi_\eta(x,y))\,dxdy-\eta|\Omega|,\end{aligned}\tag{11.23}$$

where $\phi_\eta(x,\cdot)$ is a η-minimizer of $\inf\left\{\int_Y f(\nabla u(x)+\nabla_y\phi(y))\,dy\ :\ \phi\in W_0^{1,p}(Y,\mathbf{R}^m)\right\}$. We admit that we can select a measurable map $x\mapsto\phi_\eta(x,\cdot)$ from Ω into $W_0^{1,p}(Y,\mathbf{R}^m)$. For a proof, consult Castaing and Valadier [166]. We claim that ϕ_η belongs to $L^p(\Omega,W_0^{1,p}(Y,\mathbf{R}^m))$. Indeed

$$\begin{aligned}\int_\Omega\|\nabla_y\phi_\eta(x,\cdot)\|^p_{L^p(Y,\mathbf{M}^{m\times N})}\,dx&=\int_{\Omega\times Y}|\nabla_y\phi_\eta(x,y)|^p\,dxdy\\&\leq C\left(\int_{\Omega\times Y}|\nabla u(x)+\nabla_y\phi_\eta(x,y)|^p\,dxdy+\int_\Omega|\nabla u|^p\,dx\right)\\&\leq C\left(\frac{1}{\alpha}\int_{\Omega\times Y}f(\nabla u(x)+\nabla_y\phi_\eta(x,y))\,dxdy+\int_\Omega|\nabla u|^p\,dx\right)\\&\leq C\left(\frac{1}{\alpha}QF(u)+\frac{\eta|\Omega|}{\alpha}+\int_\Omega|\nabla u|^p\,dx\right)<+\infty,\end{aligned}$$

where C is a positive constant depending only on p.

Classically $\mathbf{C}_c(\Omega,\mathcal{D}(Y,\mathbf{R}^m))$ is dense in $L^p(\Omega,W_0^{1,p}(Y,\mathbf{R}^m))$. Consequently, from the continuity assumption (11.6) fulfilled by f, it is easily seen that (11.23) yields

$$QF(u)=\int_\Omega Qf(\nabla u)\,dx\geq\int_{\Omega\times Y}f(\nabla u(x)+\nabla_y\tilde{\phi}_\eta(x,y))\,dxdy-2\eta|\Omega|\tag{11.24}$$

for some $\tilde{\phi}_\eta$ in $\mathbf{C}_c(\Omega, \mathscr{D}(Y, \mathbf{R}^m))$. We have actually established the following interchange result between infimum and integral:

$$\begin{aligned}
&\int_\Omega \left(\inf_{\phi \in W_0^{1,p}(Y, \mathbf{R}^m)} \int_Y f(\nabla u(x) + \nabla_y \phi(y))\, dy \right) dx \\
&= \inf_{\Phi \in L^p(\Omega, W_0^{1,p}(Y, \mathbf{R}^m))} \int_{\Omega \times Y} f(\nabla u(x) + \nabla_y \Phi(x,y))\, dx dy \\
&= \inf_{\Phi \in \mathbf{C}_c(\Omega, \mathscr{D}(Y, \mathbf{R}^m))} \int_{\Omega \times Y} f(\nabla u(x) + \nabla_y \Phi(x,y))\, dx dy.
\end{aligned}$$

Indeed from (11.23)

$$\begin{aligned}
&\int_\Omega \left(\inf_{\phi \in W_0^{1,p}(Y, \mathbf{R}^m)} \int_Y f(\nabla u(x) + \nabla_y \phi(y))\, dy \right) dx \\
&\geq \inf_{\Phi \in L^p(\Omega, W_0^{1,p}(Y, \mathbf{R}^m))} \int_{\Omega \times Y} f(\nabla u(x) + \nabla_y \Phi(x,y))\, dx dy
\end{aligned}$$

and the converse inequality is trivial. Note that this result also holds for $p = 1$. Because of its importance, we state it in a slightly more general form.

Lemma 11.2.2. *Let $f : \mathbf{M}^{m \times N} \to \mathbf{R}^+$ be any function satisfying* (11.5) *and* (11.6) *with $p \geq 1$. Let moreover ξ be any element of $L^p(\Omega, \mathbf{M}^{m \times N})$. Then*

$$\begin{aligned}
&\int_\Omega \left(\inf_{\phi \in W_0^{1,p}(Y, \mathbf{R}^m)} \int_Y f(\xi(x) + \nabla_y \phi(y))\, dy \right) dx \\
&= \inf_{\Phi \in L^p(\Omega, W_0^{1,p}(Y, \mathbf{R}^m))} \int_{\Omega \times Y} f(\xi(x) + \nabla_y \Phi(x,y))\, dx dy \\
&= \inf_{\Phi \in \mathbf{C}_c(\Omega, \mathscr{D}(Y, \mathbf{R}^m))} \int_{\Omega \times Y} f(\xi(x) + \nabla_y \Phi(x,y))\, dx dy.
\end{aligned}$$

For more about interchange theorems, consult, for instance, Anza Hafsa and Mandallena [34].

PROOF OF PROPOSITION 11.2.4 CONTINUED. Let us go back to (11.24). To shorten the notation we denote the previous η-minimizer $\tilde{\phi}_\eta$ in $\mathbf{C}_c(\Omega, \mathscr{D}(Y, \mathbf{R}^m))$ by ϕ_η and extend $y \mapsto \phi_\eta(x,y)$ by Y-periodicity on $\mathbf{R}^N$. Consider now the function $u_{\eta,n}$ defined by

$$u_{\eta,n}(x) = u(x) + \frac{1}{n} \phi_\eta(x, nx).$$

Note that ϕ_η is a Carathéodory function so that $x \mapsto \phi_\eta(x, nx)$ is measurable. Clearly $u_{\eta,n}$ belongs to $W^{1,p}(\Omega, \mathbf{R}^m)$ and $u_{\eta,n} \to u$ strongly in $L^p(\Omega, \mathbf{R}^m)$ when $n \to +\infty$. It's indeed a straightforward consequence of

$$\begin{aligned}
\int_\Omega |\phi_\eta(x, nx)|^p\, dx &\leq \int_\Omega \sup_{y \in Y} |\phi_\eta(x,y)|^p\, dx \\
&\leq |\Omega| \sup_{x \in \Omega} \sup_{y \in Y} |\phi_\eta(x,y)|^p < +\infty.
\end{aligned}$$

On the other hand,

$$\begin{aligned}\lim_{n\to+\infty}\int_\Omega f(\nabla u_{\eta,n})\,dx &= \lim_{n\to+\infty}\int_\Omega f\Big(\nabla u(x)+(\nabla_y\phi_\eta)(x,nx)+\frac{1}{n}\nabla_x\phi_\eta(x,nx)\Big)dx\\ &= \lim_{n\to+\infty}\int_\Omega f(\nabla u(x)+(\nabla_y\phi_\eta)(x,nx))\,dx\\ &= \int_{\Omega\times Y} f(\nabla u(x)+\nabla_y\phi_\eta(x,y))\,dxdy.\end{aligned}$$

For passing from the first to the second equality we have used the continuity assumption (11.6) on f. The last equality is a consequence of Lemma 11.2.3 stated at the end of the proof and applied to $g(x,y)=f(\nabla u(x)+\nabla_y\phi_\eta(x,y))$. (Note that $y\mapsto f(\nabla u(x)+\nabla_y\phi_\eta(x,y))$ belongs to the set $\mathbf{C}_\#(Y)$ of all the restrictions to Y of continuous and Y-periodic functions on $\mathbf{R}^N$ and that $g\in L^1(\Omega,\mathbf{C}_\#(Y))$.) Consequently, from (11.24)

$$\begin{aligned}\lim_{n\to+\infty}F(u_{\eta,n}) &= \lim_{n\to+\infty}\int_\Omega f(\nabla u(x)+(\nabla_y\phi_\eta)(x,nx))\,dx\\ &= \int_{\Omega\times Y} f(\nabla u(x)+\nabla_y\phi_\eta(x,y))\,dxdy\\ &\le QF(u)+2\eta.\end{aligned}$$

Letting $\eta\to 0$ (i.e., $k\to+\infty$), up to a subsequence, one obtains

$$\lim_{\eta\to+0}\lim_{n\to+\infty}F(u_{\eta,n})\le QF(u).$$

We apply now the diagonalization Lemma 11.1.1 for the sequence $\big(F(u_{\eta,n}),u_{\eta,n}\big)_{\eta,n}$ in the metric space $\mathbf{R}\times L^p(\Omega,\mathbf{R}^m)$: there exists a map $n\mapsto\eta(n):=\frac{1}{k(n)}$ such that

$$\begin{aligned}&\lim_{n\to+\infty}F(u_{\eta(n),n})\le QF(u),\\ &\lim_{n\to+\infty}u_{\eta(n),n}=u \quad \text{strongly in } L^p(\Omega,\mathbf{R}^m).\end{aligned}$$

The sequence $(u_n)_{n\in\mathbf{N}}$ where $u_n=u_{\eta(n),n}$ then verifies the assertion of Proposition 11.2.4.

We now state and prove Lemma 11.2.3 invoked above. Let D be any open cube of $\mathbf{R}^N$. We recall that $\mathbf{C}_\#(D)$ denotes the set of all the restrictions to D of continuous and D-periodic functions on $\mathbf{R}^N$, equipped with the uniform norm on D, and that $L^1(\Omega,\mathbf{C}_\#(D))$ denotes the space of all measurable functions h from Ω into $\mathbf{C}_\#(D)$ satisfying

$$\int_\Omega\sup_{y\in D}|h(x,y)|\,dx<+\infty.$$

Lemma 11.2.3. *For every function g in $L^1(\Omega,\mathbf{C}_\#(D))$,*

$$\lim_{n\to+\infty}\int_\Omega g(x,nx)\,dx=\frac{1}{|D|}\int_{\Omega\times D}g(x,y)\,dxdy. \tag{11.25}$$

PROOF. It is well known (see, for instance, Yosida [361]) that if g belongs to $L^1(\Omega,\mathbf{C}_\#(D))$, then g is a Carathéodory function satisfying $\sup_{y\in D}|g(.,y)|\in L^1(\Omega)$ so that $\int_\Omega g(x,nx)\,dx$

is well defined. For $k \in \mathbf{N}^*$, let us decompose the cube D as follows:

$$\overline{D} = \bigcup_{i=1}^{k^N} \overline{D}_i,$$

where D_i are small pairwise disjoint open cubes $1/k$-homothetic of D. We approximate g in $L^1(\Omega, \mathbf{C}_\#(D))$ by the following step function g_k:

$$g_k(x,y) = \sum_{i=1}^{k^N} g(x,y_i) 1_{D_i}(y),$$

where 1_{D_i} is the characteristic function of the set D_i extended by D-periodicity on $\mathbf{R}^N$ and y_i is any fixed element of D_i. Due to the periodicity of 1_{D_i}, classically $x \mapsto 1_{D_i}(nx)$ $\sigma(L^\infty, L^1)$ weakly converges to $|D_i|/|D|$ so that

$$\lim_{n\to+\infty} \int_\Omega g(x,y_i) 1_{D_i}(nx)\, dx = \frac{|D_i|}{|D|} \int_\Omega g(x,y_i)\, dx.$$

For a proof, see Example 2.4.2 or Proposition 13.2.1 and the proof of Theorem 13.2.1. From now on, to shorten notation, we assume $|D| = 1$. Summing these equalities over $i = 1, \ldots, k^N$, one obtains that (11.25) is satisfied for g_k.

In order to conclude by going to the limit on k, we use the uniform bound with respect to n:

$$\int_\Omega |g(x,nx)\, dx - g_k(x,nx)|\, dx \le ||g - g_k||_{L^1(\Omega,\mathbf{C}_\#(D))}. \tag{11.26}$$

Moreover, since

$$\sup_{y\in D} |g_k(x,y) - g(x,y)| \le 2 \sup_{y\in D} |g(x,y)| \in L^1(\Omega)$$

and $\lim_{k\to+\infty} \sup_{y\in D} |g_k(x,y) - g(x,y)| = 0$, according to the Lebesgue dominated convergence theorem, we obtain

$$\lim_{k\to+\infty} ||g - g_k||_{L^1(\Omega,\mathbf{C}_\#(D))} = 0. \tag{11.27}$$

Thus, from (11.26)

$$\begin{aligned}
\left| \int_\Omega g(x,nx)\, dx - \int_{\Omega\times D} g(x,y)\, dxdy \right| &\le \left| \int_\Omega g(x,nx)\, dx - \int_\Omega g_k(x,nx)\, dx \right| \\
&\quad + \left| \int_\Omega g_k(x,nx)\, dx - \int_{\Omega\times D} g_k(x,y)\, dxdy \right| \\
&\quad + \left| \int_{\Omega\times D} g_k(x,y)\, dxdy - \int_{\Omega\times D} g(x,y)\, dxdy \right| \\
&\le 2||g - g_k||_{L^1(\Omega,\mathbf{C}_\#(D))} \\
&\quad + \left| \int_\Omega g_k(x,nx)\, dx - \int_{\Omega\times Y} g_k(x,y)\, dxdy \right|.
\end{aligned}$$

We conclude first by letting $n\to+\infty$ and using (11.25) satisfied by g_k and then by letting $k\to+\infty$ and using (11.27). □

Now, we would like to compute the lower semicontinuous envelope of the integral functional in (11.7) by taking into account a boundary condition on a part Γ_0 of the boundary $\partial\Omega$ of Ω. More precisely, we aim to describe the lower semicontinuous envelope of the integral functional $F: L^p(\Omega,\mathbf{R}^m)\longrightarrow\mathbf{R}^+\cup\{+\infty\}$ defined by

$$F(u)=\begin{cases}\displaystyle\int_\Omega f(\nabla u)\,dx & \text{if } u\in W^{1,p}_{\Gamma_0}(\Omega,\mathbf{R}^m),\\ +\infty & \text{otherwise,}\end{cases}\tag{11.28}$$

where $W^{1,p}_{\Gamma_0}(\Omega,\mathbf{R}^m)$ denotes the subspace of all the functions u in $W^{1,p}(\Omega,\mathbf{R}^m)$ such that $u=0$ on Γ_0 in the sense of traces. In the following corollary, we state that the boundary condition is "not relaxed." We will see that the preservation of the boundary condition is not satisfied in the case $p=1$.

Corollary 11.2.1. *The lower semicontinuous envelope of the integral functional F defined in* (11.28) *is given by*

$$cl(F)(u)=\begin{cases}\displaystyle\int_\Omega Qf(\nabla u)\,dx & \textit{if } u\in W^{1,p}_{\Gamma_0}(\Omega,\mathbf{R}^m),\\ +\infty & \textit{otherwise.}\end{cases}$$

PROOF. We only have to establish the existence of a sequence $(u_n)_{n\in\mathbf{N}}$ strongly converging to u in $L^p(\Omega,\mathbf{R}^m)$ such that $\mathrm{cl}(F)(u)\geq\limsup_{n\to+\infty}F(u_n)$. Assuming $\mathrm{cl}(F)(u)<+\infty$, we must construct a sequence $(u_n)_{n\in\mathbf{N}}$ in $W^{1,p}_{\Gamma_0}(\Omega,\mathbf{R}^m)$, strongly converging to u in $L^p(\Omega,\mathbf{R}^m)$ and satisfying $\mathrm{cl}(F)(u)\geq\limsup_{n\to+\infty}F(u_n)$.

According to Theorem 11.2.1, there exists a sequence $(v_n)_{n\in\mathbf{N}}$ in $W^{1,p}(\Omega,\mathbf{R}^m)$ strongly converging to u in $L^p(\Omega,\mathbf{R}^m)$ such that

$$\int_\Omega Qf(\nabla u)\,dx\geq\limsup_{n\to+\infty}\int_\Omega f(\nabla v_n)\,dx.\tag{11.29}$$

The idea is now to modify v_n on a neighborhood of $\partial\Omega$ so that the new function belongs to $W^{1,p}_{\Gamma_0}(\Omega,\mathbf{R}^m)$ and in such a way to decrease the energy. We use again the slicing method of De Giorgi. Let $\nu\in\mathbf{N}^*$ and $\Omega_0\subset\subset\Omega$ such that

$$\int_{\Omega\setminus\Omega_0}\left(1+|\nabla u|^p\right)dx\leq\frac{1}{\nu}\tag{11.30}$$

and $(\Omega_i)_{i=0,\dots,\nu}$ an increasing sequence of open subsets strictly included in Ω, $\Omega_i\subset\subset\Omega_{i+1}\subset\subset\Omega$. Let $(\varphi_i)_{i=0,\dots,\nu-1}$ be a sequence of functions in $\mathcal{D}(\mathbf{R}^N)$ satisfying

$$\begin{aligned}&\varphi_i=1\quad\text{on }\Omega_i,\qquad \varphi_i=0\quad\text{on }\mathbf{R}^N\setminus\Omega_{i+1},\ 0\leq\varphi_i\leq1,\\ &|\nabla\varphi_i|\leq\frac{\nu}{d},\end{aligned}$$

where $d=\mathrm{dist}(\Omega_0,\mathbf{R}^N\setminus\overline{\Omega})$, and define

$$u_{n,i}=\varphi_i(v_n-u)+u.$$

Clearly $u_{n,i}$ belongs to $W^{1,p}_{\Gamma_0}(\Omega,\mathbf{R}^m)$ and

$$\begin{aligned}\int_\Omega f(\nabla u_{n,i})\,dx &= \int_{\Omega\setminus\Omega_{i+1}} f(\nabla u_{n,i})\,dx + \int_{\Omega_{i+1}\setminus\Omega_i} f(\nabla u_{n,i})\,dx + \int_{\Omega_i} f(\nabla u_{n,i})\,dx\\ &\le \int_{\Omega\setminus\Omega_0} f(\nabla u)\,dx + \int_{\Omega_{i+1}\setminus\Omega_i} f(\nabla u_{n,i})\,dx + \int_\Omega f(\nabla v_n)\,dx.\end{aligned}$$

Then, from (11.30) and the growth condition in (11.5), we obtain

$$\int_\Omega f(\nabla u_{n,i})\,dx \le C\left(\frac{1}{\nu}+\left(\frac{\nu}{d}\right)^p\int_\Omega |v_n-u|^p\,dx+\int_{\Omega_{i+1}\setminus\Omega_i}(|\nabla v_n|^p)\,dx\right)+\int_\Omega f(\nabla v_n)\,dx,$$

where, from now on, C denotes various positive constants depending only on β, p, and Ω. By averaging these ν inequalities, we obtain

$$\frac{1}{\nu}\sum_{i=0}^{\nu-1}\int_\Omega f(\nabla u_{n,i})\,dx \le C\left(\frac{1}{\nu}+\left(\frac{\nu}{d}\right)^p\int_\Omega |v_n-u|^p\,dx+\frac{1}{\nu}\int_\Omega |\nabla v_n|^p\,dx\right)+\int_\Omega f(\nabla v_n)\,dx.$$

As already said in the proof of Proposition 11.2.3, we have used a slicing method in order to control the term $\int_\Omega |\nabla v_n|^p\,dx$ by taking increasingly thin slices (i.e., $\nu\to+\infty$). We could not conclude by using a simple truncation.

From the coercivity condition (11.5), $\int_\Omega |\nabla v_n|^p\,dx$ is bounded, hence

$$\frac{1}{\nu}\sum_{i=0}^{\nu-1}\int_\Omega f(\nabla u_{n,i})\,dx \le C\left(\frac{1}{\nu}+\left(\frac{\nu}{d}\right)^p\int_\Omega |v_n-u|^p\,dx\right)+\int_\Omega f(\nabla v_n)\,dx. \tag{11.31}$$

Let $i(n,\nu)$ be the index i such that

$$\int_\Omega f(\nabla u_{n,i(n,\nu)})\,dx = \min_{i=0,\dots,\nu-1}\int_\Omega f(\nabla u_{n,i})\,dx.$$

Inequality (11.31) yields

$$\int_\Omega f(\nabla u_{n,i(n,\nu)})\,dx \le C\left(\frac{1}{\nu}+\left(\frac{\nu}{d}\right)^p\int_\Omega |v_n-u|^p\,dx\right)+\int_\Omega f(\nabla v_n)\,dx$$

so that from (11.29)

$$\limsup_{\nu\to+\infty}\limsup_{n\to+\infty} F(u_{n,i(n,\nu)}) \le \limsup_{n\to+\infty} F(v_n) \le \int_\Omega Qf(\nabla u)\,dx.$$

We conclude by a classical diagonalization argument: there exists $n\mapsto \nu(n)$ mapping $\mathbf{N}$ into $\mathbf{N}$ such that

$$\limsup_{n\to+\infty} F(u_{n,i(n,\nu(n))}) \le \limsup_{\nu\to+\infty}\limsup_{n\to+\infty} F(u_{n,i(n,\nu)}) \le \limsup_{n\to+\infty} F(v_n) \le \int_\Omega Qf(\nabla u)\,dx.$$

Obviously

$$\lim_{n\to+\infty} u_{n,i(n,\nu(n))} = u \text{ strongly in } L^p(\Omega,\mathbf{R}^m).$$

The sequence $(u_n)_{n\in\mathbf{N}}$ defined by $u_n = u_{n,i(n,\nu(n))}$ then tends to u in $L^p(\Omega,\mathbf{R}^m)$ and satisfies $\int_\Omega Qf(\nabla u)\,dx \geq \limsup_{n\to+\infty} F(u_n)$. □

As a consequence we obtain the following relaxation theorem in the case $p>1$.

Theorem 11.2.2 (relaxation theorem, $p>1$). *Let us consider a function $f:\mathbf{M}^{m\times N}\longrightarrow\mathbf{R}$ satisfying* (11.5), (11.6), *a function g in $L^q(\Omega,\mathbf{R}^m)$, where $\frac{1}{p}+\frac{1}{q}=1$, and the following problem:*

$$\inf\left\{\int_\Omega f(\nabla u)\,dx - \int_\Omega g.u\;dx\;:\;u\in W^{1,p}_{\Gamma_0}(\Omega,\mathbf{R}^m)\right\}. \qquad (\mathscr{P})$$

Then, the relaxed problem of $(\mathscr{P})$ in the sense of Theorem 11.1.2 *is*

$$\inf\left\{\int_\Omega Qf(\nabla u)\,dx - \int_\Omega g.u\;dx\;:\;u\in W^{1,p}_{\Gamma_0}(\Omega,\mathbf{R}^m)\right\}. \qquad (\overline{\mathscr{P}})$$

PROOF. With the notation of Theorem 11.2.1 one has

$$\inf\left\{\int_\Omega f(\nabla u)\,dx - \int_\Omega g.u\;dx\;:\;u\in W^{1,p}_{\Gamma_0}(\Omega,\mathbf{R}^m)\right\} = \inf_{u\in L^p(\Omega,\mathbf{R}^m)}\left\{F(u)-\int_\Omega g.u\;dx\right\}$$

and

$$\inf\left\{\int_\Omega Qf(\nabla u)\,dx - \int_\Omega g.u\;dx\;:\;u\in W^{1,p}_{\Gamma_0}(\Omega,\mathbf{R}^m)\right\} = \inf_{u\in L^p(\Omega,\mathbf{R}^m)}\left\{\mathrm{cl}(F)(u)-\int_\Omega g.u\;dx\right\}.$$

Since $G: u\mapsto \int_\Omega g.u\,dx$ is a continuous perturbation of F, one has $\mathrm{cl}(F+G)=\mathrm{cl}(F)+G$ in $L^p(\Omega,\mathbf{R}^m)$. Then, according to Theorems 11.1.2 and 11.2.1, it suffices to establish the inf-compactness of

$$u\mapsto F(u)-\int_\Omega g.u\;dx$$

in $L^p(\Omega,\mathbf{R}^m)$ equipped with its strong topology. Let $u\in L^p(\Omega,\mathbf{R}^m)$ such that $F(u)-\int_\Omega g.u\,dx \leq C$, where C is any positive constant. Then $u\in W^{1,p}(\Omega,\mathbf{R}^m)$ and

$$F(u)-\int_\Omega g.u\;dx = \int_\Omega f(\nabla u)\,dx - \int_\Omega g.u\;dx \leq C.$$

From (11.5) and Hölder's inequality, we obtain

$$\alpha\int_\Omega |\nabla u|^p\;dx \leq \|g\|_{L^q(\Omega,\mathbf{R}^m)}\|u\|_{L^p(\Omega,\mathbf{R}^m)} + C. \qquad (11.32)$$

Applying Young's inequality $ab\leq \frac{\lambda^p a^p}{p}+\frac{1}{\lambda^q}\frac{b^q}{q}$ with $a=\|u\|_{L^p(\Omega,\mathbf{R}^m)}$, $b=\|g\|_{L^q(\Omega,\mathbf{R}^m)}$, where λ is chosen so that $\frac{\lambda^p}{p}C_p<\alpha$ and C_p denotes the Poincaré constant, i.e., the best constant satisfying

$$\int_\Omega |u|^p\;dx \leq C_p\int_\Omega |\nabla u|^p\;dx,$$

the estimate (11.32) yields

$$\int_\Omega |\nabla u|^p \, dx \le C,$$

where C is a positive constant depending only on Ω, p, α, and $\|g\|_{L^q(\Omega,\mathbf{R}^m)}$. Therefore u belongs to the closed ball with radius C of $W^{1,p}_{\Gamma_0}(\Omega,\mathbf{R}^m)$ which, according to the Rellich–Kondrakov theorem, Theorem 5.4.2, is compact in $L^p(\Omega,\mathbf{R}^m)$. ☐

Remark 11.2.2. In the case $p = 1$, Theorem 11.2.1 obviously remains valid as long as one considers the restrictions of F and cl(F) to $W^{1,1}(\Omega,\mathbf{R}^m)$. More precisely, for all $u \in W^{1,1}(\Omega,\mathbf{R}^m)$,

$$\mathrm{cl}(F)(u) = \int_\Omega Qf(\nabla u)\, dx.$$

Nevertheless, we do not have a complete description of cl(F), which is indeed given in the next section.

11.3 ▪ Relaxation of integral functionals with domain $W^{1,1}(\Omega,\mathbf{R}^m)$

We show how the space $BV(\Omega,\mathbf{R}^m)$ and the notion of trace (see Remark 10.2.2) take place in the relaxation theory. For simplicity of the exposition, in a first approach we limit our study to the case $m = 1$ and $f = |.|$. The general case is treated at the end of this section. From now on Ω is a Lipschitz open bounded subset of $\mathbf{R}^N$.

It is well known that the integral functionals defined on $L^p(\Omega)$, $p > 1$, by

$$F(u) = \begin{cases} \displaystyle\int_\Omega |\nabla u|^p \, dx & \text{if } u \in W^{1,p}(\Omega), \\ +\infty & \text{otherwise,} \end{cases} \qquad G(u) = \begin{cases} \displaystyle\int_\Omega |\nabla u|^p \, dx & \text{if } u \in W^{1,p}_0(\Omega), \\ +\infty & \text{otherwise,} \end{cases}$$

are lower semicontinuous for the strong topology of $L^P(\Omega)$ or the weak topology of $W^{1,p}(\Omega)$. For instance, one may argue directly by using the convexity of these two functionals or one may apply the previous section by noticing that $Qf = f$ when $f = |.|^p$.

The case $p = 1$, where $L^1(\Omega)$ is equipped with its strong topology and the functionals F, G are given by

$$F(u) = \begin{cases} \displaystyle\int_\Omega |\nabla u| \, dx & \text{if } u \in W^{1,1}(\Omega), \\ +\infty & \text{otherwise,} \end{cases} \qquad G(u) = \begin{cases} \displaystyle\int_\Omega |\nabla u| \, dx & \text{if } u \in W^{1,1}_0(\Omega), \\ +\infty & \text{otherwise,} \end{cases}$$

is more involved. Indeed, we have seen in Section 10.4 that the sequence $(u_n)_{n\in\mathbf{N}}$ which generates the Cantor–Vitali function u satisfies $u_n \to u$ strongly in $L^1(0,1)$, $\sup_{n\in\mathbf{N}} F(u_n) < +\infty$ but $u \notin W^{1,1}(\Omega)$. Consequently, the domain of the lower closure of F strictly contains the space $W^{1,1}(\Omega)$. We see below that this domain is included in the space $BV(\Omega)$. Actually, as a consequence of Proposition 11.3.2, the domain is exactly $BV(\Omega)$. Concerning the second functional G, we will see that the boundary condition $u = 0$ is "relaxed" by a surface energy.

Proposition 11.3.1. *The domain of cl(F) and cl(G) is included in $BV(\Omega)$.*

PROOF. Let $u \in L^1(\Omega,\mathbf{R}^m)$ be such that cl$(F)(u) < +\infty$ and consider a sequence $(u_n)_{n\in\mathbf{N}}$ strongly converging to u in $L^1(\Omega,\mathbf{R}^m)$ and satisfying cl$(F)(u) = \lim_{n\to+\infty} F(u_n)$. Such a sequence exists from Proposition 11.1.1. According to $\mathrm{cl}(F)(u)=\lim_{n\to+\infty}F(u_n)<+\infty$, one obtains, for a not relabeled subsequence of $(u_n)_{n\in\mathbf{N}}$,

$$\sup_{n\in\mathbf{N}}\int_\Omega |\nabla u_n|\,dx<+\infty.$$

Thus, from Proposition 10.1.1(i), $u\in BV(\Omega)$. □

Proposition 11.3.2. *Let Ω be a Lipschitz bounded open subset of $\mathbf{R}^N$ and Γ its boundary. The lsc envelopes $cl(F)$ and $cl(G)$ of the functionals F and G defined on $L^1(\Omega)$ equipped with its strong topology are given by*

$$cl(F)(u)=\begin{cases}\displaystyle\int_\Omega |Du| & \text{if } u\in BV(\Omega),\\ +\infty & \text{otherwise,}\end{cases}$$
$$cl(G)(u)=\begin{cases}\displaystyle\int_\Omega |Du|+\int_\Gamma|\gamma_0(u)|\,d\mathscr{H}^{N-1} & \text{if } u\in BV(\Omega),\\ +\infty & \text{otherwise,}\end{cases}$$

where γ_0 is the trace operator from $BV(\Omega)$ into $L^1(\Gamma)$, defined in Section 10.2.

PROOF. Let us set

$$QF(u)=\begin{cases}\displaystyle\int_\Omega |Du| & \text{if } u\in BV(\Omega),\\ +\infty & \text{otherwise,}\end{cases}$$
$$QG(u)=\begin{cases}\displaystyle\int_\Omega |Du|+\int_\Gamma|\gamma_0(u)|\,d\mathscr{H}^{N-1} & \text{if } u\in BV(\Omega),\\ +\infty & \text{otherwise.}\end{cases}$$

We first establish that for all u in $L^1(\Omega)$,

$$\begin{cases}\text{if } u_n\to u \text{ in } L^1(\Omega), \text{ then } QF(u)\le\liminf\limits_{n\to+\infty}F(u_n),\\ \text{there exists a sequence } (u_n)_{n\in\mathbf{N}} \text{ converging to } u \text{ in } L^1(\Omega) \text{ such that}\\ QF(u)\ge\limsup_{n\to+\infty}F(u_n).\end{cases}$$

These two assertions are straightforward consequences of Proposition 10.1.1 and Theorem 10.1.2. We now deal with the functional G and establish for every $u\in L^1(\Omega)$,

$$\begin{cases}\text{if } u_n\to u \text{ in } L^1(\Omega), \text{ then } QG(u)\le\liminf\limits_{n\to+\infty}G(u_n),\\ \text{there exists a sequence } (u_n)_{n\in\mathbf{N}} \text{ converging to } u \text{ in } L^1(\Omega) \text{ such that}\\ QG(u)\ge\limsup_{n\to+\infty}G(u_n).\end{cases}$$

Proof of the first assertion. Consider a sequence $(u_n)_{n\in\mathbf{N}}$ strongly converging to u in $L^1(\Omega)$ such that $\liminf_{n\to+\infty}G(u_n)<+\infty$. For a subsequence (not relabeled), we have $u_n\in W_0^{1,1}(\Omega)$. Let $\tilde\Omega$ denote a bounded open subset of $\mathbf{R}^N$ strongly containing Ω and define, for every function v in $L^1(\Omega)$, the function $\tilde v$ in $L^1(\tilde\Omega)$ by

$$\tilde v(x)=\begin{cases}v(x) & \text{if } x\in\Omega,\\ 0 & \text{if } x\in\tilde\Omega\setminus\overline{\Omega}.\end{cases}$$

It is easily seen that $\tilde{u}_n \to \tilde{u}$ in $L^1(\tilde{\Omega})$, $\tilde{u}_n \in W^{1,1}(\tilde{\Omega})$, and $\tilde{u} \in BV(\tilde{\Omega})$ with $D\tilde{u} = Du\lfloor\Omega + \gamma_0(u)\,\nu\,\mathcal{H}^{N-1}\lfloor\Gamma$, where ν denotes the inner unit normal at $\mathcal{H}^{N-1}$ a.e. x in Γ (see Examples 10.2.1 and 10.2.2).

Since $Du\lfloor\Omega$ and $\gamma_0(u)\;\nu\;\mathcal{H}^{N-1}\lfloor\Gamma$ are mutually singular, one has $|D\tilde{u}| = |Du|\lfloor\Omega + |\gamma_0(u)|\mathcal{H}^{N-1}\lfloor\Gamma$. Thus, according to the previous result for F, one has

$$\int_\Omega |Du| + \int_\Gamma |\gamma_0(u)|\, d\mathcal{H}^{N-1} = \int_{\tilde{\Omega}} |D\tilde{u}| \leq \liminf_{n\to+\infty} \int_{\tilde{\Omega}} |D\tilde{u}_n| = \liminf_{n\to+\infty} \int_\Omega |Du_n|.$$

Proof of the second assertion. For $t > 0$ let us consider the open subset $\Omega_t = \{x \in \Omega : \operatorname{dist}(x, \mathbf{R}^N \setminus \Omega) > t\}$ of Ω and define for every u in $BV(\Omega)$ the function u_t by

$$u_t(x) = \begin{cases} u(x) & \text{if } x \in \Omega_t, \\ 0 & \text{if } x \in \Omega \setminus \overline{\Omega}_t. \end{cases}$$

It is easy to see that $u_t \in BV(\Omega)$ and that

$$Du_t\lfloor\Omega = Du\lfloor\Omega_t + \gamma_t(u)\,\nu_t\;\mathcal{H}^{N-1}\lfloor_{\Gamma_t},$$

where Γ_t denotes the boundary of Ω_t, ν_t the inner unit normal at $\mathcal{H}^{N-1}$ a.e. x in Γ_t and γ_t the trace operator from $\mathrm{BV}(\Omega_t)$ into $L^1(\Gamma_t)$ (see Examples 10.2.1 and 10.2.2). According to Theorem 10.1.2 and Remark 10.2.1, there exists $u_{t,n} \in \mathbf{C}^\infty(\Omega) \cap BV(\Omega)$ satisfying $u_{t,n} = 0$ on Γ in the trace sense and converging to u_t for the intermediate convergence of $BV(\Omega)$. We obtain, for a.e. t,

$$\begin{aligned} \lim_{n\to+\infty} \int_\Omega |Du_{t,n}| &= \int_\Omega |Du_t| \\ &= \int_{\Omega_t} |Du| + \int_{\Gamma_t} |u|\, d\mathcal{H}^{N-1}, \end{aligned}$$

where we have used that for a.e. t, $\gamma_t(u) = u$ on Γ_t (cf. Example 10.2.4). Letting $t \to 0^+$, we claim that

$$\lim_{t\to 0^+} \lim_{n\to+\infty} \int_\Omega |Du_{t,n}| = \int_\Omega |Du| + \int_\Gamma |\gamma_0(u)|\, d\mathcal{H}^{N-1}. \tag{11.33}$$

In order to justify the nontrivial limit

$$\lim_{t\to 0^+} \int_{\Gamma_t} |u|\, d\mathcal{H}^{N-1} = \int_\Gamma |\gamma_0(u)|\, d\mathcal{H}^{N-1},$$

we argue with local coordinates and make use of estimate (10.13) in the proof of Theorem 10.2.1. With the notation of this proof, we have

$$\int_\Gamma |u(\tilde{x}, t) - u(\tilde{x}, t')| \leq \int_{C_{R,t',t}} |Du|,$$

and letting $t' \to 0$,

$$\int_\Gamma |\gamma_0(u) - u(\tilde{x}, t)| \leq \int_{C_{R,t}} |Du|.$$

The result follows after letting $t \to 0$.

Now t denotes a sequence converging to 0. Going back to (11.33) and using the diagonalization Lemma 11.1.1 for the sequence $\left(\int_\Omega |Du_{t,n}|, u_{t,n}\right)$ in the metrizable space $\mathbf{R}\times L^1(\Omega)$, we conclude that there exists a sequence $n\mapsto t(n)$ such that

$$\begin{cases} \lim\limits_{n\to+\infty}\displaystyle\int_\Omega |Du_{t(n),n}| = \int_\Omega |Du| + \int_\Gamma |\gamma_0(u)|\, d\mathscr{H}^{N-1}, \\ \lim\limits_{n\to+\infty} u_{t(n),n} = u \quad \text{in } L^1(\Omega), \end{cases}$$

and the proof is complete. $\square$

Let us now consider the general case and state the following theorem, analogous to Theorem 11.2.1. Given a function $f:\mathbf{M}^{m\times N}\to\mathbf{R}$ satisfying (11.5) and (11.6) with $p=1$, we intend to compute the lsc envelope of the functional $F: L^1(\Omega,\mathbf{R}^m)\longrightarrow \mathbf{R}^+\cup\{+\infty\}$ defined by

$$F(u)=\begin{cases}\displaystyle\int_\Omega f(\nabla u)\,dx & \text{if } u\in W^{1,1}(\Omega,\mathbf{R}^m),\\ +\infty & \text{otherwise.}\end{cases} \tag{11.34}$$

Let us recall that when $u\in BV(\Omega)$, according to the Radon–Nikodým theorem, Theorem 4.2.1, one has $Du=\nabla u\mathscr{L}^N\lfloor\Omega + D^s u$, where $\nabla u\mathscr{L}^N\lfloor\Omega$ and $D^s u$ are two mutually singular measures in $\mathbf{M}(\Omega,\mathbf{R}^N)$.

Theorem 11.3.1. *The lsc envelope of the functional F defined in* (11.34) *is given, for every $u\in L^1(\Omega,\mathbf{R}^m)$, by*

$$cl(F)(u)=\begin{cases}\displaystyle\int_\Omega Qf(\nabla u)\,dx+\int_\Omega (Qf)^\infty\left(\frac{D^s u}{|D^s u|}\right)|D^s u| & \textit{if } u\in BV(\Omega,\mathbf{R}^m),\\ +\infty & \textit{otherwise,}\end{cases}$$

where Qf is the quasi-convex envelope of f defined in Proposition 11.2.2, *and $(Qf)^\infty$ is the recession function of Qf defined for every a in $\mathbf{M}^{m\times N}$ by $(Qf)^\infty(a)=\limsup_{t\to+\infty}\frac{Qf(ta)}{t}$.*

The proof of Theorem 11.3.1 is a consequence of Propositions 11.3.3 and 11.3.4. We set

$$QF(u)=\begin{cases}\displaystyle\int_\Omega Qf(\nabla u)\,dx+\int_\Omega (Qf)^\infty\left(\frac{D^s u}{|D^s u|}\right)|D^s u| & \text{if } u\in \mathrm{BV}(\Omega,\mathbf{R}^m)\\ +\infty & \text{otherwise.}\end{cases}$$

Proposition 11.3.3. *For every u in $L^1(\Omega,\mathbf{R}^m)$ and every sequence $(u_n)_{n\in\mathbf{N}}$ strongly converging to u in $L^1(\Omega,\mathbf{R}^m)$, one has*

$$QF(u)\le \liminf_{n\to+\infty} F(u_n). \tag{11.35}$$

PROOF. Our strategy is exactly the one of Proposition 11.2.3. Obviously, one may assume $\liminf_{n\to+\infty}F(u_n)<+\infty$. For a nonrelabeled subsequence, let us consider the nonnegative Borel measure $\mu_n := f(\nabla u_n(.))\mathscr{L}^N\lfloor\Omega$. Since

$$\sup_{n\in\mathbf{N}}\mu_n(\Omega)<+\infty,$$

there exists a further subsequence (not relabeled) and a nonnegative Borel measure $\mu \in \mathbf{M}(\Omega)$ such that

$$\mu_n \rightharpoonup \mu \text{ weakly in } \mathbf{M}(\Omega).$$

Let $\mu = g\mathscr{L}^N\lfloor\Omega + \mu^s$ be the Lebesgue–Nikodým decomposition of μ, where μ^s is a nonnegative Borel measure, singular with respect to the N-dimensional Lebesgue measure $\mathscr{L}^N\lfloor\Omega$. For establishing (11.35) it suffices to prove that

$$\begin{aligned} &g(x) \geq Qf(\nabla u(x)) \qquad \text{for a.e. } x \in \Omega,\\ &\mu^s \geq (Qf)^\infty\left(\frac{D^s u}{|D^s u|}\right)|D^s u|. \end{aligned}$$

Indeed, according to Alexandrov's theorem, Theorem 4.2.3, we will obtain

$$\begin{aligned} \liminf_{n\to+\infty} F(u_n) = \liminf_{n\to+\infty} \mu_n(\Omega) \geq \mu(\Omega) &= \int_\Omega g(x)\,dx + \mu_s(\Omega)\\ &\geq \int_\Omega Qf(\nabla u(x))\,dx + \int_\Omega (Qf)^\infty\left(\frac{D^s u}{|D^s u|}\right)|D^s u|. \end{aligned}$$

(a) *Proof of* $g(x) \geq Qf(\nabla u(x))$ *for a.e.* x *in* Ω. It suffices to reproduce the proof of Proposition 11.2.3, which obviously holds true for $p = 1$ (see Remark 11.2.2).

(b) *Proof of* $\mu^s \geq (Qf)^\infty(\frac{D^s u}{|D^s u|})\,|D^s u|$. The proof is based on various arguments of Ambrosio and Dal Maso [25]. The density $\frac{Du}{|Du|}$ satisfies the following property (see Alberti [10]).

Lemma 11.3.1. *The density* $\frac{Du}{|Du|}$ *is for* $|D^s u|$*-a.e.* $x_0 \in \Omega$ *a rank-one matrix, i.e., for* $|D^s u|$*-a.e.* $x_0 \in \Omega$

$$\frac{Du}{|Du|}(x_0) = a(x_0) \otimes b(x_0)$$

with $a(x_0) \in \mathbf{R}^m$, $b(x_0) \in \mathbf{R}^N$, $|a(x_0)| = |b(x_0)| = 1$.

The rank-one property of the jump part of the singular measure $D^s u$ is indeed trivial because of its structure (see Section 10.3). De Giorgi conjectured that the diffuse singular part is also a rank-one matrix valued measure. The proof was later given by Alberti in [10], to which we refer. Let us give now some notation. Let x_0 be an element of the complementary of the $|D^s u|$-null set invoked in Lemma 11.3.1 and let Q denote the unit cube of $\mathbf{R}^N$ centered at the origin, whose sides are either orthogonal or parallel to $b(x_0)$. We set $Q_\rho(x_0) := \{x_0 + \rho x \, : \, x \in Q\}$, where ρ is a positive parameter intended to tend to zero.

According to the theory of differentiation of measures (see Theorem 4.2.1), for establishing

$$\mu^s \geq (Qf)^\infty(D^s u),$$

it is enough to prove

$$\lim_{\rho\to 0} \frac{\mu(Q_\rho(x_0))}{|Du|(Q_\rho(x_0))} \geq (Qf)^\infty(a(x_0) \otimes b(x_0)). \tag{11.36}$$

We will make use of the following two estimates: for $|D^s u|$-a.e. $x_0 \in \Omega$, we have

$$\lim_{\rho\to 0} \frac{|Du|(Q_\rho(x_0))}{\rho^N} = +\infty, \tag{11.37}$$

$$\limsup_{\rho\to 0} \frac{|Du|(Q_{t\rho}(x_0))}{|Du|(Q_\rho(x_0))} \geq t^N \ \forall t \in]0,1[^d. \tag{11.38}$$

Assertion (11.37) easily follows from the theory of differentiation of measures (see Theorem 4.2.1). For a proof of (11.38), consult Ambrosio and Dal Maso [25, Theorem 2.3].

From now on, x_0 will be a fixed element of the complementary of the $|D^s u|$-null set invoked in Lemma 11.3.1, for which moreover estimates (11.37) and (11.38) hold true and the two limits

$$\lim_{\rho\to 0} \frac{\mu(Q_\rho(x_0))}{|Du|(Q_\rho(x_0))}, \quad \lim_{\rho\to 0} \frac{Du(Q_\rho(x_0))}{|Du|(Q_\rho(x_0))}$$

exist (cf. Theorem 4.2.1).

We set $t_\rho := \frac{|Du|(Q_\rho(x_0))}{\rho^N}$. According to (11.37), t_ρ will play the role of the parameter t in the definition of $(Qf)^\infty$:

$$(Qf)^\infty(a) = \limsup_{t\to+\infty} \frac{Qf(ta)}{t}.$$

Let us now define the following rescaled functions of $BV(Q,\mathbf{R}^m)$:

$$v_\rho(x) := \frac{\rho^{N-1}}{|Du|(Q_\rho(x_0))}\left(u(x_0+\rho x) - \frac{1}{|Q_\rho(x_0)|}\int_{Q_\rho(x_0)} u(y)dy \right),$$

$$v_{\rho,n}(x) := \frac{\rho^{N-1}}{|Du|(Q_\rho(x_0))}\left(u_n(x_0+\rho x) - \frac{1}{|Q_\rho(x_0)|}\int_{Q_\rho(x_0)} u_n(y)dy \right),$$

which satisfy

$$\int_Q v_\rho(y)dy = 0, \ \lim_{n\to+\infty} |v_{\rho,n} - v_\rho|_{L^1(Q,\mathbf{R}^m)} = 0$$

and, from Lemma 11.3.1,

$$Dv_\rho(Q) = \frac{Du(Q_\rho(x_0))}{|Du|(Q_\rho(x_0))} \to a(x_0)\otimes b(x_0) \text{ when } \rho \to 0.$$

The sequence $(v_\rho)_{\rho>0}$ fulfills the following properties (for a proof, consults Ambrosio and Dal Maso [25, Theorem 2.3]).

Lemma 11.3.2. *There exists a subsequence of $(v_\rho)_{\rho>0}$, not relabeled, which weakly converges in $BV(Q,\mathbf{R}^m)$ to a function v of the form $v(x) = \overline{v}(\langle b(x_0), x\rangle)a(x_0)$, where $\overline{v}$ is nondecreasing and belongs to $BV(]-1/2,1/2[)$. Moreover, for a.e. δ in $(0,1)$ one has*

$$Dv_\rho(\delta Q) \to Dv(\delta Q).$$

We are now in a position to establish $\mu^s \geq (Qf)^\infty(\frac{D^s u}{|D^s u|})\,|D^s u|$. The proof will be divided into three steps. We fix δ in $(0,1)$ outside of a set of null measure so that the last assertion of Lemma 11.3.2 holds. We will make δ tend to 1 at the end of the proof.

First step (truncation). Let $u_{\rho,n}$ in $W^{1,1}(Q,\mathbf{R}^m)$ defined by

$$u_{\rho,n} := \frac{1}{\rho}\left(u_n(x_0+\rho y) - \frac{1}{|Q_\rho(x_0)|}\int_{Q_\rho(x_0)} u_n(y)dy\right).$$

Note that $\frac{1}{t_\rho}u_{\rho,n}$ is the function $v_{\rho,n}$ previously defined. A change of scale gives

$$\frac{1}{|Du|(Q_\rho(x_0))}\int_{Q_{\delta\rho}(x_0)} f(\nabla u_n)dx = \frac{1}{t_\rho}\int_{\delta Q} f(\nabla u_{\rho,n})dx.$$

We want to modify the function $u_{\rho,n}$ in a neighborhood of δQ so that it coincides with an affine function of gradient $t_\rho Dv(\delta Q)$ on $\partial\delta Q$. The basic idea, similar to the one used in the proof of Proposition 11.2.3, consists in slicing a neighborhood of the boundary of δQ by thin slices whose size is of order $||v_\rho - v||^{1/2}_{L^1(Q,\mathbf{R}^m)}$. (Recall that $||v_\rho - v||_{L^1(Q,\mathbf{R}^m)}$ goes to zero when $\rho \to 0$.)

More precisely, we set $\alpha_\rho := ||v_\rho - v||^{\frac{1}{2}}_{L^1(Q,\mathbf{R}^m)}$ and, for $\nu \in \mathbf{N}^*$ intended to tend to $+\infty$,

$$Q_0 := (1-\alpha_\rho)\delta Q,\; Q_i := \left(1-\alpha_\rho + i\frac{\alpha_\rho}{\nu}\right)\delta Q \quad \text{for } i = 1,\dots,\nu.$$

On the other hand, for $i = 1,\dots,\nu$, we consider

$$\varphi_i \in \mathbf{C}_0^\infty(Q_i),\ 0 \leq \varphi_i \leq 1,\ \varphi_i = 1 \text{ in } Q_{i-1}, \quad |\nabla\varphi_i| \leq \frac{\nu}{\alpha_\rho}$$

and define $u_{\rho,n,i} \in t_\rho\Theta_\delta + W_0^{1,1}(\delta Q,\mathbf{R}^m)$ by

$$u_{\rho,n,i} := t_\rho\Theta_\delta + \varphi_i(u_{\rho,n} - t_\rho\Theta_\delta),$$

where Θ_δ is the affine function:

$$\Theta_\delta(x) := \frac{Dv(\delta Q)}{\delta^N}.x + \frac{\overline{v}((\frac{\delta}{2})^-) + \overline{v}((-\frac{\delta}{2})^+)}{2}a(x_0).$$

The constant $\frac{\overline{v}((\delta/2)^-)+\overline{v}(-(\delta/2)^+)}{2}a(x_0)$ has been chosen so that the traces of v and Θ agree on the faces of δQ orthogonal to $b(x_0)$ and, consequently, so that $v - \theta$ fulfills the Poincaré inequality on δQ. From (11.5), an easy computation gives

$$\frac{1}{t_\rho}\int_{\delta Q} f(\nabla u_{\rho,n})dx \geq \frac{1}{t_\rho}\int_{\delta Q} f(\nabla u_{\rho,n,i})dx - R_{\rho,n,\nu,\delta,i}, \tag{11.39}$$

where

$$R_{\rho,n,\nu,\delta,i} := O_\rho + \frac{\beta}{t_\rho}\int_{Q_i\setminus Q_{i-1}} |\nabla(u_{\rho,n} - t_\rho\Theta_\delta)|dx + \frac{\beta\nu}{\alpha_\rho}\int_{Q_i\setminus Q_{i-1}} |v_{\rho,n} - \Theta_\delta|dx$$

and O_ρ does not depend on n and tends to 0 when $\rho \to +\infty$.

Second step (averaging). According to (11.39) and to the definition of Qf, we have

$$\frac{1}{|Du|(Q_\rho(x_0))}\int_{Q_{\delta\rho}(x_0)} f(\nabla u_n)dx \geq \frac{\delta^N}{t_\rho}Qf\left(\frac{t_\rho}{\delta^N}Dv(\delta Q)\right) - O_\rho - \frac{\beta}{t_\rho}\int_{Q_i\setminus Q_{i-1}}|\nabla(u_{\rho,n}-t_\rho\Theta_\delta)|dx - \frac{\beta\nu}{\alpha_\rho}\int_{Q_i\setminus Q_{i-1}}|v_{\rho,n}-\Theta_\delta|dx.$$

Averaging these ν inequalities, we obtain

$$\begin{aligned}\frac{1}{|Du|(Q_\rho(x_0))}\int_{Q_{\delta\rho}(x_0)} f(\nabla u_n)dx &\geq \frac{\delta^N}{t_\rho}Qf\left(\frac{t_\rho}{\delta^N}Dv(\delta Q)\right)\\ - O_\rho - \frac{\beta}{\nu t_\rho}\int_Q|\nabla(u_{\rho,n}-t_\rho\Theta_\delta)|dx &- \frac{\beta}{\alpha_\rho}\int_{\delta Q\setminus(1-\alpha_\rho)\delta Q}|v_{\rho,n}-\Theta_\delta|dx\\ \geq \frac{\delta^N}{t_\rho}Qf\left(\frac{t_\rho}{\delta^N}Dv(\delta Q)\right) - O_\rho &- \frac{\beta}{\nu t_\rho}\int_Q|D(u_{\rho,n}-t_\rho\Theta_\delta)|dx\\ -\beta\left(\int_Q|v_{\rho,n}-v|dx\right)^{\frac{1}{2}} &- \frac{\beta}{\alpha_\rho}\int_{\delta Q\setminus(1-\alpha_\rho)\delta Q}|v-\Theta_\delta|dx.\end{aligned}$$

Letting successively $\nu\to+\infty$ and $n\to+\infty$, we obtain

$$\begin{aligned}&\limsup_{n\to+\infty}\frac{1}{|Du|(Q_\rho(x_0))}\int_{Q_{\delta\rho}(x_0)} f(\nabla u_n)dx\\ &\geq \frac{\delta^N}{t_\rho}Qf\left(\frac{t_\rho}{\delta^N}Dv(\delta Q)\right) - O_\rho - \beta\left(\int_Q|v_m-v|dx\right)^{\frac{1}{2}}\\ &-\frac{\beta}{\alpha_\rho}\int_{\delta Q\setminus(1-\alpha_\rho)\delta Q}|v-\Theta_\delta|dx.\end{aligned}\tag{11.40}$$

Last step. From Lipschitz continuity of Qf (cf. Proposition 11.2.2) and according to Lemmas 11.3.1 and 11.3.2 and estimates (11.37) and (11.38), we obtain

$$\begin{aligned}\limsup_{\rho\to0}\frac{\delta^N}{t\rho}Qf\left(\frac{t_\rho}{\delta^N}Dv(\delta Q)\right) &= \limsup_{\rho\to0}\frac{\delta^N}{t_\rho}Qf\left(\frac{t_\rho}{\delta^N}Dv(\delta Q)\right)\\ &\geq \limsup_{\rho\to0}\frac{\delta^N}{t_\rho}Qf\left(\frac{t_\rho}{\delta^N}a(x_0)\otimes b(x_0)\right)\\ &\quad -\liminf_{\rho\to0}L'\left(1-\frac{|Du|(C_{\delta\rho}(x_0))}{|Du|(C_\rho(x_0))}\right)\\ &\geq (Qf)^\infty(a(x_0)\otimes b(x_0)) - L'(1-\delta^N).\end{aligned}\tag{11.41}$$

On the other hand,

$$\lim_{\rho\to0}\frac{1}{\alpha_\rho}\int_{\delta Q\setminus(1-\alpha_\rho)\delta Q}|v-\Theta_\delta|dx = 0.\tag{11.42}$$

Indeed, from Poincaré's inequality

$$\frac{1}{\alpha_\rho}\int_{\delta Q\setminus(1-\alpha_\rho)\delta Q}|v-\Theta_\delta|dx \le C\int_{\delta Q\setminus(1-\alpha_\rho)\delta Q}\left|Dv-\frac{Dv(\delta Q)}{\delta^N}\right|$$

which tends to zero when ρ goes to zero. Combining (11.40), (11.41), and (11.42), we obtain

$$\limsup_{\rho\to 0}\limsup_{n\to+\infty}\frac{1}{|Du|(Q_\rho(x_0))}\int_{\delta Q_\rho(x_0)} f(\nabla u_n)dx \ge (Qf)^\infty(a(x_0)\otimes b(x_0)) - L'(1-\delta^N).$$

The lower bound is then established after letting $\delta \to 1$. $\square$

Proposition 11.3.4. *For every u in $L^p(\Omega,\mathbf{R}^m)$, there exists a sequence $(u_n)_{n\in\mathbf{N}}$ strongly converging to u in $L^1(\Omega,\mathbf{R}^m)$ such that*

$$QF(u) \ge \limsup_{n\to+\infty} F(u_n).$$

PROOF. The proof proceeds in two steps.

First step: computation of $cl(F)\lfloor W^{1,1}(\Omega,\mathbf{R}^m)$. According to Remark 11.2.2 and Theorem 11.2.1 for every $u \in W^{1,1}(\Omega,\mathbf{R}^m)$,

$$\mathrm{cl}(F)(u) = \int_\Omega Qf(\nabla u)\,dx. \tag{11.43}$$

Second step. Let us consider the functional $\tilde{F} : L^1(\Omega,\mathbf{R}^m) \longrightarrow \mathbf{R}^+ \cup \{+\infty\}$ defined by

$$\tilde{F}(u) = \begin{cases} \displaystyle\int_\Omega Qf(\nabla u)\,dx & \text{if } u \in W^{1,1}(\Omega,\mathbf{R}^m),\\ +\infty & \text{otherwise.}\end{cases}$$

We claim that it is enough to establish that for all $u \in \mathrm{BV}(\Omega,\mathbf{R}^m)$,

$$\mathrm{cl}(\tilde{F})(u) \le \int_\Omega Qf(\nabla u)\,dx + \int_\Omega (Qf)^\infty\left(\frac{D^s u}{|D^s u|}\right)|D^s u|. \tag{11.44}$$

Indeed, let us assume (11.44). For every $u \in \mathrm{BV}(\Omega,\mathbf{R}^m)$, according to Proposition 11.1.1 and Theorem 11.1.1, we obtain the existence of a sequence $(u_k)_{k\in\mathbf{N}}$ in $W^{1,1}(\Omega,\mathbf{R}^m)$ strongly converging to u in $L^1(\Omega,\mathbf{R}^m)$ and satisfying

$$\begin{aligned}\int_\Omega Qf(\nabla u)\,dx + \int_\Omega (Qf)^\infty\left(\frac{D^s u}{|D^s u|}\right)|D^s u| &\ge \mathrm{cl}(\tilde{F})(u)\\ &\ge \limsup_{k\to+\infty}\tilde{F}(u_k)\\ &= \limsup_{k\to+\infty}\int_\Omega Qf(\nabla u_k)\,dx.\end{aligned}$$

On the other hand, from (11.43), there exists a sequence $(u_{k,n})_{n\in\mathbf{N}}$ in $W^{1,1}(\Omega,\mathbf{R}^m)$ strongly converging to u_k in $L^1(\Omega,\mathbf{R}^m)$ and satisfying

$$\int_\Omega Qf(\nabla u_k)\,dx \ge \limsup_{n\to+\infty}\int_\Omega f(\nabla u_{k,n})\,dx.$$

Consequently

$$\begin{cases} \displaystyle\int_\Omega Qf(\nabla u)\,dx + \int_\Omega (Qf)^\infty\Big(\frac{D^s u}{|D^s u|}\Big)\,|D^s u| \geq \limsup_{k\to+\infty}\limsup_{n\to+\infty}\int_\Omega f(\nabla u_{k,n})\,dx, \\ \displaystyle\lim_{k\to+\infty}\lim_{n\to+\infty} u_{k,n} = u \quad \text{strongly in } L^1(\Omega,\mathbf{R}^m), \end{cases}$$

and we conclude by a diagonalization argument: there exists $n\mapsto k(n)$ mapping $\mathbf{N}$ to $\mathbf{N}$ such that

$$\begin{cases} \displaystyle\limsup_{n\to+\infty}\int_\Omega f(\nabla u_{k(n),n})\,dx \leq \int_\Omega Qf(\nabla u)\,dx + \int_\Omega (Qf)^\infty\Big(\frac{D^s u}{|D^s u|}\Big)\,|D^s u|, \\ \displaystyle\lim_{n\to+\infty} u_{k(n),n} = u \quad \text{strongly in } L^1(\Omega,\mathbf{R}^m). \end{cases}$$

We now establish (11.44). We use some arguments of Ambrosio and Dal Maso [25]. Let u be a fixed element of $\mathrm{BV}(\Omega,\mathbf{R}^m)$ and consider the set function

$$\mu : A \mapsto \mathrm{cl}(\tilde F)(u,A)$$

defined for all bounded open subset A of Ω. The notation $\mathrm{cl}(\tilde F)(\cdot,A)$ means that we consider the lower semicontinuous envelope of the functional $\tilde F$ localized on A, the space $L^1(A,\mathbf{R}^m)$ being equipped with its strong topology. Note that μ satisfies the following estimate for every open bounded subset A of Ω:

$$\mu(A) \leq \beta|A| + \beta\int_A |Du|. \tag{11.45}$$

Indeed, applying the approximation Theorem 10.1.2, there exists a sequence $(v_n)_{n\in\mathbf{N}}$ in $\mathrm{BV}(A,\mathbf{R}^m)\cap\mathbf{C}^\infty(A,\mathbf{R}^m)$ such that

$$v_n \to u \quad \text{strongly in } L^1(A,\mathbf{R}^m),$$

$$\int_A |\nabla v_n|\,dx \to \int_A |Du|.$$

Thus, from the growth condition (11.5)

$$\begin{aligned} \mu(A) = \mathrm{cl}(\tilde F)(u,A) &\leq \liminf_{n\to+\infty} F(v_n,A) \\ &\leq \beta|A| + \liminf_{n\to+\infty}\int_A |\nabla v_n|\,dx \\ &= \beta|A| + \int_A |Du|. \end{aligned}$$

According to the definition of μ and to (11.45), one can now easily establish that for all bounded open subsets A and A' of Ω, μ satisfies

$$A\subset A' \Longrightarrow \mu(A)\leq\mu(A');$$

$$A\cap A' = \emptyset \Longrightarrow \mu(A\cup A') \geq \mu(A)+\mu(A');$$

$$\mu(A) \leq \sup\{\mu(A') \,:\, A'\subset\subset A\};$$

$$\mu(A\cup A') \leq \mu(A)+\mu(A').$$

For a complete proof, consult Ambrosio and Dal Maso [25]. Consequently (consult, for instance, the book [183]), one can extend μ to a Borel measure on Ω, still denoted by μ, defined for every Borel set B of Ω by

$$\mu(B) = \inf\{\mu(A) \,:\, A \text{ open}, B \subset A\}.$$

By considering the Lebesgue–Nikodým decomposition (cf. Theorem 4.2.1) $\mu = \mu^a + \mu^s$ of μ, it suffices now to establish, for every Borel set B of Ω,

$$\mu^a(B) \leq \int_B Qf(\nabla u)\, dx, \tag{11.46}$$

$$\mu^s(B) \leq \int_B (Qf)^\infty\left(\frac{D^s u}{|D^s u|}\right) |D^s u|. \tag{11.47}$$

In order to estimate the singular part μ^s from above, i.e., (11.47), we will need the following continuity result.

Lemma 11.3.3. *Let $h : \mathbf{M}^{m\times N} \to \mathbf{R}$ be a continuous function and $(u_n)_{n\in\mathbf{N}}$ a sequence in $BV(\Omega, \mathbf{R}^m)$ converging to u for the intermediate convergence. Then*

$$\lim_{n\to+\infty} \int_\Omega h\left(\frac{Du_n}{|Du_n|}\right) |Du_n| = \int_\Omega h\left(\frac{Du}{|Du|}\right) |Du|.$$

For a proof, we refer the reader to Luckhaus and Modica [281] or, in the convex case, to Demengel and Temam [199]. Let us set, for every $a \in \mathbf{M}^{m\times N}$,

$$g(a) = \sup_{t>0} \frac{Qf(ta) - Qf(0)}{t}.$$

According to the rank-one convexity of Qf (cf. Remark 11.2.1), we note that if $\operatorname{rank}(a) = 1$, one has $g(a) = (Qf)^\infty(a)$. For every open subset A of Ω, consider $u_n \in \mathrm{BV}(A, \mathbf{R}^m) \cap \mathbf{C}^\infty(A, \mathbf{R}^m)$, strongly converging to u in $L^1(A, \mathbf{R}^m)$, and such that

$$\lim_{n\to+\infty} |Du_n|(A) = |Du|(A).$$

Such a sequence exists by Theorem 10.1.2. From $f \leq f(0) + g$ and Lemma 11.3.3, one has

$$\begin{aligned}\mu(A) &\leq Qf(0)|A| + \lim_{n\to+\infty} \int_A g\left(\frac{Du_n}{|Du_n|}\right)|Du_n| \\ &= Qf(0)|A| + \int_A g\left(\frac{Du}{|Du|}\right)|Du|.\end{aligned}$$

The same inequality now holds true for any Borel set B of Ω. Taking the singular part of the two members and according to the rank-one property of the singular part of Du, we obtain

$$\mu^s(B) \leq \int_B (Qf)^\infty\left(\frac{D^s u}{|D^s u|}\right)|D^s u|.$$

We are going to estimate from above the absolutely continuous part μ^a, i.e., (11.46). Let ρ_n be a regularizing kernel (see Theorem 4.2.2) and set $u_n = \rho_n * u$. From

$$\nabla u_n = \rho_n * Du = \rho_n * \nabla u + \rho_n * D^s u$$

and the local Lipschitz continuity of Qf, one has, for every open set $A' \subset\subset A$,

$$\begin{aligned}\int_{A'} Qf(\nabla u_n)\, dx &\le \int_{A'} Qf(\rho_n * \nabla u)\, dx + L'|\rho_n * D^s u|(A') \\ &\le \int_{A'} Qf(\rho_n * \nabla u)\, dx + L'|D^s u|(A' + \mathrm{spt}(\rho_n)).\end{aligned}$$

Letting $n \to +\infty$ yields

$$\mu(A') \le \int_{A'} Qf(\nabla u)\, dx + L'|D^s u|(A)$$

and, since $A' \subset\subset A$ is arbitrary,

$$\mu(A) \le \int_{A} Qf(\nabla u)\, dx + L'|D^s u|(A)$$

for every open subset A of Ω. Since the Lebesgue measure and the nonnegative Borel measure $|D^s u|$ are regular, the same inequality now holds true for every Borel subset B of Ω. Taking the absolutely continuous part of each member, we finally obtain

$$\mu^a(B) \le \int_B Qf(\nabla u)\, dx$$

and the proof is complete. $\square$

We now compute the lsc envelope of the integral functional in (11.34) by taking into account a boundary condition on part Γ_0 of the boundary $\partial\Omega$ of Ω. More precisely, we aim to describe the lsc envelope of the integral functional $F : L^1(\Omega, \mathbf{R}^m) \longrightarrow \mathbf{R}^+ \cup \{+\infty\}$ defined by

$$F(u) = \begin{cases} \displaystyle\int_\Omega f(\nabla u)\, dx & \text{if } u \in W^{1,1}_{\Gamma_0}(\Omega, \mathbf{R}^m), \\ +\infty & \text{otherwise.}\end{cases}$$

Corollary 11.3.1. *The lsc envelope of the integral functional F is given by*

$$cl(F)(u) = \begin{cases} \displaystyle\int_\Omega Qf(\nabla u)\, dx + \int_\Omega (Qf)^\infty\left(\frac{D^s u}{|D^s u|}\right)|D^s u| \\ \qquad\qquad + \displaystyle\int_{\Gamma_0} (Qf)^\infty(\gamma_0(u)\otimes \nu)\, d\mathcal{H}^{N-1} & \text{if } u \in BV(\Omega, \mathbf{R}^m), \\ +\infty \quad \text{otherwise,} \end{cases}$$

where ν denotes the outer unit normal to Γ_0 and γ_0 the trace operator.

In other words, in this relaxation process, the boundary condition is translated in terms of surface energy $\int_\Omega (Qf)^\infty([u]\otimes\nu)\, d\mathcal{H}^{N-1}\lfloor\Gamma_0$. The proof is very similar to that of Proposition 11.3.2. For details, consult Abddaimi, Licht, and Michaille [1] and the references therein. As a consequence we obtain the relaxation theorem in the case $p = 1$.

Theorem 11.3.2 (relaxation theorem, $p = 1$). *Let us consider a function $f : \mathbf{M}^{m\times N} \longrightarrow \mathbf{R}$ satisfying* (11.5) *and* (11.6) *and g in $L^\infty(\Omega, \mathbf{R}^m)$ satisfying $\|g\|_\infty < \frac{\alpha}{C_p}$, where C_p is the*

Poincaré constant in Ω. Then the relaxed problem of

$$\inf\left\{\int_\Omega f(\nabla u)\,dx - \int_\Omega g.u\,dx \;:\; u \in W^{1,1}_{\Gamma_0}(\Omega,\mathbf{R}^m)\right\}, \qquad (\mathscr{P})$$

in the sense of Theorem 11.1.2, *is given by*

$$\inf\left\{\int_\Omega Qf(\nabla u)\,dx - \int_\Omega (Qf)^\infty\left(\frac{D^s u}{|D^s u|}\right)|D^s u| + \int_{\Gamma_0}(Qf)^\infty(\gamma_0(u)\otimes\nu)\,d\mathscr{H}^{N-1} \right. \qquad (\overline{\mathscr{P}})$$
$$\left. - \int_\Omega g.u\,dx \;:\; u \in BV(\Omega,\mathbf{R}^m)\right\}.$$

PROOF. Arguing as in the proof of Theorem 11.2.2, according to Theorem 11.1.2, it suffices to establish the inf-compactness of

$$u \mapsto F(u) - \int_\Omega g.u\,dx$$

in $L^1(\Omega,\mathbf{M}^{m\times N})$. This property is a straightforward consequence of the estimate $\|g\|_\infty < \frac{\alpha}{C_p}$ and the compactness of the embedding of $W^{1,1}_{\Gamma_0}(\Omega,\mathbf{R}^m)$ into $L^1(\Omega,\mathbf{M}^{m\times N})$. □

11.4 ▪ Relaxation in the space of Young measures in nonlinear elasticity

The strategy described in Sections 11.2 and 11.3 has the disadvantage to quasi-convexify the density function f so that the relaxed functional, with density Qf, does not provide information on the oscillations of the gradient minimizing sequences. In this section, we describe an alternative way for relaxing the free energy by using the notion of the Young measure introduced in Section 4.3, well adapted for capturing oscillations of minimizing sequences (see Subsection 4.3.6). According to the point of view of Ball and James in [81] or Bhattacharya and Kohn in [99], the density of the relaxed free energy obtained this way is the microscopic free energy density corresponding to the macroscopic free energy density Qf.

11.4.1 ▪ Young measures generated by gradients

To shorten notation, we denote the N-dimensional Lebesgue measure restricted to the open bounded subset $\Omega \subset \mathbf{R}^N$ by $\mathscr{L}$.

Definition 11.4.1. *Let us denote by $E := \mathbf{R}^{mN} \sim \mathbf{M}^{m\times N}$ the set of $m \times N$ matrices. A Young measure μ in $\mathscr{Y}(\Omega;E)$ is called a $W^{1,p}$-Young measure if there exists a bounded sequence $(u_n)_{n\in\mathbf{N}}$ in $W^{1,p}(\Omega,\mathbf{R}^m)$, $p \geq 1$, such that μ is generated by the sequence of gradients $(\nabla u_n)_{n\in\mathbf{N}}$, i.e.,*

$$(\delta_{\nabla u_n(x)})_{x\in\Omega}\otimes\mathscr{L} \overset{nar}{\rightharpoonup} \mu,$$

or equivalently, from Theorem 4.3.1,

$$(\delta_{\nabla u_n(x)})_{x\in\Omega} \overset{L_w}{\rightharpoonup} (\mu_x)_{x\in\Omega}.$$

We admit the following important technical lemma. For a proof, consult Kinderlehrer and Pedregal [257], Pedregal [316], or Fonseca, Müller, and Pedregal [218].

Lemma 11.4.1. *Let $(u_n)_{n\in\mathbf{N}}$ be a sequence in $W^{1,p}(\Omega,\mathbf{R}^m)$, weakly converging to some u in $W^{1,p}(\Omega,\mathbf{R}^m)$. Then, there exists a sequence $(v_n)_{n\in\mathbf{N}}$ in $W^{1,p}(\Omega,\mathbf{R}^m)$ satisfying*

(i) $v_n \in u + W_0^{1,p}(\Omega,\mathbf{R}^m)$;

(ii) $(|\nabla v_n|^p)_{n\in\mathbf{N}}$ *is uniformly integrable;*

(iii) $v_n - u_n \to 0$ *and* $\nabla(v_n - u_n) \to 0$ *in measure.*

Let us point out that according to item (iii) and to Proposition 4.3.8, the two sequences $(\nabla u_n)_{n\in\mathbf{N}}$ and $(\nabla v_n)_{n\in\mathbf{N}}$ generate the same $W^{1,p}$-Young measure.

The main result of this section is the following characterization of $W^{1,p}$-Young measures, established by Kinderlehrer and Pedregal [257] (see Pedregal [316] and Sychev [344]).

Theorem 11.4.1. *Let $p > 1$; then $\mu \in \mathscr{Y}(\Omega;E)$ is a $W^{1,p}$-Young measure iff there exists $u \in W^{1,p}(\Omega,\mathbf{R}^m)$ such that the three following assertions hold:*

(i) $\nabla u(x) = \int_E \lambda \, d\mu_x(\lambda)$ *for a.e. x in Ω;*

(ii) *for all quasi-convex function $\phi : E \to \mathbf{R}$ for which there exist some $\gamma \in \mathbf{R}$ and $\beta > 0$ such that $\gamma \le \phi(\lambda) \le \beta(1+|\lambda|^p)$ for all $\lambda \in E$, one has*

$$\phi(\nabla u(x)) \le \int_E \phi(\lambda) \, d\mu_x(\lambda) \quad \textit{for a.e. } x \textit{ in } \Omega;$$

(iii) $\int_{\Omega\times E} |\lambda|^p \, d\mu(x,\lambda) < +\infty$.

The function u will be referred to as the underlying deformation of the Young measure μ.

PROOF. We split the proof into several steps.

PROOF OF THE NECESSARY CONDITIONS.

First step: Necessity of (i) *and* (iii). By definition, there exists a bounded sequence $(u_n)_{n\in\mathbf{N}}$ in $W^{1,p}(\Omega,\mathbf{R}^m)$ weakly converging to some $u \in W^{1,p}(\Omega,\mathbf{R}^m)$, whose sequence of gradients generates μ. Since $\nabla u_n \rightharpoonup \nabla u$ weakly in $L^p(\Omega,E)$, one obtains (i) by applying Proposition 4.3.6. Take now $\varphi(x,\lambda) = |\lambda|^p$. Since $\lambda \mapsto \varphi(x,\lambda)$ is lsc (actually continuous), according to Proposition 4.3.3, we deduce

$$\int_{\Omega\times E} |\lambda|^p \, d\mu(x,\lambda) \le \liminf_{n\to+\infty} \int_\Omega |\nabla u_n|^p \, dx < +\infty.$$

Second step: Necessity of (ii). Let ϕ be a quasi-convex function satisfying the growth condition in (ii) and x_0 a fixed element in Ω such that the two following limits exist:

$$\phi(\nabla u(x_0)) = \lim_{\rho\to 0} \frac{1}{\mathscr{L}(B_\rho(x_0))} \int_{B_\rho(x_0)} \phi(\nabla u(x)) \, dx,$$

$$\int_E \phi(\lambda) \, d\mu_{x_0}(\lambda) = \lim_{\rho\to 0} \frac{1}{\mathscr{L}(B_\rho(x_0))} \int_{B_\rho(x_0)} \Big(\int_E \phi(\lambda) \, d\mu_x(\lambda) \Big) \, dx.$$

Such x_0 exists outside a negligible set, from the Lebesgue differentiation theorem. Let us set $\varphi(x,\lambda)=\frac{1}{\mathscr{L}(B_\rho(x_0))}1_{B_\rho(x_0)}(x)\phi(\lambda)$ which defines a $\mathscr{B}(\Omega)\otimes\mathscr{B}(E)$-measurable function, continuous with respect to λ. Consider now the sequence $(v_n)_{n\in\mathbf{N}}$ given by Lemma 11.4.1 with $\Omega=B_\rho(x_0)$. From the growth condition fulfilled by ϕ, the sequence $(\varphi(x,\nabla v_n(x)))_{n\in\mathbf{N}}$ is uniformly integrable. Therefore, by applying Theorem 4.3.3, we obtain

$$\begin{aligned}\int_{\Omega\times E}\varphi(x,\lambda)\,d\mu(x,\lambda)&=\lim_{n\to+\infty}\frac{1}{\mathscr{L}(B_\rho(x_0))}\int_{B_\rho(x_0)}\phi(\nabla v_n(x))\,dx\\&\geq\frac{1}{\mathscr{L}(B_\rho(x_0))}\int_{B_\rho(x_0)}\phi(\nabla u(x))\,dx.\end{aligned}\tag{11.48}$$

The last inequality is a consequence of lower semicontinuity of the integral functional

$$v\mapsto\int_{B_\rho(x_0)}\phi(\nabla v(x))\,dx$$

when $W^{1,p}(B_\rho(x_0),\mathbf{R}^m)$ is equipped with its weak convergence, due to the quasi-convexity assumption on ϕ (see Theorem 13.2.1). But, according to the slicing theorem, Theorem 4.2.4,

$$\int_{\Omega\times E}\varphi(x,\lambda)\,d\mu(x,\lambda)=\frac{1}{\mathscr{L}(B_\rho(x_0))}\int_{B_\rho(x_0)}\left(\int_E\phi(\lambda)\,d\mu_x(\lambda)\right)dx$$

so that (11.48) yields

$$\frac{1}{\mathscr{L}(B_\rho(x_0))}\int_{B_\rho(x_0)}\left(\int_E\phi(\lambda)\,d\mu_x(\lambda)\right)dx\geq\frac{1}{\mathscr{L}(B_\rho(x_0))}\int_{B_\rho(x_0)}\phi(\nabla u(x))\,dx.$$

The conclusion then follows by letting $\rho\to 0$.

PROOF OF THE SUFFICIENT CONDITIONS. The proof of the sufficient conditions is more involved. Furthermore, we prove that any Young measure satisfying conditions (i), (ii), and (iii) is generated by the sequence of gradients of a bounded sequence in $u+W_0^{1,p}(\Omega,\mathbf{R}^m)$. For the convenience of the reader we divide the proof into several steps.

First step: The Young measure μ is assumed to be homogeneous. Let a be a fixed matrix in E and Q a fixed open bounded cube of $\mathbf{R}^N$, and let l_a be the linear function defined for every $x\in Q$ by $l_a(x)=a.x$. We consider the closed set $\mathscr{H}_a(E)$ of probability measures μ on E satisfying the three following conditions:

(i) $\int_E\lambda\,d\mu=a$;

(ii) for all quasi-convex function ϕ satisfying for all $\lambda\in E$, $\gamma\leq\phi(\lambda)\leq\beta(1+|\lambda|^p)$ for some $\gamma\in\mathbf{R}$ and $\beta>0$, one has $\phi(a)\leq\int_E\phi(\lambda)\,d\mu(\lambda)$;

(iii) $\int_E|\lambda|^p\,d\mu<+\infty$.

Note that these three conditions are exactly the one stated in Theorem 11.4.1, fulfilled by the homogeneous Young measure $\mu\otimes\mathscr{L}\lfloor\Omega$. In this step we set $\Omega=Q$ and we aim to establish the existence of a bounded sequence $(u_n)_{n\in\mathbf{N}}$ in $l_a+W_0^{1,p}(Q,\mathbf{R}^m)$ such that the sequence $(\nabla u_n)_{n\in\mathbf{N}}$ generates the Young measure $\mu\otimes\mathscr{L}\lfloor Q$. In the second step, we will apply this result in the case when $a=0$.

For every $v \in l_a + W^{1,p}(Q,\mathbf{R}^m)$, let us consider the probability measure on E

$$\mu_v := \frac{1}{\mathscr{L}(Q)} \int_Q \delta_{\nabla v(x)} \, dx,$$

which acts on every $\varphi \in \mathbf{C}_0(E)$ as follows:

$$\langle \mu_v, \varphi \rangle = \frac{1}{\mathscr{L}(Q)} \int_Q \varphi(\nabla v(x)) \, dx.$$

Note that $\int_E |\lambda|^p \, d\mu_v$ is well defined and precisely $\int_E |\lambda|^p \, d\mu_v = \frac{1}{\mathscr{L}(Q)} \int_Q |\nabla v(x)|^p \, dx$. We finally consider the following convex subset of $\mathscr{H}_a(E)$:

$$\mathscr{C}_a(E) := \left\{ \mu_v \ : \ v \in l_a + W^{1,p}(Q,\mathbf{R}^m) \right\}.$$

Let now v be a fixed element of $l_a + W_0^{1,p}(Q,\mathbf{R}^m)$, extended by Q-periodicity on $\mathbf{R}^N$, and, for every $n \in \mathbf{N}^*$, let us define the function $v_n : x \mapsto v_n(x) := \frac{1}{n} v(nx)$ in $l_a + W^{1,p}(Q,\mathbf{R}^m)$. According to a classical result on oscillating functions, rephrased in terms of narrow convergence of Young measures, one has

$$(\delta_{\nabla v_n(x)})_{x \in \mathbf{Q}} \otimes \mathscr{L}\lfloor Q \overset{nar}{\rightharpoonup} \mu_v \otimes \mathscr{L}\lfloor Q \quad \text{when } n \to +\infty. \tag{11.49}$$

Note that the norm of ∇v_n in $L^p(Q,E)$ is exactly the one of ∇v. To conclude, it is enough to show that for every $\mu \in \mathscr{H}_a(E)$, there exists a sequence $(\mu_{w_k})_{k\in\mathbf{N}}$ in $\mathscr{C}_a(E)$, $(w_k)_{k\in\mathbf{N}}$ bounded in $l_a + W_0^{1,p}(Q,\mathbf{R}^m)$ such that

$$\mu_{w_k} \rightharpoonup \mu \text{ weakly in } \mathbf{M}(E), \quad \text{i.e., } \sigma(\mathbf{C}_0'(E), \mathbf{C}_0(E));$$

hence, according to Theorem 4.3.1, and since μ_{w_k} and μ are homogeneous,

$$\mu_{w_k} \otimes \mathscr{L}\lfloor Q \overset{nar}{\rightharpoonup} \mu \otimes \mathscr{L}\lfloor Q \quad \text{when } k \to +\infty. \tag{11.50}$$

Indeed, since the space $\mathscr{Y}(Q;E)$ endowed with the topology of the narrow convergence is metrizable (see [162, Proposition 2.3.1]), combining (11.49), (11.50), and by using a diagonalization argument (cf. Lemma 11.1.1), we will show that there exists a bounded sequence $(u_n)_{n\in\mathbf{N}^*}$ in $l_a + W^{1,p}(Q,\mathbf{R}^m)$, whose sequence of gradients generates the Young measure $\mu \otimes \mathscr{L}\lfloor Q$.

We are going to establish (11.50) or, equivalently, the density of $\mathscr{C}_a(E)$ in $\mathscr{H}_a(E)$ for the $\sigma(\mathbf{C}_0'(E), \mathbf{C}_0(E))$ topology. Let us point out that we want to prove the existence of a sequence $(\mu_{w_m})_{m\in\mathbf{N}}$ weakly converging to μ, such that moreover $(w_m)_{m\in\mathbf{N}}$ is bounded in $l_a + W_0^{1,p}(Q,\mathbf{R}^m)$ or, equivalently, such that $(\nabla w_m)_{m\in\mathbf{N}}$ is bounded in $L^p(Q,E)$. In order to take into account this condition, we establish $\overline{\mathscr{C}_a(E)} = \mathscr{H}_a(E)$ for a metric d' finer than the classical metric d inducing the $\sigma(\mathbf{C}_0'(E), \mathbf{C}_0(E))$ topology in the set $\mathbf{P}(E)$ of probability measures on E. Let us recall that given a dense countable family $(\varphi_i)_{i\in\mathbf{N}^*}$ in $\mathbf{C}_0(E)$, the distance d is given by

$$\forall (\mu,\nu) \in \mathbf{P}(E) \times \mathbf{P}(E) \quad d(\mu,\nu) := \sum_{i=1}^{+\infty} \frac{1}{2^i \|\varphi_i\|_\infty} |\langle \mu, \varphi_i \rangle - \langle \nu, \varphi_i \rangle|.$$

We define now the distance d' by setting, for every $(\mu,\nu)\in \mathbf{P}(E)\times\mathbf{P}(E)$,

$$d'(\mu,\nu):=\left|\int_E |\lambda|^p\,d\mu-\int_E|\lambda|^p\,d\nu\right|+\sum_{i=1}^{+\infty}\frac{1}{2^i\|\varphi_i\|_\infty}|\langle\mu,\varphi_i\rangle-\langle\nu,\varphi_i\rangle|,$$

and we argue by contradiction. Let us assume that $\mathscr{C}_a(E)$ is not dense in $\mathscr{H}_a(E)$ for the metric associated with the distance d'. Then, there exists $\mu_0\in\mathscr{H}_a(E)$, $k\in\mathbf{N}^*$ and $\eta>0$ such that

$$\forall\nu\in\mathscr{C}_a(E)\quad \left|\int_E|\lambda|^p\,d\mu_0-\int_E|\lambda|^p\,d\nu\right|+\sum_{i=1}^{k}\frac{1}{2^i\|\varphi_i\|_\infty}|\langle\mu_0,\varphi\rangle-\langle\nu,\varphi\rangle|>\eta. \quad (11.51)$$

We reason now in the finite dimensional space $\mathbf{R}^{k+1}$. From (11.51), the vector

$$\left(\int_E|\lambda|^p\,d\mu_0,\frac{1}{2^i\|\varphi_i\|_\infty}\langle\mu_0,\varphi_i\rangle\right)_{i=1,\dots,k}$$

does not belong to the closure of the following convex set of $\mathbf{R}^{k+1}$:

$$\mathbf{C}_a:=\left\{\left(\int_E|\lambda|^p\,d\nu,\frac{1}{2^i\|\varphi_i\|_\infty}\langle\nu,\varphi_i\rangle\right)_{i=1,\dots k}\;:\;\nu\in\mathscr{C}_a(E)\right\}.$$

Consequently, according to the Hahn–Banach separation theorem, Theorem 9.1.1, there exist $(c_i)_{i=0,\dots k}$ in $\mathbf{R}^{k+1}$ and $\eta'>0$ such that for all $\nu\in\mathscr{C}_a(E)$,

$$\begin{aligned}&c_0\int_E|\lambda|^p\,d\nu+\sum_{i=1}^{k}\frac{1}{2^i\|\varphi_i\|_\infty}c_i\langle\nu,\varphi_i\rangle\\ &\ge\eta'+c_0\int_E|\lambda|^p\,d\mu_0+\sum_{i=1}^{k}\frac{1}{2^i\|\varphi_i\|_\infty}c_i\langle\mu_0,\varphi_i\rangle.\end{aligned}\quad (11.52)$$

Let us set

$$\phi:=c_0|\cdot|^p+\sum_{i=1}^{k}\frac{c_i}{\|\varphi_i\|_\infty}\varphi_i.$$

We claim that $c_0\ge0$. Otherwise, taking in (11.52) $\nu=\mu_{v_t}$ with $v_t:=tw+l_a$, $t>0$, where w is a fixed function in $W_0^{1,p}(Q,\mathbf{R}^m)$, and letting $t\to+\infty$, the right-hand side in (11.52) would be $-\infty$. Replacing if necessary (when $c_0=0$), the function ϕ by the function $\phi+\delta|\cdot|^p$ with $\delta>0$ small enough (choose precisely δ satisfying $\eta'-\delta\int_E|\lambda|^p\,d\mu_0>0$), one may assume $c_0>0$. Thus the function ϕ still satisfies (11.52) for $\eta'>0$ replaced by $\eta'-\delta\int_E|\lambda|^p\,d\mu_0>0$ if necessary, together with the growth conditions $c_0|\lambda|^p+\gamma\le\phi(\lambda)\le\beta(1+|\lambda|^p)$ for some constants $\gamma\in\mathbf{R}$ and $\beta>0$. Finally, replacing ϕ by $\phi-\gamma$, one may assume that ϕ satisfies (11.52) and the growth conditions $c_0|\lambda|^p\le\phi(\lambda)\le\beta(1+|\lambda|^p)$ of Proposition 11.2.2. Therefore, (11.52) and condition (ii) satisfied by μ_0 yield

$$\inf\left\{\frac{1}{\mathscr{L}(Q)}\int_Q\phi(\nabla v(x))\,dx:v\in l_a+W_0^{1,p}(Q,\mathbf{R}^m)\right\}>\int_E\phi\,d\mu_0\ge\int_E Q\phi\,d\mu_0\ge Q\phi(a).$$

But, according to the classical variational principle (Proposition 11.2.2), since ϕ satisfies appropriate growth conditions,

$$\inf\left\{\frac{1}{\mathcal{L}(Q)}\int_Q \phi(\nabla v(x))\,dx : v \in l_a + W_0^{1,p}(Q,\mathbf{R}^m)\right\} = Q\phi(a),$$

a contradiction.

Second step: The Young measure μ satisfies conditions (i), (ii), *and* (iii) *of Theorem* 11.4.1, *and $u = 0$*. We want to construct a bounded sequence $(u_n)_{n\in\mathbf{N}}$ in $W^{1,p}(\Omega,\mathbf{R}^m)$ whose gradients generate μ. The idea of the proof consists in "localizing" $(\mu_x)_{x\in\Omega}$ thanks to Vitali's covering theorem, to apply the previous step for each localization, then to stick together the sequences of functions whose gradients generate each localized Young measures.

According to Vitali's covering lemma, Lemma 4.1.2, and Remark 4.1.4, for every $k \in \mathbf{N}^*$, there exists a finite family $(Q_{i,k})_{i\in I_k}$ of pairwise disjoint open cubes included in Ω satisfying

$$\mathcal{L}\left(\Omega \setminus \bigcup_{i\in I_k} Q_{i,k}\right) \le \frac{1}{k}, \quad \operatorname{diam}(Q_{i,k}) < \frac{1}{k}. \tag{11.53}$$

For every $i \in I_k$, let us define the probability measure $\mu_{i,k}$ on E by

$$\mu_{i,k} := \frac{1}{\mathcal{L}(Q_{i,k})}\int_{Q_{i,k}} \mu_x\,dx,$$

which acts on every $\varphi \in \mathbf{C}_0(E)$ as follows:

$$\langle \mu_{i,k},\varphi\rangle = \frac{1}{\mathcal{L}(Q_{i,k})}\int_{Q_{i,k}}\left(\int_E \varphi(\lambda)\,d\mu_x(\lambda)\right)dx.$$

With the notation of the first step, it is easy to show that $\mu_{i,k}$ belongs to $\mathcal{H}_0(E)$. Consequently, according to the first step, for every $i \in I_k$, there exists a bounded sequence $(v_{i,k,n})_{n\in\mathbf{N}}$ in $W_0^{1,p}(Q_{i,k},\mathbf{R}^m)$ such that

$$(\delta_{\nabla v_{i,k,n}(x)})_{x\in Q_{i,k}} \otimes \mathcal{L}\lfloor Q_{i,k} \overset{nar}{\rightharpoonup} \mu_{i,k} \otimes \mathcal{L}\lfloor Q_{i,k} \quad \text{when } n \to +\infty. \tag{11.54}$$

By using Lemma 11.4.1, one may furthermore assume the sequence $(|\nabla v_{i,k,n}|^p)_{n\in\mathbf{N}}$ uniformly integrable so that, from Theorem 4.3.3, (11.54) yields

$$\begin{aligned}\lim_{n\to+\infty}\int_{Q_{i,k}} |\nabla v_{i,k,n}|^p\,dx &= \int_{Q_{i,k}}\int_E |\lambda|^p\,d\mu_{i,k}\,dx \\ &= \frac{1}{\mathcal{L}_{1\lfloor(0,1)}(Q_{i,k})}\int_{Q_{i,k}}\left(\int_E |\lambda|^p\,d\mu_x\right)dx.\end{aligned} \tag{11.55}$$

We now stick together the functions $v_{i,k,n}$, $i \in I_k$, by setting

$$v_{k,n}(x) := \begin{cases} v_{i,k,n} & \text{if } x \in Q_{i,k}, \\ 0 & \text{if } x \in \Omega \setminus \bigcup_{i\in I_k} Q_{i,k}. \end{cases}$$

Clearly $(v_{k,n})_{n\in\mathbf{N}}$ is a bounded sequence in $W_0^{1,p}(\Omega,\mathbf{R}^m)$. Take now θ in a dense subset of regular functions of $L^1(\Omega)$, $\varphi \in C_0(E)$ and set $R_k := \int_{\Omega\setminus\bigcup_{i\in I_k} Q_{i,k}} \theta(x)\varphi(0)\,dx$. Note that from (11.53), $\lim_{k\to+\infty} R_k = 0$. From (11.54), the definition of $\mu_{i,k}$, and according to the mean value theorem, one has

$$\begin{aligned}
\lim_{n\to+\infty}\int_\Omega \theta(x)\varphi(\nabla v_{k,n}(x))\,dx &= \lim_{n\to+\infty}\sum_{i\in I_k}\int_{Q_{i,k}} \theta(x)\varphi(\nabla v_{k,n}(x))\,dx + R_k\\
&= \sum_{i\in I_k}\int_{Q_{i,k}} \theta(x)\langle \mu_{i,k},\varphi\rangle\,dx + R_k\\
&= \sum_{i\in I_k}\int_{Q_{i,k}} \left(\int_E \varphi\,d\mu_y\right)dy\,\frac{1}{\mathscr{L}(Q_{i,k})}\int_{Q_{i,k}}\theta(x)\,dx + R_k\\
&= \sum_{i\in I_k}\int_{Q_{i,k}} \theta(x_{i,k})\left(\int_E \varphi\,d\mu_y\right)dy + R_k
\end{aligned} \tag{11.56}$$

for some $x_{i,k}\in Q_{i,k}$. We stress the fact that the convergence in (11.56) may be taken in the narrow convergence sense. Indeed, setting

$$\tilde{\mu}_x = \begin{cases} \mu_{i,k} & \text{if } x\in Q_{i,k},\\ \delta_0 & \text{if } x\in\Omega\setminus\bigcup_{i\in I_k} Q_{i,k}, \end{cases}$$

one defines a Young measures $(\tilde{\mu}_x)_{x\in\Omega}\otimes\mathscr{L}$ and estimate (11.56) yields

$$\lim_{n\to+\infty}(\delta_{\nabla v_{k,n}(x)})_{x\in\Omega}\otimes\mathscr{L} = (\tilde{\mu}_x)_{x\in\Omega}\otimes\mathscr{L}$$

in the narrow convergence sense in $\mathscr{Y}(\Omega;E)$. Letting $k\to+\infty$ in (11.56), from (11.53), one obtains

$$\begin{aligned}
\lim_{k\to+\infty}\lim_{n\to+\infty}\int_\Omega \theta(x)\varphi(\nabla v_{k,n}(x))\,dx &= \lim_{k\to+\infty}\sum_{i\in I_k}\int_{Q_{i,k}}\theta(x_{i,k})\left(\int_E\varphi\,d\mu_y\right)dy\\
&= \lim_{k\to+\infty}\sum_{i\in I_k}\int_{Q_{i,k}}\theta(y)\left(\int_E\varphi\,d\mu_y\right)dy\\
&= \int_\Omega \theta(y)\left(\int_E\varphi\,d\mu_y\right)dy.
\end{aligned}$$

This shows that $\lim_{k\to+\infty}\lim_{n\to+\infty}(\delta_{\nabla v_{k,n}(x)})_{x\in\Omega}\otimes\mathscr{L} = (\mu_x)_{x\in\Omega}\otimes\mathscr{L}$, where each limit must be taken in the narrow convergence sense in $\mathscr{Y}(\Omega;E)$. According to the diagonalization lemma, Lemma 11.1.1, there exists a map $n\mapsto k(n)$ such that

$$(\delta_{\nabla v_{k(n),n}(x)})_{x\in\Omega}\otimes\mathscr{L} \overset{nar}{\rightharpoonup} (\mu_x)_{x\in\Omega}\otimes\mathscr{L}$$

when $n\to+\infty$. Furthermore, the sequence $(v_{k(n),n})_{n\in\mathbf{N}}$ is bounded in $W^{1,p}(\Omega,\mathbf{R}^m)$. Setting $u_n := v_{k(n),n}$ shows what required.

Last step: The Young measure μ satisfies conditions (i), (ii), *and* (iii) *of Theorem* 11.4.1 *without condition on u in $W^{1,p}(\Omega,\mathbf{R}^m)$.* Given $\mu\in\mathscr{Y}(\Omega;E)$ satisfying (i), (ii), and (iii), let us consider $\tilde{\mu} = (\tilde{\mu}_x)_{x\in\Omega}\otimes\mathscr{L}$ in $\mathscr{Y}(\Omega;E)$ defined by $\langle\tilde{\mu}_x,\varphi\rangle = \langle\mu_x,\varphi(.-\nabla u(x))\rangle$ for every $\varphi\in L^p_{\mu_x}(E)$ and for a.e. x in Ω. It is straightforward to check that $\tilde{\mu}$ satisfies the conditions of the second step. Thus, there exists a sequence $(v_n)_{n\in\mathbf{N}}$ in $W_0^{1,p}(\Omega,\mathbf{R}^m)$

such that $(\nabla v_n)_{n\in\mathbf{N}}$ generates $\tilde{\mu}$. Consider $u_n := v_n + u$ in $W^{1,p}(\Omega,\mathbf{R}^m)$. For each $\psi \in \mathbf{C}_b(\Omega;E)$, let us define $\tilde{\psi}$ in $\mathbf{C}_b(\Omega;E)$ by setting $\tilde{\psi}(x,\lambda) := \psi(x,\lambda+\nabla u(x))$. Thus, for every $\psi \in \mathbf{C}_b(\Omega;E)$ we have

$$\begin{aligned}\lim_{n\to+\infty}\int_\Omega \psi(x,\nabla u_n(x))\,dx &= \lim_{n\to+\infty}\int_\Omega \tilde{\psi}(x,\nabla v_n(x))\,dx\\ &= \int_{\Omega\times E}\tilde{\psi}(x,\lambda)\,d\tilde{\mu}(x,\lambda)\\ &= \int_{\Omega\times E}\psi(x,\lambda)d\mu(x,\lambda),\end{aligned}$$

which proves that $(\nabla u_n)_{n\in\mathbf{N}}$ generates μ. ☐

Remark 11.4.1. When considering functions depending on x and u, the necessary condition (ii) must be replaced by condition (ii)′ below: for every $\Phi : \Omega\times\mathbf{R}^m\times E \longrightarrow \mathbf{R}$ such that for a.e. x in Ω, $\phi(x,u(x),.)$ is quasi-convex and satisfies

$$\gamma \le \phi(x,\xi,\lambda) \le \beta(1+|\xi|^p+|\lambda|^p)$$

for some constants γ and $\beta > 0$, one has

$$\phi(x,u(x),\nabla u(x)) \le \int_E \phi(x,u(x),\lambda)\,d\mu_x(\lambda) \quad \text{for a.e. } x\in\Omega.$$

Indeed, it suffices to argue as in the second step of the proof of the necessary conditions by setting

$$\varphi(x,\lambda) = \frac{1}{\mathscr{L}(B_\rho(x_0))}1_{B_\rho(x_0)}(x)\phi(x,u(x),\lambda),$$

where x_0 is such that the two following limits exist:

$$\begin{aligned}\phi(x_0,u(x_0),\nabla u(x_0)) &= \lim_{\rho\to 0}\frac{1}{\mathscr{L}(B_\rho(x_0))}\int_{B_\rho(x_0)}\phi(x,u(x),\nabla u(x))\,dx,\\ \int_E \phi(x_0,u(x_0),\lambda)\,d\mu_{x_0}(\lambda) &= \lim_{\rho\to 0}\frac{1}{\mathscr{L}(B_\rho(x_0))}\int_{B_\rho(x_0)}\left(\int_E \phi(x,u(x),\lambda)\,d\mu_x\right)dx.\end{aligned}$$

11.4.2 ▪ Relaxation of classical integral functionals in $\mathscr{Y}(\Omega;E)$

We intend to apply the relaxation procedure for the integral functionals of Section 11.2, but considered as living in the space $X = \mathscr{Y}(\Omega;E)$, equipped with the topology of the narrow convergence. The generalized solutions of the relaxed problem may be interpreted as microstructures: they capture highly oscillatory minimizing sequences on smaller and smaller spatial scales and describe fine phase mixtures in elastic crystals. Fundamental applications to real materials and polycrystals may be found in [96], [97], [98], [99], and [172]. For papers dealing with laminates and multiwell problems, see [81], [307], [343], [317], and [365].

We consider the problem

$$(\mathscr{P}) \qquad \inf\left\{F(u) \,:\, u\in u_0 + W_0^{1,p}(\Omega,\mathbf{R}^m)\right\} \qquad (:= \inf(\mathscr{P})),$$

where u_0 is a given function in $W^{1,p}(\Omega,\mathbf{R}^m)$, $p>1$, and $F: W^{1,p}(\Omega,\mathbf{R}^m)\longrightarrow\mathbf{R}$ is the integral functional defined by

$$F(u)=\int_\Omega f(x,u,\nabla u)\,dx.$$

The density $f:\Omega\times\mathbf{R}^N\times E\longrightarrow\mathbf{R}$ is assumed to be $\mathscr{B}(\Omega)\otimes\mathscr{B}(\mathbf{R}^N)\otimes\mathscr{B}(E)$-measurable, continuous with respect to the third variable, to satisfy the continuity assumption

$$|f(x,\xi,\lambda)-f(x,\xi',\lambda)|\le L|\xi-\xi'|(1+|\xi|^{p-1}+|\xi'|^{p-1}) \tag{11.57}$$

with respect to the second variable, and the usual bounds

$$\alpha(|\lambda|^p-1)\le f(x,\xi,\lambda)\le\beta(1+|\xi|^p+|\lambda|^p),$$

where $L>0$, $0<\alpha<\beta$ are three given constants. Take, for example, $f(x,\xi,\lambda)=f_0(x,\xi)+f_1(x,\lambda)$. For less restrictive continuity assumptions on f, see Acerbi and Fusco [3], Dacorogna [182], and Pedregal [316].

In general $(\mathscr{P})$ has no solution and the relaxation procedure applied to this problem with $X=W^{1,p}(\Omega,\mathbf{R}^m)$ equipped with its weak convergence leads to the classical relaxed problem

$$(\overline{\mathscr{P}})\qquad \min\left\{\overline{F}(u)\,:\,u\in u_0+W_0^{1,p}(\Omega,\mathbf{R}^m)\right\}\qquad (:=\min(\overline{\mathscr{P}})),$$

where $\overline{F}$ is the integral functional defined on $W^{1,p}(\Omega,\mathbf{R}^m)$ by

$$\overline{F}(u)=\int_\Omega Qf(x,u,\nabla u)\,dx,$$

and, for a.e. x in Ω, $Qf(x,\xi,.)$ is the quasi-convex envelope of $f(x,\xi,.)$. According to Theorem 11.1.2 and arguing as in the proof of Theorem 11.2.2, one can establish $\inf(\mathscr{P})=\min(\overline{\mathscr{P}})$.

Because of the quasi-convexification, this procedure has the disadvantage of erasing the possible potential wells of $f(x,\xi,.)$ so that the relaxed problem does not provide much information on the behavior of minimizing sequences (see Example 11.4.1 below or examples given in Geymonat [224]). An alternative way, as we will see now, is to relax $(\mathscr{P})$ in the space $\mathscr{Y}(\Omega;E)$ equipped with the narrow convergence.

Let us first give some definitions and notation. For a fixed u_0 in $W^{1,p}(\Omega,\mathbf{R}^m)$, let $G\mathscr{Y}_0(\Omega;E)$ be the set of $W^{1,p}$-Young measures such that the underlying deformation u defined in Theorem 11.4.1 belongs to $u_0+W_0^{1,p}(\Omega,\mathbf{R}^m)$. We define its subset of elementary $W^{1,p}$-Young measures by

$$EG\mathscr{Y}_0(\Omega;E):=\left\{\mu\in\mathscr{Y}(\Omega;E):\exists u\in u_0+W_0^{1,p}(\Omega,\mathbf{R}^m),\ \mu=(\delta_{\nabla u(x)})_{x\in\Omega}\otimes\mathscr{L}\right\}.$$

Since $u\in u_0+W_0^{1,p}(\Omega,\mathbf{R}^m)$ is uniquely defined by its gradient $\nabla u\in L^p(\Omega,E)$, the operator

$$T:G\mathscr{Y}_0(\Omega;E)\longrightarrow u_0+W_0^{1,p}(\Omega,\mathbf{R}^m),\quad \mu\mapsto T\mu:=u,\ \nabla u(x)=\int_E\lambda\,d\mu_x(\lambda),$$

is well defined.

We now reformulate the problem $(\mathscr{P})$ in terms of Young measures by considering the integral functional $G : \mathscr{Y}(\Omega; E) \longrightarrow \mathbf{R} \cup \{+\infty\}$, defined by

$$G(\mu) = \begin{cases} \displaystyle\int_{\Omega \times E} f(x, T\mu(x), \lambda)\, d\mu(x, \lambda) & \text{if } \mu \in EG\mathscr{Y}_0(\Omega; E), \\ +\infty & \text{otherwise.} \end{cases}$$

Note that, in its domain, the functional G is nothing but the functional F. More precisely,

$$G(\mu) = \int_\Omega f(x, u(x), \nabla u(x))\, dx = F(u)$$

when $u \in u_0 + W_0^{1,p}(\Omega, \mathbf{R}^m)$ and $\mu = (\delta_{\nabla u(x)})_{x\in\Omega} \otimes \mathscr{L}$. The problem $(\mathscr{P})$ is then equivalent to

$$\inf\left\{G(\mu) \;:\; \mu \in \mathscr{Y}(\Omega; E)\right\}.$$

On the other hand, let us define the integral functional $\overline{G} : \mathscr{Y}(\Omega; E) \longrightarrow \mathbf{R} \cup \{+\infty\}$ by setting

$$\overline{G}(\mu) = \begin{cases} \displaystyle\int_{\Omega \times E} f(x, T\mu(x), \lambda)\, d\mu(x, \lambda) & \text{if } \mu \in G\mathscr{Y}_0(\Omega; E), \\ +\infty & \text{otherwise.} \end{cases}$$

According to Theorem 11.4.1(iii) and to the growth conditions fulfilled by f, the domain of $\overline{G}$ is $G\mathscr{Y}_0(\Omega; E)$. The functional $\overline{G}$ is nothing but the natural extension of G to $G\mathscr{Y}_0(\Omega; E)$. In Theorem 11.4.2, we show that $\overline{G}$ is the sequential lower semicontinuous envelope of G when $\mathscr{Y}(\Omega; E)$ is equipped with the narrow convergence and we make precise the relationship between the two relaxed problems in $W^{1,p}(\Omega, \mathbf{R}^m)$ and $\mathscr{Y}(\Omega; E)$.

Theorem 11.4.2. *The sequential lower semicontinuous envelope of G in $\mathscr{Y}(\Omega; E)$ equipped with the narrow convergence is the extended functional $\overline{G}$. Moreover, we have*

$$\inf(\mathscr{P}) = \min(\overline{\mathscr{P}}) = \min(\overline{\mathscr{P}}^{young}),$$

where

$$\min(\overline{\mathscr{P}}^{young}) := \min\left\{\overline{G}(\mu) \;:\; \mu \in \mathscr{Y}(\Omega; E)\right\},$$

and, if $\overline{\mu}$ is a solution of $\min(\overline{\mathscr{P}}^{young})$, then $\overline{u} = T\overline{\mu}$ is a solution of $\min(\overline{\mathscr{P}})$. Furthermore, if $(u_n)_{n\in\mathbf{N}}$ is a sequence of $1/n$-minimizers of the problem $\inf(\mathscr{P})$, then every cluster point of $(\delta_{\nabla u_n(x)})_{x\in\Omega} \otimes \mathscr{L}$ converges narrowly to a solution of $\min(\overline{\mathscr{P}}^{young})$.

PROOF. We begin by proving that $\overline{G}$ is the lsc envelope of G. According to Proposition 11.1.1 and Theorem 11.1.1 we must establish the two following assertions: for every $\mu \in \mathscr{Y}(\Omega; E)$,

$$\forall \mu_n \overset{nar}{\rightharpoonup} \mu,\ \overline{G}(\mu) \leq \liminf_{n\to+\infty} G(\mu_n), \tag{11.58}$$

$$\text{there exists a sequence } (\nu_n)_{n\in\mathbf{N}} \text{ in } \mathscr{Y}(\Omega; E),\ \nu_n \overset{nar}{\rightharpoonup} \mu, \text{ such that } \overline{G}(\mu) \geq \limsup_{n\to+\infty} G(\nu_n). \tag{11.59}$$

In a second step, in order to apply Theorem 11.1.1, we will establish the compactness of every minimizing sequence of $\inf(\mathscr{P})$, in terms of Young measures. We will conclude by giving the relations between a solution of $\min(\overline{\mathscr{P}}^{young})$ and the corresponding underlying deformation.

First step: Proof of (11.58). One may assume $\liminf_{n\to+\infty} G(\mu_n) < +\infty$, so that, for a nonrelabeled subsequence, one has $G(\mu_n) < +\infty$, $\mu_n \in EG\mathscr{Y}_0(\Omega;E)$, and

$$G(\mu_n) = \int_\Omega f(x, u_n(x), \nabla u_n(x))\, dx.$$

According to the coerciveness assumption on f, there exists $u \in u_0 + W_0^{1,p}(\Omega, \mathbf{R}^m)$ such that $u_n \rightharpoonup u$ weakly in $W^{1,p}(\Omega, \mathbf{R}^m)$ and strongly in $L^p(\Omega, \mathbf{R}^m)$. Let us write

$$\begin{aligned} G(\mu_n) &= \int_\Omega f(x, u_n(x), \nabla u_n(x))\, dx \\ &= \int_{\Omega\times E} f(x, u(x), \lambda)\, d\mu_n(x, \lambda) \\ &\quad + \left(\int_\Omega f(x, u_n(x), \nabla u_n(x))\, dx - \int_\Omega f(x, u(x), \nabla u_n(x))\, dx \right). \end{aligned}$$

From continuity assumption (11.57), and the boundedness of the sequence $(u_n)_{n\in\mathbf{N}}$ in $W^{1,p}(\Omega, \mathbf{R}^m)$, the second term in the right-hand side tends to zero. Since, for a.e. $x \in \Omega$, $\lambda \mapsto f(x, u(x), \lambda)$ is lsc (actually continuous), the conclusion then follows from Proposition 4.3.3.

Second step: Proof of (11.59). One may assume $\overline{G}(\mu) < +\infty$ so that $\mu \in G\mathscr{Y}_0(\Omega;E)$. According to the definition of $G\mathscr{Y}_0(\Omega;E)$, there exists $u \in u_0 + W_0^{1,p}(\Omega, \mathbf{R}^m)$ and a bounded sequence $(u_n)_{n\in\mathbf{N}}$ in $W^{1,p}(\Omega, \mathbf{R}^m)$ such that $\mu_n = (\delta_{\nabla u_n(x)})_{x\in\Omega} \otimes \mathscr{L}$ narrowly converges to μ. But, from Lemma 11.4.1, there exists a sequence $(v_n)_{n\in\mathbf{N}}$ in $W^{1,p}(\Omega, \mathbf{R}^m)$ satisfying

(i) $v_n \in u + W_0^{1,p}(\Omega, \mathbf{R}^m)$;

(ii) $(|\nabla v_n|^p)_{n\in\mathbf{N}}$ is uniformly integrable;

(iii) $v_n - u_n \to 0$ and $\nabla(v_n - u_n) \to 0$ in measure.

Hence by (iii) and Proposition 4.3.8, $\nu_n := (\delta_{\nabla v_n(x)})_{x\in\Omega} \otimes \mathscr{L}$ narrowly converges to μ. Moreover, it is easily seen that, up to a subsequence, v_n strongly converges to u in $L^p(\Omega, \mathbf{R}^m)$. From now on we consider a nonrelabeled subsequence of $(v_n)_{n\in\mathbf{N}}$ such that v_n strongly converges to u in $L^p(\Omega, \mathbf{R}^m)$. Let us write

$$\begin{aligned} \int_\Omega f(x, v_n, \nabla v_n)\, dx &= \int_\Omega f(x, u, \nabla v_n)\, dx \\ &\quad + \left(\int_\Omega f(x, v_n, \nabla v_n)\, dx - \int_\Omega f(x, u, \nabla v_n)\, dx \right). \end{aligned} \tag{11.60}$$

According to the growth condition satisfied by f, the sequence $(f(x, u, \nabla v_n))_{n\in\mathbf{N}}$ is also uniformly integrable, so that, by Theorem 4.3.3,

$$\lim_{n\to+\infty} \int_\Omega f(x, u, \nabla v_n)\, dx = \int_{\Omega\times E} f(x, u, \lambda)\, d\mu(x, \lambda). \tag{11.61}$$

On the other hand, from (11.57) and the boundedness of the sequence $(v_n)_{n\in\mathbf{N}}$ in $L^p(\Omega,\mathbf{R}^m)$, the second term in the right-hand side of (11.60) tends to zero. Collecting (11.60) and (11.61), we obtain

$$\lim_{n\to+\infty} G(\nu_n) = \overline{G}(\mu),$$

which completes the proof of the second step.

Third step: Compactness of minimizing sequences. Let $(\mu_n)_{n\in\mathbf{N}}$ be a minimizing sequence of $\inf(\mathscr{P})$. Since $G(\mu_n) < +\infty$, there exists $(u_n)_{n\in\mathbf{N}}$ in $u_0 + W_0^{1,p}(\Omega,\mathbf{R}^m)$ such that $\mu_n = (\delta_{\nabla u_n(x)})_{x\in\Omega} \otimes \mathscr{L}$ and $G(\mu_n) = F(u_n)$. According to the coerciveness assumption on f, the sequence $(\nabla u_n)_{n\in\mathbf{N}}$ is bounded in $L^p(\Omega,E)$; thus, from Remark 4.3.3, the sequence $(\mu_n)_{n\in\mathbf{N}}$ is tight. The conclusion then follows by applying Prokhorov's theorem, Theorem 4.3.2.

Last step. Let $\overline{\mu}$ be a solution of $\min(\overline{\mathscr{P}}^{young})$ and $\overline{u} = T\overline{\mu}$. According to Theorem 11.4.1(ii) and Remark 11.4.1, we have, for a.e. $x \in \Omega$,

$$\begin{aligned} Qf(x,\overline{u}(x),\nabla\overline{u}(x)) &\le \int_E Qf(x,\overline{u}(x),\lambda)\, d\overline{\mu}_x(\lambda) \\ &\le \int_E f(x,\overline{u}(x),\lambda)\, d\overline{\mu}_x(\lambda). \end{aligned}$$

Therefore

$$\begin{aligned} \int_\Omega Qf(x,\overline{u}(x),\nabla\overline{u}(x))\, dx &\le \int_{\Omega\times E} f(x,T\overline{\mu}(x),\lambda)\, d\overline{\mu}(x,\lambda) \\ &= \min(\overline{\mathscr{P}}^{young}) = \inf(\mathscr{P}) = \min(\overline{\mathscr{P}}). \end{aligned}$$

This shows that $\overline{u} = T\overline{\mu}$ is a solution of $\min(\overline{\mathscr{P}})$. The last assertion is a straightforward consequence of Theorem 11.1.2. □

Example 11.4.1. Let us illustrate Theorem 11.4.2 with the example treated in Geymonat [224]: take $\Omega = (0,1)$, $p = 4$, $u_0 = 0$, and $\varphi(\xi,\lambda) = (\lambda^2-1)^2 + \xi^2$; see Figure 11.1. The map $\lambda \mapsto \varphi(\xi,\lambda)$ possesses the two potential wells ± 1 and $\inf(\mathscr{P}) = 0$ has no solution.

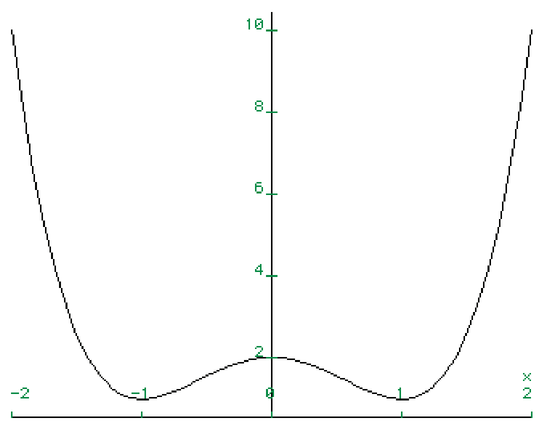

Figure 11.1. *The graph of* $\lambda \mapsto \varphi(\xi,\lambda)$ *with* $\xi = 1$.

For each fixed ξ, the quasi-convex envelope of $\lambda \mapsto (\lambda^2-1)^2+\xi^2$ (equivalently, its convex envelope) is given by

$$Q\varphi(x,\xi,\lambda)=\begin{cases}(\lambda^2-1)^2+\xi^2 & \text{if } |\lambda|\geq 1,\\ \xi^2 & \text{otherwise,}\end{cases}$$

so that $\min(\overline{\mathscr{P}})=0$ possesses a unique solution $\overline{u}=0$. Therefore, from Theorem 11.4.2 we have $\min(\overline{\mathscr{P}}^{young})=0$ and if $\overline{\mu}$ is a solution

$$\int_{(0,1)}\left(\int_{\mathbf{R}}((\lambda^2-1)^2+(T\overline{\mu})^2)d\overline{\mu}_x(\lambda)\right)dx=0, \quad x \text{ a.e. on } (0,1),$$
$$0=\int_{\mathbf{R}}\lambda\, d\overline{\mu}_x(\lambda).$$

From the first equality, we deduce that the support of $\overline{\mu}_x$ is include in $\{-1,1\}$. Therefore, the second equality gives $\overline{\mu}_x=\frac{1}{2}\delta_{-1}+\frac{1}{2}\delta_1$ and $\overline{\mu}=(\frac{1}{2}\delta_{-1}+\frac{1}{2}\delta_1)\otimes\mathscr{L}$.

11.5 ▪ Mass transportation problems

The mass transportation theory goes back to Gaspard Monge: in 1781 he proposed in [295] a model to describe the work necessary to move a mass distribution $\mu_1=\rho_1\,dx$ into a final destination $\mu_2=\rho_2\,dx$ (respectively, *déblais* and *remblais* in Monge's terminology), once the unitary transportation cost function $c(x,y)$, which measures the work to move a unit mass from x to y, is given. Monge's goal was to find, among all possible *transportation maps* T which move μ_1 into μ_2, i.e., such that

$$\mu_2(E)=\mu_1\big(T^{-1}(E)\big) \qquad \text{for every measurable set } E,$$

a map with minimal total transportation cost, defined as

$$\int c\big(x,T(x)\big)\,d\mu_1.$$

The measures μ_1 and μ_2 on $\mathbf{R}^N$ (more generally on metric spaces) have equal mass (normalized to one for simplicity) and are called *marginals*; the operator

$$T^{\#}\mu(E)=\mu\big(T^{-1}(E)\big)$$

is called the *push forward* operator and is characterized by the fact that

$$\int \phi\, dT^{\#}\mu=\int \phi\circ T\, d\mu$$

for every Borel function $\phi:\mathbf{R}^N\to\overline{\mathbf{R}}$ which makes the integrals above meaningful.

Note that when the measures μ_1 and μ_2 are absolutely continuous with densities ρ_1 and ρ_2, respectively, and T is smooth and injective, the push forward equality reads

$$\rho_2(T(x))|\det\nabla T(x)|=\rho_1(x) \qquad \text{for a.e. } x\in\mathbf{R}^N.$$

With the notation above, the optimal mass transportation problems is written as

$$\min\left\{\int c\big(x,T(x)\big)\,d\mu_1\ :\ T^{\#}\mu_1=\mu_2\right\}. \tag{11.62}$$

The literature on mass transportation problems is very wide and several excellent publications are available; our goal here is not to be exhaustive but only to give to the reader a short presentation of the field, referring, for instance, to Ambrosio [20], Brenier [133], Evans and Gangbo [210], Villani [358], [359], and the references therein.

The natural framework for this kind of problems is the one where X is a metric space and μ_1, μ_2 are probabilities on X; however, the existence of an optimal transport map is a very delicate question, even in the classical Monge case, where X is the Euclidean space $\mathbf{R}^N$ and $c(x,y) = |x-y|$. When μ_1, μ_2 are singular measures, easy counterexamples show that an optimal transport map may not exist.

Example 11.5.1. On the real line $\mathbf{R}$ consider the probabilities $\mu_1 = \delta_0$ and $\mu_2 = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$. Then no transport map between μ_1 and μ_2 exists, and so the Monge problem (11.62) is meaningless for any cost function $c(x,y)$. The same situation occurs in general when μ_1 has h atoms and μ_2 has k atoms with $h < k$.

When the marginal measures are discrete (i.e., finite sums of Dirac masses) and with the same number of atoms,

$$\mu_1 = \frac{1}{k}\sum_{i=1}^{k}\delta_{x_i}, \qquad \mu_2 = \frac{1}{k}\sum_{i=1}^{k}\delta_{y_i},$$

the Monge problem becomes a problem in combinatorial optimization. In fact, a transport map T corresponds to a permutation of the indices and, taking, for instance, $c(x,y) = |x-y|$, the Monge problem becomes

$$\min\left\{\sum_{i=1}^{k}|x_i - y_{h(i)}| \ : \ h \text{ permutation of } 1,\dots,k\right\}.$$

An example with $k = 4$ is shown in Figure 11.2.

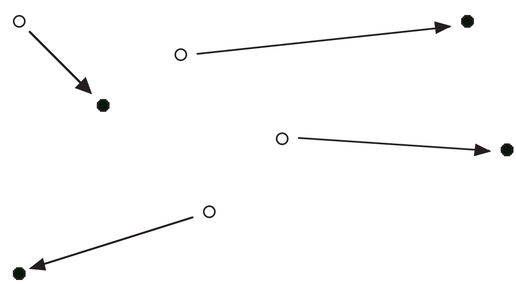

Figure 11.2. *A transport map in a discrete case.*

Example 11.5.2. It may also happen that the family of admissible transport maps in nonempty but an optimal transport map does not exist: take, for instance, in $\mathbf{R}^2$ the segments $A = \{(0,y) \ : \ y \in [0,1]\}$, $B = \{(1,y) \ : \ y \in [0,1]\}$, $C = \{(-1,y) \ : \ y \in [0,1]\}$, and the probabilities $\mu_1 = \mathcal{H}^1\lfloor A$, $\mu_2 = \frac{1}{2}\mathcal{H}^1\lfloor B + \frac{1}{2}\mathcal{H}^1\lfloor C$ (see Figure 11.3). In this case the family of transport maps between μ_1 and μ_2 is nonempty; for instance, the map

$$T(0,y) = \begin{cases} (-1,2y) & \text{if } y \in [0,1/2], \\ (1,2y-1) & \text{if } y \in]1/2,1] \end{cases}$$

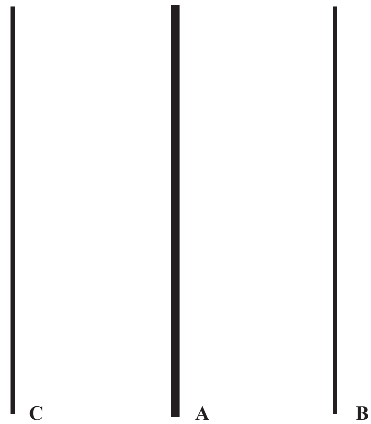

Figure 11.3. *A case in which an optimal transport map does not exist.*

moves μ_1 into μ_2 with cost

$$\int |x-T(x)|\,d\mu_1 = \frac{\sqrt{5}}{4} + \log\frac{1+\sqrt{5}}{2} \approx 1.04.$$

Taking the maps

$$T_n(0,y) = \begin{cases} (-1, 2y-\frac{k}{n}) & \text{if } y \in [\frac{k}{n}, \frac{k+1}{n}] \text{ and } k \text{ is even,} \\ (1, 2y-\frac{k+1}{n}) & \text{if } y \in]\frac{k}{n}, \frac{k+1}{n}[\text{ and } k \text{ is odd,} \end{cases}$$

an easy computation shows that

$$\lim_{n\to\infty} \int |x-T_n(x)|\,d\mu_1 = 1$$

so that

$$\inf\left\{ \int |x-T(x)|\,d\mu_1 \; : \; T^{\#}\mu_1 = \mu_2 \right\} \le 1. \tag{11.63}$$

On the other hand, for every transport map T we have

$$|x-T(x)| \ge 1 \qquad \forall x \in A,$$

which implies that the infimum in (11.63) is equal to 1. However, the infimum in (11.63) cannot be attained, because this would imply that for a transport map T we have

$$|x-T(x)| = 1 \qquad \text{for a.e. } x \in A,$$

which is impossible.

Example 11.5.3. In general, we cannot expect the uniqueness of optimal transport maps; for instance, in the one-dimensional case and with $c(x,y)=|x-y|$, the following example, known as *book shifting* for the analogy of the movement of volumes between shelves, shows that there are infinitely many optimal transport maps. Let $\mathscr{L}^1$ be the Lebesgue measure on $\mathbf{R}$; consider $\mu_1 = \mathscr{L}^1\lfloor[0,1]$ and $\mu_2 = \mathscr{L}^1\lfloor[1,2]$. It is not difficult to verify that both the transport maps $T_1(x) = 1 + x$ (translation) and $T_2(x) = 2 - x$ (reflection) are optimal with cost

$$\int_0^1 |x - T_1(x)|\, dx = \int_0^1 |x - T_2(x)|\, dx = 1.$$

In a similar way we may construct infinitely many optimal transport maps, for instance, subdividing the interval $[0,1]$ into n equal intervals I_k and the interval $[1,2]$ into n equal intervals J_k, and transporting every I_k into J_k by maps similar to T_1 and T_2.

Actually, in this case every transport map is optimal. In fact, by the definition of the push forward operator, we have

$$\int_0^1 T(x)\, dx = \int_1^2 x\, dx = \frac{3}{2};$$

therefore, since $T(x) \in [1,2]$ and $x \in [0,1]$, we obtain

$$\int_0^1 |x - T(x)|\, dx = \int_0^1 \big(T(x) - x\big)\, dx = \frac{3}{2} - \frac{1}{2} = 1.$$

The original problem was stated by Monge with the cost function $c(x,y) = |x-y|$; in this case the existence of an optimal transport map is a very delicate question; the following result was shown by Sudakov [339] in an incomplete form and successively completely proved by several authors (see Ambrosio and Pratelli [28], Evans and Gangbo [210], Caffarelli, Feldman, and McCann [163], and Trudinger and Wang [353]).

Theorem 11.5.1. *Let μ_1, μ_2 be two probabilities on the Euclidean space $\mathbf{R}^N$ with a compact support or more generally with finite first-order moments $\int |x|\, d\mu_i$. Assume that μ_1 is absolutely continuous with respect to the Lebesgue measure. Then the optimal transport problem*

$$\min\left\{ \int |x - T(x)|\, d\mu_1 \;:\; T^{\#}\mu_1 = \mu_2 \right\}$$

admits a solution.

To avoid the difficulties related to the existence of optimal transport maps, in 1942 Kantorovich proposed [252] a relaxed formulation of the Monge transport problem: the goal is now to find a probability on the product space, which minimizes the relaxed transportation cost

$$K(\mu_0, \mu_1) = \int c(x,y)\, \gamma(dx, dy) \tag{11.64}$$

over all admissible *transport plans* γ, that is, probabilities on $\mathbf{R}^N \times \mathbf{R}^N$ such that the projections $\pi_1^{\#}\gamma$ and $\pi_2^{\#}\gamma$ coincide with the marginals μ_1 and μ_2, respectively. Note that the projection condition above on the transport plan γ can be equivalently stated as

$$\gamma(B \times \mathbf{R}^N) = \mu_1(B) \quad \text{and} \quad \gamma(\mathbf{R}^N \times B) = \mu_2(B) \qquad \text{for every Borel set } B \subset \mathbf{R}^N.$$

Moreover, every transport map T can be seen as a transport plan γ, given by $\gamma = (Id \times T)^{\#}\mu_1$.

The Kantorovich problem then reads

$$\min\left\{\int c(x,y)\gamma(dx,dy)\ :\ \pi_j^{\#}\gamma=\mu_j \text{ for } j=1,2\right\}. \tag{11.65}$$

The cases $c(x,y)=|x-y|^p$ with $p\geq 1$ have been particularly studied, and the cost $K(\mu_0,\mu_1)$ in (11.64) provides, through the relation

$$W_p(\mu_0,\mu_1)=\big(K(\mu_0,\mu_1)\big)^{1/p},$$

the so-called *Wasserstein distance*. This distance metrizes the weak* convergence on the space $\mathbf{P}(\Omega)$ of probabilities supported in a compact set Ω. A very wide literature on the subject is available; we mention, for instance, the books [27], [358], [359], where one can find a complete list of references.

Theorem 11.5.2. *Assume that the cost function c is lower semicontinuous and bounded from below. Then the optimization Kantorovich problem* (11.65) *admits a solution.*

PROOF. Let $(\gamma_n)_{n\in\mathbf{N}}$ be a minimizing sequence for the Kantorovich optimization problem (11.65). If H,K are compact subsets of $\mathbf{R}^N$ we have

$$\begin{aligned}\gamma_n(\mathbf{R}^N\times\mathbf{R}^N\setminus H\times K)&=\gamma_n\Big(\big((\mathbf{R}^N\setminus H)\times\mathbf{R}^N\big)\cup\big(\mathbf{R}^N\times(\mathbf{R}^N\setminus K)\big)\Big)\\&\leq\gamma_n\big((\mathbf{R}^N\setminus H)\times\mathbf{R}^N\big)+\gamma_n\big(\mathbf{R}^N\times(\mathbf{R}^N\setminus K)\big)\\&=\mu_1(\mathbf{R}^N\setminus H)+\mu_2(\mathbf{R}^N\setminus K),\end{aligned}$$

where the last equality follows from the projection property

$$\pi_1^{\#}\gamma=\mu_1,\qquad \pi_2^{\#}\gamma=\mu_2.$$

Therefore, the sequence of measures $(\gamma_n)_{n\in\mathbf{N}}$ is tight, in the sense of Theorem 4.2.3 (see also Definition 4.3.2), and the Prokhorov compactness, Theorem 4.2.3, allows to extract a subsequence $(\gamma_{n_k})_{k\in\mathbf{N}}$ which converges narrowly to a probability γ on $\mathbf{R}^N\times\mathbf{R}^N$ in the sense of Definition 4.2.2. We show now that γ is a transport plan, that is, it satisfies the projection property: take a Borel set E of $\mathbf{R}^N$; for every compact set $K\subset E$ and every open set $A\supset E$, by using the Alexandrov proposition, Proposition 4.2.3, and the projection property of γ_n, we have

$$\begin{aligned}\mu_1(K)\ &=\limsup_{n\to\infty}\gamma_n(K\times\mathbf{R}^N)\leq\gamma(K\times\mathbf{R}^N)\leq\gamma(E\times\mathbf{R}^N)\\&\leq\gamma(A\times\mathbf{R}^N)\leq\liminf_{n\to\infty}\gamma_n(A\times\mathbf{R}^N)=\mu_1(A).\end{aligned}$$

Since K and A were arbitrarily chosen, we have

$$\gamma(E\times\mathbf{R}^N)=\mu_1(E).$$

In a similar way we obtain the second projection property

$$\gamma(\mathbf{R}^N\times E)=\mu_2(E).$$

To conclude, it remains to show that the transport plan γ is optimal for the Kantorovich problem (11.65). Since the cost function $c(x,y)$ is lower semicontinuous, there exists an increasing sequence $\left(c_k(x,y)\right)_{k\in\mathbf{N}}$ of continuous and bounded functions which converge to $c(x,y)$ pointwise. Then, by the narrow convergence of γ_n to γ, we have for every $k\in\mathbf{N}$

$$\begin{aligned}\int c_k(x,y)\,d\gamma(x,y) &= \lim_{n\to\infty}\int c_k(x,y)\,d\gamma_n(x,y)\\ &\leq \liminf_{n\to\infty}\int c(x,y)\,d\gamma_n(x,y) = \inf(11.65).\end{aligned}$$

Passing now to the limit as $k\to\infty$ gives the conclusion. □

Optimal transportation problems have strong links with several research fields and applications; we list some of them, pointing the interested reader to the related references:

- shape and density optimization for elastic structures (see, for instance, [113], [115]);
- optimization of transportation networks (see, for instance, [130], [156], [161]);
- problems in urban planning (see, for instance, [158], [159], [157]);
- optimal location and irrigation problems (see, for instance, [116], [117], [132], [154], [160]);
- traffic models with congestion effects (see, for instance, [164], [165], [360]);
- curves in the space of measures and applications to crowd movements (see, for instance, [131], [244], [289], [290]).