Lecture notes for Calculus of Variations

Lukas Koch[1]

[1] *Max Planck Institute for Mathematics in the Sciences, 04103 Leipzig,Germany*
`lkoch@mis.mpg.de`

August 8, 2023

Abstract

These are informal lecture notes for a 3rd year undergraduate course I taught in SS2023 at Leipzig University. They are largely based on (sections of) the books Introduction to the calculus of variations by Bernard Dacorogna, Direct Methods in the Calculus of Variations by Enrico Giusti and Calculus of Variations by Filip Rindler.

Contents

1 Introduction

Calculus of variations studies energy minimisation problems and properties of the mininisers. Many models of real-world problems, in physics, biology, chemistry, data science, and many more, can be seen as some system seeking to minimise an energy. In this course, we will develop a general theory to study such problems. We will begin by linking energy minimisation to a PDE, the Euler-Lagrange equation. The classical method of the calculus of variations is concerned with studying the Euler-Lagrange equations and to link them back to the minimisation problem, developing for example criteria to ascertain whether solutions of the Euler-Lagrange equation are a minimiser. We will go on to consider the direct method in the calculus of variations, which is a powerful tool to obtain existence results for minimisers. Moreover, we will study the regularity of minimisers in this framework. Finally, we will encounter Young measures as a useful tool to study the limits of sequences of integral functionals.

We begin by considering some important examples of the types of energy that we wish to study.

Example 1.1 (Shortest path). *Consider two points $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$. We ask to find the shortest path connecting them. Of course, it is clear that the solution is the straight line segment between x and y. However, it is less clear how to see this as an energy minimisation problem. In fact there are two natural approaches:*
a) (parametric approach) *Consider parametrised C^1-curves connecting x and y, that is C^1-functions $w\colon [0,1] \to \mathbb{R}^2$ such that $w(0) = x$ and $w(1) = y$. The length of w is given by*

$$L(w) = \int_0^1 |w'| \,\mathrm{d}t.$$

Thus, the minimisation problem we wish to consider is

$$\min_{\mathscr{A}} L(w) \qquad \text{where } \mathscr{A} = \{w \in C^1([0,1], \mathbb{R}^2)\colon w(0) = x, w(1) = y\}.$$

In this context, we expect that the following 'theorem' holds: u minimises L (in the sense that it is a minimum of the problem above) if and only if u is a monotone parametrisation of the line segment connecting x and y, i.e.

$$u = (1 - \tau(t))x + \tau(t)y \quad \text{where } \tau \text{ is a monontone map such that } \tau(0) = 0, \tau(1) = 1.$$

b) (non-parametric approach) *Here we view curves connecting x and y as graphs, i.e. as the image of a C^1-function $w\colon [x_1, y_1] \to \mathbb{R}$ with $w(x_1) = x_2$ and $w(y_1) = y_2$. The length of the curve is then given by*

$$L(w) = \left| \int_{x_1}^{x_2} \sqrt{1 + (w')^2} \,\mathrm{d}x \right|.$$

The theorem we now expect to hold is: u minimises L if and only if u is affine.

Example 1.2 (Planar isoperimetric problem)**.** *The planar isoperimetric problem can be formulated in three ways:*

- *Fix $L \geq 0$. Then the largest area A that it is possible to enclose with a closed curve of length L is $A = \frac{1}{4\pi}L^2$ with equality if and only if the curve is a circle.*
- *Fix $A \geq 0$. The shortest length of a closed curve needed to enclose an area of size A is $L = \sqrt{4\pi A}$ with equality if and only if L is a circle.*
- *If a closed curve of length L encloses an area of size A, then $A \leq \frac{1}{4\pi}L^2$ with equality if and only if the curve is a circle.*

It is easy to check that all three statements are equivalent. Again, we want to view this problem as minimising an energy. We will show how to do so in the setting of the second statement.
a) (parametric approach) *Parametrise the curve using $w \in C^1([0,1],\mathbb{R}^2)$. We encode the fact that w is closed, by demanding that $w(0) = w(1)$. Note that we may then compute the enclosed area as follows:*

$$\int_0^1 w_1 w_2' \,\mathrm{d}t = -\int_0^1 w_2 w_1' = \frac{1}{2}\int_0^1 w_1 w_2' - w_2 w_1' \,\mathrm{d}t.$$

The class of admissible functions is thus

$$\mathscr{A} = \left\{ w \in C^1([0,1],\mathbb{R}^2) \colon \int_0^1 w_1 w_2' \,\mathrm{d}t = A \text{ and } w(0) = w(1) \right\}$$

and the energy minimisation problem is

$$\min_{w \in \mathscr{A}} L(w)$$

where L is the parametric length functional defined in Example 1.1.
b) (non-parametric approach) *Note that graphs cannot be closed. Thus, we consider the following variant: Given $A \geq 0$, we consider the class of admissible functions*

$$\mathscr{A} = \{w \in C^1([x_1,x_2],\mathbb{R}) \colon w(x_1) = y_1,\ w(x_2) = y_2 \ \int_{x_1}^{x_2} w \,\mathrm{d}x = A\}.$$

The minimisation problem is then

$$\min_{w \in \mathscr{A}} L(w),$$

where L is the non-parametric length functional defined in Example 1.1.

Example 1.3. *[Brachistrone problem] The Brachistrone is the shape of the graph with endpoints (x_1, y_1) and (x_2, y_2) such that a mass starting at rest at (x_1, y_1) sliding along the graph reaches the point (x_2, y_2) in the shortest amount of time. Note that in order to model the situation we may assume that $y_2 \leq y_1$, as otherwise the particle cannot possibly reach (x_2, y_2) starting from (x_1, y_1). Further if $x_1 = x_2$, the solution is clearly described by free-fall and reflecting the x-axis if necessary, we may assume that $x_1 < x_2$. If we describe the position of the mass at time q as $(q, w(q))$, its velocity is $(q', w'(q)q')$. In particular, it has scalar velocity $\sqrt{1 + w'(q)^2}q'$. By conservation of energy we obtain*

$$\frac{1}{2}(1 + w'(q)^2)(q')^2 + gw(q) = \textit{const.} = gy_1. \tag{1.1}$$

From a simple geometric/physical consideration, we see that we must have $q' > 0$. Thus q admits an inverse τ with $\tau' > 0$. In particular, (1.1) *can be re-arranged to give*

$$\tau' = \sqrt{\frac{1 + (w')^2}{2g(y_1 - w)}}.$$

The time the mass takes to traverse the graph is then

$$\int_{x_1}^{x_2} \tau' = \frac{1}{\sqrt{2g}} \int_{x_1}^{x_2} \sqrt{\frac{1 + (w')^2}{y_1 - w}}.$$

In light of these examples, we fix the class of problems we will study: Let $\Omega \subset \mathbb{R}^n$. Given $f \colon \Omega \times \mathbb{R}^n \times \mathbb{R}^{N\times n} \to \mathbb{R}$ seek to minimise

$$\mathscr{F}[w] = \int_\Omega f(x, w(x), \mathrm{D}w(x))\,\mathrm{d}x \tag{P}$$

amongst all suitable $w \colon \Omega \to \mathbb{R}^N$.

Remark 1.4 (Side conditions)**.** *Often the class of admissible functions w is restricted by imposing further conditions. Common examples are:*

- *It is common to impose Dirichlet boundary conditions in the form $w = \phi$ on $\partial\Omega$ for some function $\phi \colon \Omega \to \mathbb{R}^N$.*
- *We can impose holonomic/equality constraints, that is for some $g \colon \Omega \times \mathbb{R}^N \to \mathbb{R}$, impose $g(x, w(x)) = 0$. An example of this is $|w(x)| = 1$, i.e. w needs to map into the sphere. Thus an important subclass of this type of constraint is to constrain w to map into a manifold.*
- *We can relax to inequality constraints, $g(x, w(x)) \geq 0$ for some $g \colon \Omega\times\mathbb{R}^N \to \mathbb{R}$. This type of constraint includes obstacle problems.*
- *A further common type of constraints are integral constraints taking the form $\int_\Omega g(x, w(x))\,\mathrm{d}x = 0$ for $g \colon \Omega \times \mathbb{R}^N \to \mathbb{R}$. An example are Neumann boundary conditions, which are usually normalised through the constraint $\int_\Omega w(x)\,\mathrm{d}x = 0$.*

Perhaps the most important example of an integrand that fits the framework of (P) is the Dirichlet integral:

$$\mathscr{F}[w] = \int_\Omega |\mathrm{D}w|^2\,\mathrm{d}x \tag{1.2}$$

For simplicity, we consider this functional in the scalar case $N = 1$. The key insight into linking energy minimisation problems and PDE's is then contained in the following calculation: w is a minimiser of (1.2) if and only if for any $\varepsilon > 0$ and $\phi \colon \Omega \to \mathbb{R}$ sufficiently smooth with $\phi = 0$ on $\partial\Omega$,

$$\begin{aligned} 0 \geq \mathscr{F}[w] - \mathscr{F}[w + \varepsilon\phi] = &\int_\Omega |\mathrm{D}w|^2 - |\mathrm{D}w + \varepsilon\mathrm{D}\phi|^2\,\mathrm{d}x \\ =&2\varepsilon \int_\Omega \mathrm{D}\phi \cdot \mathrm{D}w - \varepsilon^2 \int_\Omega |\mathrm{D}\phi|^2\,\mathrm{d}x. \end{aligned}$$

As this holds for arbitrary ε, this statement is true if and only if for any such ϕ,

$$\int_\Omega \mathrm{D}w \cdot \mathrm{D}\phi\,\mathrm{d}x \leq 0.$$

Swapping $\phi \to -\phi$ even gives

$$0 = \int_\Omega \mathrm{D}w \cdot \mathrm{D}\phi \,\mathrm{d}x = \int_\Omega -\Delta w \phi \,\mathrm{d}x$$

As ϕ was arbitrary, we deduce that $-\Delta w = 0$. This argument in fact proves the following theorem:

Theorem 1.5. *Let $\mathscr{F}[w] = \int_\Omega |\mathrm{D}w|^2 \,\mathrm{d}x$. For $w \in C^2(\Omega)$,*

$$\mathscr{F}[w] \leq \mathscr{F}[v] \quad \forall v \in C^2(\Omega) \text{ such that } v = w \text{ on } \partial\Omega$$

if and only if

$$-\Delta w = 0.$$

We close this introduction by listing some more examples of functionals of the form (P).

Example 1.6. ***a)*** *We can replace the Euclidean norm in* (1.2) *by a different quadratic form:*

$$\mathscr{F}[w] = \int_\Omega A(\mathrm{D}w, \mathrm{D}w) \,\mathrm{d}x = \frac{1}{2} \sum_{l,m=1}^{N} \sum_{i,j=1}^{n} \int_\Omega a_{ij}^{lm} \partial_i w_l \partial_j w_m \,\mathrm{d}x.$$

If A is symmetric, this is related to the PDE

$$\sum_{l=1}^{N} \sum_{i,j=1}^{n} A_{ij}^{lm} \partial_j \partial_i u_l = 0 \qquad \text{for } m = 1, \dots N.$$

b) *We can include lower-order terms:*

$$f(x,y,z) = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) z_i z_j - \frac{1}{2} c(x) y^2 + g(x) y.$$

c) *We can replace the square in* (1.2) *with $p \geq 1$:*

$$\mathscr{F}[w] = \frac{1}{p} \int_\Omega |\mathrm{D}w|^p \,\mathrm{d}x.$$

This functional is linked to the p-Laplace equation

$$\operatorname{div}\left(|\mathrm{D}u|^{p-2} \mathrm{D}u\right) = 0.$$

2 Classical methods

In this section, we study (P) through studying the linked PDE, which is known as the Euler-Lagrange equation. Our first goal is to establish this link for a wide class of problems, where for simplicity we restrict to the case $N = 1$. In order to achieve this, convexity of the integrand $f(x,y,z)$ in the (y,z)-variables will play a crucial role. Hence, we begin by establishing some basic facts of convex analysis.

2.1 Convex analysis

The essential object in convex analysis are convex sets and convex functions:

Definition 2.1.
- *$\Omega \subset \mathbb{R}^n$ is convex if for all $x, y \in \Omega$, $\lambda \in [0,1]$, $\lambda x + (1-\lambda)y \in \Omega$.*
- *If $\Omega \subset \mathbb{R}^n$ is convex and $f\colon \Omega \to \mathbb{R}$, then f is convex if for all $x, y \in \Omega$, $\lambda \in [0,1]$, $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$. We say that f is strictly convex if the inequality is strict.*

A number of useful characterisations of convexity exist, which we collect in the following theorem:

Theorem 2.2. *Let $f\colon \mathbb{R}^n \to \mathbb{R}$, $f \in C^1(\mathbb{R}^n)$. Then the following are equivalent:*

(i) *f is convex,*

(ii) *for all $x, y \in \mathbb{R}^n$, $f(x) \geq f(x) + \langle \nabla f(y), x-y \rangle$,*

(iii) *for all $x, y \in \mathbb{R}^n$, $\langle f(x) - f(y), x - y \rangle \geq 0$.*

In case of strict convexity, all inequalities are strict. Moreover, if $f \in C^2(\mathbb{R}^n)$, then f is convex if and only if $\nabla^2 f$ is positive semi-definite (positive definite in the case of strict convexity).

Proof. Suppose f is convex. Then

$$\begin{aligned}\langle \nabla f(x), x-y \rangle = \lim_{\lambda \to 0} \frac{f(y+\lambda(x-y)) - f(y)}{\lambda} \leq &\lim_{\lambda \to 0} \frac{\lambda f(x) + (1-\lambda)f(y) - f(y)}{\lambda} \\ =& f(x) - f(y).\end{aligned}$$

Conversely, if (ii) holds, for any $x, y \in \Omega$ and $\lambda \in [0,1]$, denoting $x_\lambda = \lambda x + (1-\lambda)y$

$$\begin{aligned}f(x_\lambda) &\leq f(y) + \langle \nabla f(x_\lambda), y - x_\lambda \rangle \\ f(x_\lambda) &\leq f(x) + \langle \nabla f(x_\lambda), x - x_\lambda \rangle.\end{aligned}$$

Multiplying the first line with $(1-\lambda)$, the second line with λ and adding shows that f is convex.

Suppose (ii) holds. Then for all $x, y \in \Omega$,

$$\begin{aligned}f(x) \geq& f(y) + \langle \nabla f(y), x-y \rangle \\ f(y) \geq& f(x) + \langle \nabla f(x), y-x \rangle.\end{aligned}$$

Adding the two lines gives (iii). Conversely, assume (iii) holds. Then

$$\begin{aligned}f(y) =& f(x) + \int_0^1 \langle \nabla f(x + t(y-x)), y-x \rangle \,\mathrm{d}t \\ =& f(x) + \langle \nabla f(x), y-x \rangle + \int_0^1 \langle \nabla f(x+t(y-x)) - \nabla f(x), y-x \rangle \,\mathrm{d}t \\ \geq& f(x) + \langle \nabla f(x), y-x \rangle.\end{aligned}$$

To see the moreover part, assume first that f is convex and C^2. Then for any $x, s \in \mathbb{R}^n$ and $t > 0$,

$$\begin{aligned}0 \leq \frac{1}{t}\langle \nabla f(x+ts) - \nabla f(x), x+ts-x \rangle =& \langle \nabla f(x+ts) - \nabla f(x), s \rangle \\ =& \frac{1}{t} \int_0^t \langle \nabla^2 f(x+\lambda s)s, s \rangle.\end{aligned}$$

Letting $t \to 0$, we obtain $\langle \nabla^2 f(x)s, s\rangle \geq 0$, which gives the desired positive semi-definiteness of $\nabla^2 f$. Conversely, if $\nabla^2 f$ is positive semi-definite, we write

$$\begin{aligned} f(y) =& f(x) + \langle \nabla f(x), y-x\rangle + \int_0^1 \int_0^t \langle \nabla^2 f(x+\lambda(y-x))(y-x), y-x\rangle \,\mathrm{d}\lambda\,\mathrm{d}t \\ \geq& f(x) + \langle \nabla f(x), y-x\rangle, \end{aligned}$$

which is (ii). Hence f is convex. □

Finally we record an important integral inequality connected to convex functions:

Theorem 2.3 (Jensen's inequality)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and bounded, $u \in L^1(\Omega)$ and $f\colon \mathbb{R} \to \mathbb{R}$ be convex. Then*

$$f\left(\fint_\Omega u \,\mathrm{d}x\right) \leq \fint_\Omega f(u)\,\mathrm{d}x.$$

Proof. We give the proof in the case when $f \in C^1$. Note that then by Theorem 2.2 for $x \neq y$,

$$f'(x) \leq \frac{f(y)-f(x)}{y-x} \leq f\prime(y).$$

Setting $y = u$, $x = \fint_\Omega u$ and integrating gives

$$\fint_\Omega f(u) - f\left(\fint_\Omega u\right)\,\mathrm{d}x \geq f'\left(\fint_\Omega u\right)\fint_\Omega \left(\fint_\Omega u - u\right)\,\mathrm{d}x = 0.$$

□

Finally, we require the concept of the Fenchel conjugate f^* of a function f. This is sometimes also referred to as convex conjugate or Legendre conjugate of f.

Definition 2.4. *Let $f\colon \mathbb{R}^n \to \mathbb{R}$. Define the Fenchel conjugate f^* by setting for $\xi \in \mathbb{R}^n$,*

$$f^*(\xi) = \sup_{x\in\mathbb{R}^n} \langle x, \xi\rangle - f(x).$$

The bi-dual f^{**} of f is the Fenchel conjugate of f^*, that is for $x \in \mathbb{R}^n$,

$$f^{**}(x) = \sup_{\xi\in\mathbb{R}^n} \langle x, \xi\rangle - f^*(\xi).$$

We summarise the properties of f^* that will be important to us without giving a proof.

Theorem 2.5. *Let $f\colon \mathbb{R}^n \to \mathbb{R}$. Then the following statements hold:*

(i) *f^* is convex.*
(ii) *$f^{**} = f$ if and only if f is convex.*
(iii) *For any $x, \xi \in \mathbb{R}^n$, the Fenchel-Young inequality holds:*

$$\langle x, \xi\rangle \leq f(x) + f^*(\xi).$$

For $f \in C^1(\mathbb{R}^n)$, equality holds if and only if $\xi = f'(x)$.

2.2 The Euler-Lagrange equation

We now turn to establishing the equivalent of Theorem 1.5 in the general setting of (P). For simplicity, we will work in the case $n = 1$ and recall the problem we study:

$$\inf\{\mathscr{F}(u)\colon u \in X\} \tag{P1d}$$

where

$$\mathscr{F}(u) = \int_a^b f(x, u, u')\,\mathrm{d}x \qquad \text{and} \qquad X = \{u \in C^1([a,b])\colon u(a) = \alpha, u(b) = \beta\}.$$

The proof will proceed along the lines of that given for Theorem 1.5, the main difference being that we cannot expand the square. Instead, we will use convexity to obtain a similar estimate.

Theorem 2.6. *Let $f \in C^2([a,b] \times \mathbb{R} \times \mathbb{R})$, $f \equiv f(x,u,\xi)$ and suppose we have that $m = \inf\{\mathscr{F}[u]\colon u \in X\}$ is finite.*

(i) *If* (P1d) *admits a minimiser $\overline{u} \in X \cap C^2([a,b])$, then for $x \in (a,b)$*

$$\frac{\mathrm{d}}{\mathrm{d}x} f_\xi(x, \overline{u}, \overline{u}') = f_u(x, \overline{u}, \overline{u}'). \tag{EL}$$

(ii) *Conversely, if $\overline{u} \in X$ solves* (EL) *and $(u,\xi) \to f(x,u,\xi)$ is convex for every $x \in [a,b]$, then $\overline{u}$ is a minimiser of* (P1d).

(iii) *If $(u,\xi) \to f(x,u,\xi)$ is strictly convex for every $x \in [a,b]$, then the minimiser, if it exists, is unique.*

Remark 2.7. (i) *Note that we do not prove existence of a minimiser. This is a main drawback of the classical method and will motivate us to seek minimisers in weaker spaces of functions than C^2, where proving existence of minimisers is easier. The direct method in the calculus of variations arises from this approach.*

(ii) *In general, minimisers need not be C^2, as we will see on the problem sheet.*

(iii) *Without the convexity assumptions, solutions of* (EL) *may fail to be minimisers. Instead, they are usually referred to as stationary points or extremals.*

(iv) *Theorem 2.6 generalises to many more general set-ups, including vectorial problems where $n, N > 1$. We will see some examples on the problem sheet.*

Proof of Theorem 2.6. Suppose $\overline{u} \in X \cap C^2([a,b])$ minimises (P1d). Then for all $h \in \mathbb{R}$ and $v \in C^1([a,b])$ with $v(a) = v(b) = 0$,

$$\mathscr{F}(\overline{u}) \leq \mathscr{F}(\overline{u} + hv).$$

Introducing $\Phi(h) = \mathscr{F}(\overline{u} + hv) \in C^1(\mathbb{R})$, we may rewrite this as $\Phi(0) \leq \Phi(h)$ for all $h \in \mathbb{R}$. Consequently

$$\Phi'(0) = \frac{\mathrm{d}}{\mathrm{d}h}\Phi(\overline{u} + hv)\Big|_{h=0} = 0.$$

Calculating Φ' and integrating by parts, we find

$$\begin{aligned} 0 &= \int_a^b f_\xi(x, \overline{u}, \overline{u}')v' + f_u(x, \overline{u}, \overline{u}')v\,\mathrm{d}x \\ &= \int_a^b \left(-\frac{\mathrm{d}}{\mathrm{d}x} f_\xi(x, \overline{u}, \overline{u}') + f_u(x, \overline{u}, \overline{u}')\right) v = 0. \end{aligned} \tag{2.1}$$

Since v was arbitrary, we deduce (EL) holds. We will formally prove this statement, which is known as the fundamental lemma in the calculus of variations in Lemma 2.8.

Suppose now $\overline{u} \in X$ solves (EL). Using the joint convexity of f in its second and third variable, by Theorem 2.2, we have for any $u \in X$,

$$f(x,u,u') \geq f(x,\overline{u},\overline{u}') + f_u(x,\overline{u},\overline{u}')(u-\overline{u}) + f_\xi(x,\overline{u},\overline{u}')(u'-\overline{u}').$$

Integrating, after an integration by parts, we deduce

$$\begin{aligned}\mathscr{F}(u) \geq & \mathscr{F}(\overline{u}) + \int_a^b \left(f_u(x,\overline{u},\overline{u}') - \frac{\mathrm{d}}{\mathrm{d}x} f_\xi(x,\overline{u},\overline{u}') \right)(u-\overline{u})\,\mathrm{d}x \\ = & \mathscr{F}(\overline{u}).\end{aligned}$$

Thus $\overline{u}$ solves (P1d).

Finally, suppose for any $x \in [a,b]$, $(u,\xi) \to f(x,u,\xi)$ is strictly convex. Assume $u, v \in X$ solve (P1d). Set $w = \frac{u+v}{2}$. Note $w \in X$ and by strict convexity, for any $x \in [a,b]$,

$$\frac{f(x,u,u') + f(x,v,v')}{2} \geq f(x,w,w'),$$

with equality if and only if $u(x) = v(x)$. In particular,

$$m = \frac{\mathscr{F}(u) + \mathscr{F}(v)}{2} \geq \mathscr{F}(w) \geq m.$$

In other words,

$$\int_a^b \frac{f(x,u,u') - f(x,v,v')}{2} - f(x,w,w')\,\mathrm{d}x = 0.$$

Thus $u = v$. □

It remains to fully justify that (2.1) implies that (EL) holds for every $x \in [a,b]$. This will follow from the following theorem.

Theorem 2.8 (Fundamental lemma of the calculus of variations)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and $u \in L^1_{\mathrm{loc}}(\Omega)$ such that*

$$\int_\Omega u(x)\phi(x)\,\mathrm{d}x = 0 \quad \forall \phi \in C_0^\infty(\Omega).$$

Then $u = 0$ almost everywhere in Ω.

Proof. We give the proof in the case $u \in L^2(\Omega)$. The general case can be done using a technical approximation argument. Let $\varepsilon > 0$. There exists $\psi \in C_0^\infty(\Omega)$ such that

$$\|u - \psi\|_{L^2(\Omega)} \leq \varepsilon.$$

In particular,

$$\|u\|^2_{L^2(\Omega)} = \int_\Omega u^2\,\mathrm{d}x = \int_\Omega u(u-\psi)\,\mathrm{d}x \leq \|u\|_{L^2(\Omega)} \|u-\psi\|_{L^2(\Omega)}.$$

We deduce $\|u\|_{L^2(\Omega)} \leq \varepsilon$. As ε was arbitrary, $\|u\|_{L^2(\Omega)} = 0$ and hence $u = 0$ almost everywhere. □

2.3 A 2nd form of the Euler-Lagrange equation

Occasionally, a slightly different form of the Euler-Lagrange equation is useful.

Theorem 2.9. *Let $f \in C^2([a,b] \times \mathbb{R} \times \mathbb{R})$, $f \equiv f(x,u,\xi)$ and*

$$\inf_{u \in X} \left\{ \mathscr{F}(u) = \int_a^b f(x,u,u')\,\mathrm{d}x \right\} = m \tag{P}$$

be finite, where $X = \{u \in C^1([a,b])\colon u(a) = \alpha,\ u(b) = \beta\}$. Let $u \in X \cap C^2([a,b])$ be a minimiser of (P)*, then for every $x \in [a,b]$ it holds that*

$$\frac{\mathrm{d}}{\mathrm{d}x}\big(f(x,u,u') - u' f_\xi(x,u,u')\big) = f_x(x,u,u').$$

Proof. Using Theorem 2.6 the proof is easy. In fact, for any $u \in C^2([a,b])$,

$$\frac{\mathrm{d}}{\mathrm{d}x}\big(f(x,u,u') - u' f_\xi(x,u,u')\big) = f_x(u,u,u') + u'\left(f_u(x,u,u') - \frac{\mathrm{d}}{\mathrm{d}x} f_\xi(x,u,u')\right).$$

If $u \in X \cap C^2([a,b])$ is a minimiser of (P), by Theorem 2.6, the term in parentheses on the right-hand side vanishes, giving the claimed result. □

There is a second proof of Theorem 2.9, that is of interest due to the technique of its proof, which turns out to be useful in a number of circumstances. While in the proof of Theorem 2.6, we used variations of the form $u + \varepsilon\phi$, that is we used variations of the dependent variables, we will prove Theorem 2.9 using variations of the independent variables. This approach is also known under the name inner variations. Thus, given $\varepsilon \in \mathbb{R}$, $\phi \in C_0^\infty(a,b)$ and $\lambda = (2\|\phi'\|_{L^\infty})^{-1}$, we set

$$\xi(x,\varepsilon) = x + \varepsilon\lambda\phi(x) = y. \tag{2.2}$$

For $\varepsilon \le 1$, $\xi(\cdot,\varepsilon)\colon [a,b] \to [a,b]$ is a diffeomorphism with $\xi(a,\varepsilon) = a$, $\xi(b,\varepsilon) = b$ and $\xi_x(x,\varepsilon) > 0$. We may introduce its inverse $\eta(\cdot,\varepsilon)\colon [a,b] \to [a,b]$ by demanding

$$\xi(\eta(y,\varepsilon),\varepsilon) = y. \tag{2.3}$$

We want to study the competitor given by

$$u^\varepsilon(x) = u(\xi(x,\varepsilon)).$$

In order to carry out the argument, we will see that we require estimates on $\eta_y(y,\varepsilon)$ and $\eta_\varepsilon(y,\varepsilon)$. However due to (2.3),

$$\begin{gathered} \xi_x(\eta(y,\varepsilon),\varepsilon)\eta_y(y,\varepsilon) = 1 \\ \xi_x(\eta(y,\varepsilon),\varepsilon)\eta_\varepsilon(y,\varepsilon) + \xi_\varepsilon(\eta(y,\varepsilon),\varepsilon) = 0. \end{gathered} \tag{2.4}$$

Next consider (2.2) and use a Taylor-expansion to deduce,

$$\begin{aligned} \eta_y(y,\varepsilon) &= 1 - \varepsilon\lambda\phi'(y) + O(\varepsilon^2) \\ \eta_\varepsilon(y,\varepsilon) &= -\lambda\phi(y) + O(\varepsilon). \end{aligned} \tag{2.5}$$

Using the change of variables $y = \xi(x,\varepsilon)$ and the first line of (2.4), we thus calculate

$$\mathscr{F}(u^\varepsilon) = \int_a^b f(x, u(\xi(x,\varepsilon)), u'(\xi(x,\varepsilon))\xi_x(x,\varepsilon))\,\mathrm{d}x$$

$$
\begin{aligned}
&= \int_a^b f(\eta(y,\varepsilon), u(y), u'(y)/\eta_y(y,\varepsilon))\eta_y(y,\varepsilon)\,\mathrm{d}y \\
&:= \int_a^b g(\varepsilon)\,\mathrm{d}y.
\end{aligned}
$$

Since u is a minimiser of (P) and $u^\varepsilon \in X$, $\mathscr{F}(u^\varepsilon) \geq \mathscr{F}(u)$ and hence

$$
0 = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathscr{F}(u^\varepsilon)\Big|_{\varepsilon=0} = \int_a^b g'(0)\,\mathrm{d}x.
$$

Using (2.5) we find

$$
g'(0) = \lambda\left(-f_x\phi + (u' f_\xi - f)\phi'\right),
$$

from which we deduce after an integration by parts,

$$
0 = \lambda \int_a^b \left(-f_x(x,u,u') + \frac{\mathrm{d}}{\mathrm{d}x}\left(-u' f_\xi(x,u,u') + f(x,u,u')\right)\right)\phi(x)\,\mathrm{d}x.
$$

Appealing to Theorem 2.8, this gives the desired identity.

2.4 The Hamilton-Jacobi equation

Solutions of the Euler-Lagrange equation can be linked to solutions of a first order PDE which is known as Hamilton-Jacobi equation. This connection is especially useful for some energies arising in classical mechanics. We begin by studying a different energy functional, involving the Hamiltonian H of the energy f. The Hamiltonian is nothing but the convex conjugate of f with respect to the ξ-variable, i.e.

$$
H(x,u,v) = \sup_{\xi\in\mathbb{R}} \langle v, \xi\rangle - f(x,u,\xi).
$$

Consider the energy functional

$$
J(u,v) = \int_a^b u'(x)v(x) - H(x,u(x),v(x))\,\mathrm{d}x.
$$

Note that formally the Euler-Lagrange equation for J is the system

$$
\begin{cases} u' = H_v(x,u,v) \\ v' = -H_u(x,u,v). \end{cases} \tag{H}
$$

We want to link solutions of (H) to solutions of (EL). We begin our analysis by studying the regularity of H and finding convenient expressions for H_v and H_u.

Lemma 2.10. *Let $f \in C^2([a,b]\times\mathbb{R}\times\mathbb{R})$, $f \equiv f(x,u,\xi)$ such that $f_{\xi\xi}(x,u,\xi) > 0$ for every $(x,u,\xi) \in [a,b]\times\mathbb{R}\times\mathbb{R}$ and $\lim_{|\xi|\to\infty} \frac{f(x,u,\xi)}{|\xi|} = +\infty$ for every $(x,u) \in [a,b]\times\mathbb{R}$. Then $H \in C^2([a,b]\times\mathbb{R}\times\mathbb{R})$ and*

$$
\begin{aligned}
H_x(x,u,v) &= -f_x(x,u,H_v(x,u,v)) \\
H_u(x,u,v) &= -f_u(x,u,H_v(x,u,v)) \\
H(x,u,v) &= vH_v(x,u,v) - f(x,u,H_v(x,u,v)) \\
v = f_\xi(x,u,\xi) &\Leftrightarrow \xi = H_v(x,u,v).
\end{aligned}
$$

Proof. **Step 1: *H* is a maximum** Fix $(x,u) \in [a,b] \times \mathbb{R}$. Due to the super-linear growth of $f(x,u,\cdot)$ at infinity and the definition of H, we find $\xi = \xi(x,u,v)$ such that

$$\begin{cases} H(x,u,v) = \langle v, \xi \rangle - f(x,u,\xi) \\ v = f_\xi(x,u,\xi). \end{cases} \tag{2.6}$$

Step 2: *H* is continuous Let $(x,u,v),(x',u',v') \in [a,b] \times \mathbb{R} \times \mathbb{R}$. Then for some $\xi = \xi(u,v)$ using the definition of H,

$$\begin{gathered} H(x,u,v) = \langle v, \xi \rangle - f(x,u,\xi) \\ H(x',u',v') \geq \langle v', \xi \rangle - f(x',u',\xi), \end{gathered}$$

so that subtracting the two lines, we find

$$H(x,u,v) - H(x',u',v') \leq \langle v - v', \xi \rangle + f(x',u',\xi) - f(x,u,\xi).$$

Reversing the roles of primed and unprimed variables, we can prove

$$H(x,u,v) - H(x',u',v') \geq \langle v - v', \xi \rangle + f(x',u',\xi) - f(x,u,\xi).$$

In particular

$$|H(x,u,v) - H(x',u',v')| \leq |v - v'||\xi| + |f(x',u',\xi) - f(x,u,\xi)|.$$

In particular, since f is continuous in the (x,u)-variables, we deduce that H is continuous.
Step 3: $\boldsymbol{\xi}$ is $\boldsymbol{C^1}$ Considering the second line of (2.6) and noting that $f \in C^2$ with $f_{\xi\xi}(x,u,\xi) > 0$, we may apply the inverse function theorem to deduce that $\xi \in C^1([a,b] \times \mathbb{R} \times \mathbb{R})$.
Step 4: conclusion From our work so far, we conclude that the functions

$$(x,u,v) \to \xi(x,u,v),\ f_x(x,u,\xi(x,u,v)),\ f_u(x,u,\xi(x,u,v))$$

are C^1. Combining this with (2.6), the first three claimed identities follow immediately. Indeed

$$\begin{gathered} H_x = v\xi_x - f_x - f_\xi \xi_x = (v - f_\xi)\xi_x - f_x = -f_x \\ H_u = v\xi_u - f_u - f_\xi \xi_u = (v - f_\xi)\xi_u - f_u = -f_u \\ H_v = \xi + v\xi_v - f_\xi \xi_v = (v - f_\xi)\xi_v + \xi = \xi. \end{gathered}$$

The last claim of the Lemma follows from Theorem 2.5. □

Using Lemma 2.10, we can relate solutions of (H) and (EL).

Theorem 2.11. *Let f and H be as in Lemma 2.10. Let $(u,v) \in C^2([a,b]) \times C^2([a,b])$ satisfy for every $x \in [a,b]$,*

$$\begin{cases} u' = H_v(x,u,v) \\ v' = -H_u(x,u,v). \end{cases} \tag{H}$$

Then u satisfies the Euler-Lagrange equation

$$\frac{\partial}{\partial x} f_\xi(x,u,u') = f_u(x,u,u') \tag{EL}$$

for all $x \in [a,b]$. Conversely, if $u \in C^2([a,b])$ satisfies (EL) *for all $x \in [a,b]$, then with*

$$v = f_\xi(x,u,u'),$$

the couple (u,v) satisfies (H) *for every $x \in [a,b]$.*

Proof. Suppose (u,v) satisfy (EL). Using Lemma 2.10, we find

$$\begin{aligned} u' = H_v(x,u,v) \quad &\Leftrightarrow \quad v = f_\xi(x,u,u') \\ v' = -H_u(x,u,v) &= f_u(x,u,u'). \end{aligned}$$

In particular,

$$\frac{\mathrm{d}}{\mathrm{d}x} f_\xi(x,u,u') = v' = f_u(x,u,u'),$$

that is (EL) holds. Conversely, if u satisfies (EL) and we set $v = f_\xi(x,u,u')$, by Lemma 2.10, we have

$$v = f_\xi(u,u,u') \quad \Leftrightarrow \quad u' = H_v(x,u,v),$$

which is the first equation in (H). Moreover, we calculate, since u satisfies (EL),

$$v' = \frac{\mathrm{d}}{\mathrm{d}x} f_\xi(x,u,u') = f_u(x,u,u').$$

Thus (u,v) solves (H). $\square$

We consider two examples inspired by classical mechanics that make the relation to the Hamiltonian as used in physics clearer.

Example 2.12. *We consider an example inspired by a particle in free fall. Let $m > 0$, $g \in C^1([a,b])$ and $f(x,u,\xi) = \frac{m}{2}\xi^2 - g(x)u$. The Euler-Lagrange equation associated to the integral*

$$\int_a^b f(x,u,u')\,\mathrm{d}x$$

is

$$mu'' = -g$$

for $x \in (a,b)$. Moreover, by direct calculation,

$$H(x,u,v) = \frac{v^2}{2m} + g(x)u.$$

Note that in this example, along trajectories, that is when $v = f_\xi(x,u,u') = u'$, the Hamiltonian corresponds to the total energy of the system as the sum of kinetic and potential energy. Indeed, the Hamiltonian formulation is given by

$$\begin{cases} u' = \frac{v}{m} \\ v' = -g, \end{cases}$$

so that

$$H(x,u,v) = \frac{m}{2}(u')^2 + g(x)u.$$

Example 2.13. *We consider a system with N particles of masses $m_i > 0$ and whose positions at time t are described by $u_i(t) = (x_i(t), y_i(t), z_i(t))$. The kinetic energy of the system is then given by*

$$E_{\mathrm{kin}}(u') = \frac{1}{2}\sum_{i=1}^N m_i |u_i'|^2.$$

Suppose $U = U(t,u)$ is the potential energy of the system. Consider the energy (=Lagrangian)

$$L(t,u,u') = T(u') - U(t,u).$$

A similar calculation to that of Example 2.12 shows

$$H(x,u,v) = \frac{1}{2}\sum_{i=1}^{N} \frac{1}{m_i}|v_i|^2 + U(t,u).$$

In particular, the Euler-Lagrange equation for the system is

$$\sum_{i=1}^{N} m_i \partial_i u_i' = U_u(t,u),$$

while the Hamilton-Jacobi system is given by

$$\begin{cases} u_i' = \frac{1}{m_i} v_i \\ v_i' = -U_{u_i}(t,u). \end{cases}$$

for $i = 1,\dots N$. Note that we are using here an extension of the arguments of this section to vector-valued functions. This is however not difficult to achieve. Along trajectories, where $v = L_\xi(t,u,u')$, we thus find

$$H(x,u,v) = \frac{1}{2}\sum_{i=1}^{N} m_i|u_i|^2 + U(t,u).$$

As in the first example, we recognize this expression as the total energy of the system.

We next want to relate solutions of (H) to solutions of the Hamilton-Jacobi equation, which is a first order PDE. Our main result in this direction will be

Theorem 2.14. *Let $H \in C^1([a,b]\times\mathbb{R}\times\mathbb{R})$, $H \equiv H(x,u,v)$. Assume there exists a solution $S \in C^2([a,b]\times\mathbb{R})$, $S \equiv S(x,u)$ of the Hamilton-Jacobi equation*

$$S_x + H(x,u,S_u) = 0 \tag{2.7}$$

for all $(x,u) \in [a,b]\times\mathbb{R}$. Assume in addition that there is $u \in C^1([a,b])$ which solves

$$u'(x) = H_v(x,u,S_u(x,u)) \tag{2.8}$$

for all $x \in (a,b)$. Set $v = S_u(x,u)$. Then $(u,v) \in C^1([a,b]) \times C^1([a,b])$ is a solution of

$$\begin{cases} u' = H_v(x,u,v) \\ v' = -H_u(x,u,v). \end{cases}$$

Moreover, if there is a one-parameter family of solutions

$$S \equiv S(x,u,\alpha) \in C^2([a,b]\times\mathbb{R}\times\mathbb{R})$$

solving (2.7) for every $(x,u,\alpha) \in [a,b]\times\mathbb{R}\times\mathbb{R}$, then any solution of (2.8) satisfies

$$\frac{\mathrm{d}}{\mathrm{d}x} S_\alpha(x,u,\alpha) = 0$$

for all $(x,\alpha) \in [a,b] \in \mathbb{R}$.

Remark 2.15. *In general, it is very difficult to solve* (2.7) *and the Hamilton-Jacobi equation has been the subject of intense study.*

In the case where H is independent of x, every solution $S^*(u,\alpha)$ of

$$H(u, S_u^*) = \alpha$$

for every $(u,\alpha) \in \mathbb{R} \times \mathbb{R}$ gives a solution of (2.7) *by defining*

$$S(x,u,\alpha) = S^*(u,\alpha) - \alpha x.$$

Before proceeding with the proof, we consider a simple example.

Example 2.16. *Let $g \in C^1(\mathbb{R})$ with $g(u) \geq g_0 > 0$. Consider the energy functional*

$$f(u,\xi) = \frac{1}{2}\xi^2 + g(u),$$

with the corresponding Hamiltonian

$$H(u,v) = \frac{1}{2}v^2 - g(u).$$

The Hamilton-Jacobi equation is then given by

$$S_x + \frac{1}{2}S_u^2 - g(u) = 0.$$

Considering Remark 2.15, it suffices to consider

$$\frac{1}{2}(S_u^*)^2 = g(u).$$

In particular, a solution of the Hamilton-Jacobi equation is given by

$$S = S(u) = \int_0^u \sqrt{2g(s)}\,\mathrm{d}s.$$

We further need to solve (2.8)*, which now reads*

$$u' = H_v(u, S_u(u)) = S_u(u) = \sqrt{2g(u)}.$$

A solution of this equation is implicitly given by

$$x = \int_{u(0)}^{u(x)} \frac{\mathrm{d}s}{\sqrt{2g(s)}}.$$

Thus, setting $v = S_u(u)$, we have found a solution of the Hamiltonian system

$$\begin{aligned} u' &= H_v(u,v) = v \\ v' &= -H_u(u,v) = g'(u). \end{aligned}$$

Note in particular, that

$$u'' = v' = g'(u),$$

which is the Euler-Lagrange equation associated to the integrand f.

Proof of Theorem 2.14. Differentiating the definition of v, we see that for all $x \in [a, b]$,

$$v' = S_{xu}(x, u) + S_{uu}(x, u)u'.$$

Differentiating the Hamilton-Jacobi equation with respect to u, it holds that for every $(x, u) \in [a, b] \times \mathbb{R}$,

$$S_{xu} + H_u(x, u, S_u(x, u)) + H_v(x, u, S_u(x, u))S_{uu}(x, u) = 0.$$

In particular, we deduce, using also (2.8),

$$\begin{aligned} v' =& - H_u(x, u, S_u(x, u)) - H_v(x, u, S_u(x, u))S_{uu}(x, u) + S_{uu}(x, u)u' \\ =& - H_u(x, u, S_u(x, u)) = -H_u(x, u, v). \end{aligned}$$

Thus the couple (u, v) does indeed solve the Hamiltonian system.

Concerning the moreover part, we note

$$\begin{aligned} 0 =& \frac{\mathrm{d}}{\mathrm{d}\alpha} \big(S_x(x, u, \alpha) + H(x, u, S_u(x, u, \alpha)) \big) \\ =& S_{x\alpha}(x, u, \alpha) + H_v(x, u, S_u(x, u, \alpha))S_{u\alpha}(x, u, \alpha) = 0. \end{aligned}$$

In particular, if u solves (2.8),

$$0 = S_{x\alpha}(x, u, \alpha) + u' S_{u\alpha}(x, u, \alpha) = \frac{\mathrm{d}}{\mathrm{d}x} S_\alpha(x, u, \alpha)).$$

This was exactly our claim. □

In fact Theorem 2.14 admits a converse, which we prove to close this section.

Theorem 2.17. *Let $H \in C^1([a, b] \times \mathbb{R} \times \mathbb{R})$, $S \in C^2([a, b] \times \mathbb{R} \times \mathbb{R})$, $S \equiv S(x, u, \alpha)$. Suppose S solves* (2.7) *for every $(x, u, \alpha) \in [a, b] \times \mathbb{R} \times \mathbb{R}$ with*

$$S_{u\alpha}(x, u, \alpha) \neq 0$$

for all $(x, u, \alpha) \in [a, b] \times \mathbb{R} \times \mathbb{R}$. If u satisfies for every $(x, \alpha) \in [a, b] \times \mathbb{R}$,

$$\frac{\mathrm{d}}{\mathrm{d}x} S_\alpha(x, u, \alpha) = 0,$$

then for all $(x, \alpha) \in [a, b] \times \mathbb{R}$,

$$u' = H_v(x, u, S_u(x, u, \alpha)).$$

In particular, if $v = S_u(x, u, \alpha)$, then $(u, v) \in C^1([a, b]) \times C^1([a, b])$ is a solution of the Hamiltonian system (H)*.*

Proof. We explicitly calculate

$$0 = \frac{\mathrm{d}}{\mathrm{d}x} S_\alpha(x, u, \alpha) = S_{x\alpha}(x, u, \alpha) + S_{u\alpha}(x, u, \alpha)u'.$$

Using the Hamilton-Jacobi equation (2.7),

$$\begin{aligned} 0 =& \frac{\mathrm{d}}{\mathrm{d}\alpha} \big(S_x(x, u, \alpha) + H(x, u, S_u(x, u, \alpha)) \big) \\ =& S_{x\alpha}(x, u, \alpha) + H_v(x, u, S_u(x, u, \alpha))S_{u\alpha}(x, u, \alpha). \end{aligned}$$

Since $S_{u\alpha}(x, u, \alpha) \neq 0$, we deduce

$$u' = H_v(x, u, S_u(x, u, \alpha))$$

for all $(x, \alpha) \in [a, b] \times \mathbb{R}$. It remains to prove $v' = -H_u$. Differentiating the definition of v,

$$\begin{aligned} v' =& S_{xu}(x, u, \alpha) + S_{uu}(x, u, \alpha) u' \\ =& S_{xu}(x, u, \alpha) + S_{uu}(x, u, \alpha) H_v(x, u, S_u(x, u, \alpha)). \end{aligned}$$

Differentiating the Hamilton-Jacobi equation with respect to u, we find

$$\begin{aligned} 0 =& \frac{\mathrm{d}}{\mathrm{d}u} \left(S_x(x, u, \alpha) + H(x, u, S_u(x, u, \alpha)) \right) \\ =& S_{xu}(x, u, \alpha) + H_u(x, u, S_u(x, u, \alpha)) + H_v(x, u, S_u(x, u, \alpha)) S_{uu}(x, u, \alpha). \end{aligned}$$

In particular, we infer

$$v' = -H_u(x, u, S_u(x, u, \alpha)) = -H_u(x, u, v).$$

□

3 Noether's theorem

Many physical systems have symmetry properties. For example, in classical mechanics we expect our systems to be invariant under translations and rotations of the coordinate system. Moreover, we generally expect each symmetry of a system to give rise to a conserved quantity. This raises the question of whether there is a systematic way of exploiting this relationship. In other words, given an energy with a symmetry property, can we derive a conservation law that the minimiser satisfies? Noether's theorem gives a way of ensuring this.

Example 3.1 (Symmetries of the Dirichlet energy)**.** *Let $\Omega \subset \mathbb{R}^n$. Consider the minimisation problem*

$$\min_{u \in W^{1,2}(\Omega)} \frac{1}{2} \int_\Omega |\mathrm{D}u|^2 \, \mathrm{d}x$$

Suppose $\overline{u} \in \mathrm{W}^{1,2}(\Omega) \cap \mathrm{W}^{2,2}_{\mathrm{loc}}(\Omega)$ is a minimiser. Note that the Dirichlet energy is invariant under translations: If $\tau \in \mathbb{R}$ and $k \in \{1, \ldots, n\}$ set

$$x_\tau = x + \tau e_k, \quad u_\tau(x) = u(x + \tau e_k),$$

where e_k is the k-th unit vector. If $D \subset \Omega$ is such that $D_\tau = D + \tau e_k \subset \Omega$ for all τ sufficiently small, then for such τ,

$$\frac{1}{2} \int_D |\mathrm{D}u|^2 \, \mathrm{d}x = \frac{1}{2} \int_{D_\tau} |\mathrm{D}u_\tau|^2 \, \mathrm{d}x_\tau.$$

A further symmetry is expressed by setting for $\lambda \geq 0$,

$$x_\lambda = \lambda x, \quad u_\lambda = \lambda^{\frac{n-2}{2}} u(\lambda x), \quad D_\lambda = \lambda D.$$

Then if $D \subset \Omega$ such that $D_\lambda \subset \Omega$ for all λ sufficiently small, then for such λ,

$$\frac{1}{2} \int_D |\mathrm{D}u|^2 \, \mathrm{d}x = \frac{1}{2} \int_{D_\lambda} |\mathrm{D}u_\lambda|^2 \, \mathrm{d}x_\lambda$$

Note that, if we prefer this symmetry to be parametrised by $\tau \in \mathbb{R}$, we can set $\lambda = e^\tau$.

The general set-up we consider in this section is the following: Consider the energy

$$\mathscr{F}[u] = \int_\Omega f(x, u, \mathrm{D}u)\,\mathrm{d}x, \tag{3.1}$$

where $f \equiv f(x, y, z)\colon \Omega \times \mathbb{R}^n \times \mathbb{R}^{m\times n} \to \mathbb{R}$, with $f(x, \cdot, \cdot)$ being twice continuously differentiably for every $x \in \Omega$. Assume moreover that

$$|\partial_y F(x, y, z)| + |\partial_z F(x, y, z)| \le C(1 + |y|^2 + |z|^2) \tag{3.2}$$

for some $C > 0$ and all $(x, y, z) \in \Omega \times \mathbb{R}^n \times \mathbb{R}^{m\times n}$.

We first note, that for minimisers $\overline{u} \in \mathrm{W}^{1,2}(\Omega) \cap \mathrm{W}^{2,2}_{\mathrm{loc}}(\Omega)$ the Euler-Lagrange equation holds pointwise almost-everywhere:

Proposition 3.2. *If $\overline{u} \in \mathrm{W}^{1,2}(\Omega) \cap \mathrm{W}^{2,2}_{\mathrm{loc}}(\Omega)$ minimises* (3.1)*, then for almost every $x \in \Omega$,*

$$-\mathrm{div}\,\partial_z f(x, \overline{u}, \mathrm{D}\overline{u}) + \partial_y f(x, \overline{u}, \mathrm{D}\overline{u}) = 0.$$

Proof. Fix a domain $\omega \Subset \Omega$. Let $\phi \in C^\infty(\Omega)$ with $\phi = 0$ in $\Omega \setminus \omega$ and $\varepsilon > 0$. Then since $\overline{u}$ minimizes (3.1),

$$0 \ge \lim_{\varepsilon\to 0} \varepsilon^{-1}\left(\mathscr{F}[u] - \mathscr{F}[u + \varepsilon\phi]\right) = \lim_{\varepsilon\to 0} \int_\omega f(x, \overline{u}, \mathrm{D}\overline{u}) - f(x, \overline{u} + \varepsilon\phi, \mathrm{D}\overline{u} + \varepsilon \mathrm{D}\phi)\,\mathrm{d}x.$$

Due to (3.2) and the mean-value theorem, the dominated convergence theorem applies and we deduce using the divergence theorem,

$$\begin{aligned} 0 &\ge \int_\omega \langle \partial_z f(x, \overline{u}, \mathrm{D}\overline{u}), \mathrm{D}\phi\rangle + \partial_y f(x, \overline{u}, \mathrm{D}\overline{u})\phi\,\mathrm{d}x \\ &= \int_\omega \left(-\mathrm{div}\,\partial_z f(x, \overline{u}, \mathrm{D}\overline{u}) + \partial_y f(x, \overline{u}, \mathrm{D}\overline{u})\right)\phi\,\mathrm{d}x. \end{aligned}$$

Applying Lemma 2.8 and since $\omega \Subset \Omega$ was arbitrary, we conclude the proof. $\square$

We next need to generalise what we mean by the energy having a symmetry. For this purpose, let $g\colon \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$, $H\colon \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ with

$$g(x, 0) = x, \quad H(x, 0) = \overline{u}(x).$$

Suppose $g(x, \cdot)$, $H(x, \cdot)$ are continuously differentiable for almost every $x \in \Omega$. We write for any domain $D \subset \Omega$,

$$x_\tau = g(x, \tau), \quad u_\tau(x) = H(x, \tau), \quad D_\tau = g(D, \tau)$$

One should think of (g, H) as a sort of homotopy.

Definition 3.3. *We say $\mathscr{F}$ is invariant under (g, H) if*

$$\int_D f(x, u_\tau, \mathrm{D}u_\tau)\,\mathrm{d}x = \int_{D_\tau} f(x', \overline{u}(x'), \mathrm{D}\overline{u}(x'))\,\mathrm{d}x' \tag{3.3}$$

for any Lipschitz domains (that is $\Omega \subset \mathbb{R}^n$ open, bounded, connected such that $\partial\Omega$ is the union of a finite number of Lipschitz curves) $D \Subset \Omega$ with $D_\tau \Subset \Omega$.

Theorem 3.4 (Noether's theorem)**.** *Let $f\colon \Omega\times\mathbb{R}^n\times\mathbb{R}^{m\times n}\to\mathbb{R}$. Suppose $f(x,\cdot,\cdot)$ is C^2 for almost every $x\in\Omega$ and there exists $C>0$ such that*

$$|\partial_y f(x,y,z)| + |\partial_z f(x,y,z)| \leq C(1+|y|^2+|z|^2)$$

for all $(x,y,z)\in\Omega\times\mathbb{R}^n\times\mathbb{R}^{m\times n}$. Assume $\mathscr{F}$ is invariant under (g,H) and that there is $h\in \mathrm{L}^2(\Omega)$ such that for almost every $x\in\Omega$ and every $\tau\in\mathbb{R}$,

$$|\partial_t H(x,\tau)| + |\partial_\tau g(x,\tau)| \leq h(x).$$

Then for any minimiser or critial point $\overline{u}\in \mathrm{W}^{1,2}(\Omega)\cap \mathrm{W}^{2,2}_{\mathrm{loc}}(\Omega)$ the conservation law

$$\operatorname{div}\mu^T\partial_z f(x,\overline{u},\mathrm{D}\overline{u}) - \nu f(x,\overline{u},\mathrm{D}\overline{u}) = 0$$

holds for almost every $x\in\Omega$. Here $\mu(x)=\partial_\tau H(x,0)$ and $\nu(x)=\partial_\tau g(x,0)$ are the Noether multipliers.

Proof. Fix a Lipschitz domain $D\Subset\Omega$ such that for sufficiently small τ, $D_\tau\Subset\Omega$. Differentiate (3.3) with respect to τ, noting that integration and differentiation may be interchanged due to our growth assumptions. This gives

$$\int_D \partial_y f(x,u_\tau,\mathrm{D}u_\tau)\partial_\tau u_\tau + \partial_z f(x,u_\tau,\mathrm{D}u_\tau)\mathrm{D}\partial_\tau u_\tau\,\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}\tau}\int_{D_\tau} f(x',\overline{u},\mathrm{D}\overline{u})\,\mathrm{d}x'.$$

Due to Reynold's transport theorem, or alternatively by checking the following identity using a change of coordinates,

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\int_{D_\tau} f(x',\overline{u},\mathrm{D}\overline{u})\,\mathrm{d}x' = -\int_{\partial D_\tau} f(x',\overline{u},\mathrm{D}\overline{u})\partial_\tau g(x,\tau)\cdot n\,\mathrm{d}\mathscr{H}^{n-1}.$$

Using this identity, evaluated at $\tau=0$ and the divergence theorem, we obtain

$$\begin{aligned}
&\int_D (-\operatorname{div}\partial_z f(x,\overline{u},\mathrm{D}\overline{u}) + \partial_y f(x,\overline{u},\mathrm{D}\overline{u}))\mu\,\mathrm{d}x\\
&= \int_{\partial D} (\mu^T\partial_z f(x,\overline{u},\mathrm{D}\overline{u}) - \nu f(x,\overline{u},\mathrm{D}\overline{u}))\cdot n\,\mathrm{d}\mathscr{H}^{n-1}\\
&= -\int_D \operatorname{div}(\mu^T\partial_z f(x,\overline{u},\mathrm{D}\overline{u}) - \nu f(x,\overline{u},\mathrm{D}\overline{u}))\,\mathrm{d}x.
\end{aligned}$$

Due to Proposition 3.2, the left-hand side vanishes and we deduce that almost everywhere in D,

$$\operatorname{div}(\mu^T\partial_z f(x,\overline{u},\mathrm{D}\overline{u}) - \nu f(x,\overline{u},\mathrm{D}\overline{u}) = 0.$$

As D was arbitrary, this concludes the proof. $\square$

Noting that if f does not explicitly depend on x, any minimiser is invariant under translations, we obtain the following corollary:

Corollary 3.5. *Under the assumptions of Theorem 3.4, if $f\equiv f(y,z)\colon\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}$, then any minimiser or critical point $\overline{u}\in \mathrm{W}^{1,2}(\Omega)\cap\mathrm{W}^{2,2}_{\mathrm{loc}}(\Omega)$ satisfies for almost every $x\in\Omega$,*

$$\sum_{i=1}^d \partial_i\left(\partial_k u)\cdot\partial_{z_k} f(\overline{u},\mathrm{D}\overline{u}) - \delta_{ik} f(\overline{u},\mathrm{D}\overline{u})\right) = 0,$$

for $k=1,\dots,n$.

Proof. Apply Theorem 3.4 with the choice

$$g(x,\tau) = x + \tau e_k, \quad H(x,\tau) = \overline{u}(x_\tau).$$

Note in particular, that

$$\mu = \partial_\tau\big|_{\tau=0}\overline{u} = \mathrm{D}\overline{u}(x)\cdot e_k$$

and

$$\nu = \partial_\tau\big|_{\tau=0}x_\tau = e_k.$$

$\square$

Example 3.6 (conserved quantities related to Dirichlet energy)**.** *Note first, that in the case of the Dirichlet energy, Corollary 3.5 only gives a trivial conserved quantity. Applying Noether's theorem with respect to the scaling symmetry*

$$g(x,\tau) = e^\tau x, \quad H(x,\tau) = e^{\tau\frac{d-2}{2}}\overline{u}(e^\tau x),$$

gives non-trivial information. In particular, we find

$$\mu = (n-2)(\overline{u}(x) + x\cdot \mathrm{D}\overline{u}(x)), \quad \nu(x) = x$$

and hence that

$$\operatorname{div}\left((2\mathrm{D}\overline{u}\cdot x + (n-2)\overline{u})\mathrm{D}\overline{u} - |\mathrm{D}\overline{u}|^2 x\right) = 0$$

Noting that due to Proposition 3.2, $\operatorname{div}\mathrm{D}\overline{u} = 0$ *almost every, we find using the Divergence theorem in combination with the conservation law,*

$$\begin{aligned}(n-2)\int_{B(x_0,r)}|\mathrm{D}\overline{u}|^2\,\mathrm{d}x &= \int_{B(x_0,r)}\operatorname{div}\left(|\mathrm{D}\overline{u}|^2 x - 2\mathrm{D}\overline{u}\cdot x\mathrm{D}u_0\right)\mathrm{d}x\\ &= \int_{\partial B(x_0,r)}|\mathrm{D}\overline{u}|^2 x\cdot\frac{x}{|x|} - 2\mathrm{D}\overline{u}\cdot x\mathrm{D}\overline{u}\cdot\frac{x}{|x|}\,\mathrm{d}\mathscr{H}^{n-1}\\ &= r\int_{\partial B(x_0,r)}|\mathrm{D}\overline{u}|^2 - \left(\mathrm{D}\overline{u}\cdot\frac{x}{|x|}\right)^2\mathrm{d}\mathscr{H}^{n-1}\end{aligned}$$

In particular, using Reynold's transport theorem, we can calculate,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}r}\frac{1}{r^{n-2}}\int_{B(x_0,r)}|\mathrm{D}\overline{u}|^2\,\mathrm{d}x &= -\frac{n-2}{r^{n-1}}\int_{B(x_0,r)}|\mathrm{D}\overline{u}|^2 + \frac{1}{r^{n-2}}\int_{\partial B(x_0,r)}|\mathrm{D}\overline{u}|^2\,\mathrm{d}\mathscr{H}^{n-1}\\ &= \frac{2}{r^{n-2}}\int_{\partial B(x_0,r)}\left(\mathrm{D}\overline{u}\cdot\frac{x}{|x|}\right)^2\mathrm{d}\mathscr{H}^{n-1}\geq 0.\end{aligned}$$

Thus, the quantity $\frac{1}{r^{d-2}}\int_{B(x_0,r)}|\mathrm{D}\overline{u}|^2\,\mathrm{d}x$ *is increasing in* r*. For* $n>2$*, this is a nontrivial information about* $\overline{u}$*.*

Example 3.7 (Brachistrone)**.** *Recall that in Example 1.3 we studied the minimisation problem*

$$\mathscr{F}[u] = \int_0^1\sqrt{\frac{1+(y')^2}{-y}}\,\mathrm{d}x$$

amongst graphs $y\colon [0,1] \to \mathbb{R}$ with $y(0) = 0$ and $y(1) = \overline{y} < 0$. Corollary 3.5 now tells us that with $f(y, y') = \sqrt{\frac{1+(y')^2}{-y}}$,

$$y' \cdot \partial_z f(y, y') - f(y, y') = \text{const.} = -\frac{1}{\sqrt{2r}},$$

for some $r > 0$. We take the constant to be negative here, as the following calculation will show that for a positive constant, the obtained y is not admissible. By elementary calculations, the above identity is equivalent to

$$(y')^2 = -\frac{2r}{y} - 1.$$

It is now straightforward to check that this equation has a unique solution given by

$$\begin{cases} x(t) = r(t - \sin(t)) \\ y(t) = -r(1 - \cos(t)) \end{cases}$$

Note however that due to technical issues (e.g. growth conditions) we have not rigorously shown this to be the unique solution. These issues can however be remedied.

4 Existence theory using the direct method

We now turn our attention towards direct methods. The philosophy behind this approach is to consider the minimisation problem in sufficiently weak spaces that existence can be proven relatively easily. Having obtained existence, the idea is to then prove a priori estimates that show that such weak solutions of our problem actually are classical solutions and thus the results of the classical methods we have seen so far apply. We will begin by studying in some detail the question of existence of minimsiers in this framework.

Throughout this section we will fix, unless otherwise specified, a Lipschitz domain Ω. The general setting we wish to study is the following:

$$\min_{u \in X} \mathscr{F}[u] = \min_{u \in X} \int_\Omega f(x, u, \mathrm{D}u)\,\mathrm{d}x, \qquad \text{where } X = \{u \in \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m) \colon u\big|_{\partial\Omega} = g\} \tag{4.1}$$

where $p \in (1, \infty)$ and $g \in \mathrm{W}^{1-\frac{1}{p},p}(\partial\Omega, \mathbb{R}^m)$. The continuity assumptions on f are summarised in saying that f is a Carathéodory integrand:

Definition 4.1. *$f\colon \Omega \times \mathbb{R}^n \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is a Carathéodory integrand if*

- *$f(\cdot, y, z)$ is measurable for all $(y, z) \in \mathbb{R}^n \times \mathbb{R}^{m\times n}$*
- *$f(x, \cdot, \cdot)$ is continuous for almost every $x \in \Omega$.*

4.1 Existence for the Dirichlet energy

We begin our study by considering the model case of the Dirichlet energy. We consider the problem

$$\min_{u \in X} \int_\Omega |\mathrm{D}u|^2\,\mathrm{d}x, \quad \text{where } X = \{u \in \mathrm{W}^{1,2}(\Omega) \colon u \in \mathrm{W}^{1,2}_{u_0}(\Omega)\}$$

where $u_0 \in \mathrm{W}^{\frac{1}{2},2}(\partial\Omega)$. Note that, since Ω is Lipschitz, u_0 can be extended to $u_0 \in \mathrm{W}^{1,2}(\Omega)$. Conversely, if $u_0 \in \mathrm{W}^{1,2}(\Omega)$, u_0 admits a trace on $\partial\Omega$ in $\mathrm{W}^{\frac{1}{2},2}(\partial\Omega)$.

In particular, a common, essentially equivalent formulation of the problem, is given by

$$\min_{u\in X}\int_\Omega |\mathrm{D}u|^2\,\mathrm{d}x, \quad \text{where } X=\{u\in \mathrm{W}^{1,2}(\Omega)\colon u\in u_0+\mathrm{W}^{1,2}_0(\Omega)\}$$

where $u_0\in \mathrm{W}^{1,2}(\Omega)$.

Set $\mathscr{F}[u]=\int_\Omega |\mathrm{D}u|^2\,\mathrm{d}x$. Note that $0\leq \mathscr{F}[u]$ for any $u\in X$ and $\mathscr{F}[u_0]<\infty$. In particular,

$$0\leq \inf_{u\in X}\mathscr{F}[u]=m<\infty.$$

Thus, we may find a sequence $(u_j)\subset X$ with $\mathscr{F}[u_j]\to m$ as $j\to\infty$. Due to Poincaré's inequality, we find

$$\|u_j\|_{\mathrm{W}^{1,2}(\Omega)}\lesssim \mathscr{F}[u_j]\leq \sup_j \mathscr{F}[u_j]<\infty.$$

Thus (u_j) is a bounded sequence in $\mathrm{W}^{1,2}(\Omega)$. Note further that X is an affine, closed subspace of $\mathrm{W}^{1,2}(\Omega)$, which is reflexive and separable. Hence, by Banach-Alaoglu theorem, (u_j) admits a (non-relabeled) subsequence such that $u_j\rightharpoonup u$ for some $u\in X$. Now, by Fatou's lemma,

$$m=\liminf_{j\to\infty}\int_\Omega |\mathrm{D}u_j|^2\,\mathrm{d}x\geq \int_\Omega |\mathrm{D}u|^2\,\mathrm{d}x.$$

Thus u is a minimiser of our problem.

There is an obvious trade-off in this strategy for proving existence. On the one hand, we need to establish the convergence of an appropriate subsequence of (u_j) with respect to some topology. On the other hand, the energy needs to be sequentially lower semi-continuous with respect to this topology (expressed in the above by Fatou's lemma). In general, establishing convergence is easier in weaker topologies, but establishing continuity properties is easier in stronger topologies. It is remarkable that for many energies the topology that enables carrying out the scheme of the direct method to prove existence agrees with the topology of the physically relevant spaces of solutions.

4.2 An abstract existence result

We now extract the essence of our existence proof with regards to the Dirichlet energy and establish an abstract existence result.

Let X be a complete metric space. Consider $\mathscr{F}\colon X\to\mathbb{R}\cup\{+\infty\}$ satisfying the following properties

(H1) $\mathscr{F}$ is coercive: For all $\Lambda\in\mathbb{R}$, $\{u\in X\colon \mathscr{F}[u]\leq\Lambda\}$ is sequentially pre-compact. That is whenever $(u_j)\subset X$ is such that $\mathscr{F}[u_j]\leq\lambda$ for all j, then (u_j) has a convergent subsequence.

(H2) $\mathscr{F}$ is sequentially lower semi-continuous (slsc): Whenever $(u_j)\subset X$ with $u_j\to u$, then $\mathscr{F}[u]\leq\liminf_{j\to\infty}\mathscr{F}[u_j]$.

We focus on the sequential definitions of (H1) and (H2) rather than their topological versions, since these are more convenient to us.

Theorem 4.2. *Let X be a complete metric space. Assume $\mathscr{F}\colon X\to\mathbb{R}\cup\{+\infty\}$ is coercive and sequentially lower semi-continuous. Then*

$$\min_{u\in X}\mathscr{F}[u]$$

admits at least one solution.

Proof. Assume there exists $u_0 \in X$ with $\mathscr{F}[u_0] < \infty$. Otherwise, any $u \in X$ is a solution. Then there exists $(u_j) \subset X$ such that $\mathscr{F}[u_j] \to \inf_{u\in X} \mathscr{F}[u] = m < \infty$. In particular, there is $\Lambda > 0$ such that $\mathscr{F}[u_j] \leq \Lambda$ for all j and thus due to (H1), (u_j) has a (non-relabeled) convergent subsequence such that $u_j \to u$. Due to (H2),

$$\mathscr{F}[u] \leq \liminf_{j\to\infty} \mathscr{F}[u_j] = m.$$

Thus u is a minimiser. □

As mentioned in the discussion of the previous section, we can usually not work with the strong topology and hence will often use a version of Theorem 4.2 phrased with respect to weak convergence.

Theorem 4.3. *Let X be a reflexive Banach space or a closed affine subset of a reflexive Banach space. Let $\mathscr{F}\colon X \to \mathbb{R} \cup \{+\infty\}$. Assume*

(wH1) $\mathscr{F}$ is weakly coercive: For any $\Lambda > 0$, $\{u \in X \colon \mathscr{F}[u] \leq \Lambda\}$ is sequentially weakly pre-compact.

(wH2) $\mathscr{F}$ is sequentially weakly lower semi-continuous: For all $(u_j) \subset X$ such that $u_j \rightharpoonup u \in X$,

$$\mathscr{F}[u] \leq \liminf_{j\to\infty} \mathscr{F}[u_j].$$

Then

$$\min_{u\in X} \mathscr{F}[u]$$

admits at least one solution.

Proof. The proof is exactly as in Theorem 4.2 using that any strongly closed affine subset of a Banach space is also weakly closed. □

4.3 Existence for integrands $f(x, z)$

We now wish to apply Theorem 4.3 to our problem (4.1). We first consider integrands $f \equiv f(x, z)$. In light of Theorem 4.3, we need to establish weak coercivity and sequential weak lower semicontinuity of $\mathscr{F}$. Before doing so, we are required to address a technical issue: we need to check that Carathéodory integrands are Lebesgue-measurable and hence that $\mathscr{F}$ is well-defined.

Lemma 4.4. *Let $f\colon \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}$ be Carathéodory. Then for any Borel-measurable $V\colon \Omega \to \mathbb{R}^{m\times n}$, $x \to f(x, V(x))$ is Lebesgue-measurable.*

Proof. Suppose first that V is a simple function, that is $V = \sum_{k=1}^m v_k 1_{E_k}$ where $E_k \subset \Omega$ are pairwise-disjoint Borel measurable sets such that $\Omega = \cup_{k=1}^m E_k$ and $v_k \in \mathbb{R}^{m\times n}$. Then

$$\{x \in \Omega \colon f(x, V(x)) > t\} = \cup_{k=1}^m \{x \in E_k \colon f(x, v_k) > t\}.$$

The right-hand side is a union of sets that are Lebesgue-measurable since $f(\cdot, v_k)$ is Lebesgue-measurable.

For a general V, approximate V by simple functions V_k so that

$$f(x, V_k(x)) \to f(x, V(x))$$

pointwise in Ω. Thus $f(x, V)$ is the pointwise limit of Lebesgue-measurable functions and hence Lebesgue-measurable. □

In general, it is possible that $x \to f(x, u, \mathrm{D}u)$ is measurable, but $\mathscr{F}[u]$ is not well-defined. This can be avoided, if e.g. $f \geq 0$ or

$$f(x, y, z) \lesssim (1 + |y|^p + |z|^p).$$

This growth assumption will however play no role in this chapter, except to exclude such pathological cases.

The most commonly used coercivity assumption and the only one we will consider is the existence of $\lambda > 0$ such that

$$\lambda |z|^p \leq f(x, z) \quad \forall (x, z) \in \Omega \times \mathbb{R}^{m \times n}. \tag{4.2}$$

Often, this condition is stated as

$$\lambda |z|^p - c \leq f(x, z)$$

for some $c > 0$. However, by setting $\tilde{f}(x, z) = f(x, z) + c$, this recovers (4.2) without changing the minimisers. Further, (4.2) specifies the space $\mathrm{W}^{1,p}(\Omega)$ in which we look for solutions.

Proposition 4.5. *If $f \colon \Omega \times \mathbb{R}^{m \times n} \to [0, \infty)$ is Carathéodory and satisfies* (4.2) *for some $\lambda > 0$ and $p \in (1, \infty)$, then $\mathscr{F}$ is weakly coercive on*

$$\mathrm{W}^{1,p}_g(\Omega, \mathbb{R}^m) = \{u \in \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m) \colon u\big|_{\partial\Omega} = g\},$$

where $g \in \mathrm{W}^{1-\frac{1}{p},p}(\partial\Omega)$.

Proof. Let $(u_j) \subset \mathrm{W}^{1,p}_g(\Omega)$ with $\sup_j \mathscr{F}[u_j] < \infty$. Due to (4.2),

$$\mu \sup_j \int_\Omega |\mathrm{D}u_j|^p \,\mathrm{d}x \leq \sup_j \mathscr{F}[u_j] < \infty.$$

Fix $u_0 \in \mathrm{W}^{1,p}_g(\Omega, \mathbb{R}^m)$. Then by Poincaré's inequality,

$$\begin{aligned} \|u_j\|_{\mathrm{W}^{1,p}(\Omega)} \leq & \|u_j - u_0\|_{\mathrm{W}^{1,p}(\Omega)} + \|u_0\|_{\mathrm{W}^{1,p}(\Omega)} \\ \lesssim & \|\mathrm{D}u_j - \mathrm{D}u_0\|_{\mathrm{L}^p(\Omega)} + \|u_0\|_{\mathrm{W}^{1,p}(\Omega)} \\ \leq & 2\|u_0\|_{\mathrm{W}^{1,p}(\Omega)} + \sup_j \|\mathrm{D}u_j\|_{\mathrm{L}^{1,p}(\Omega)} < \infty. \end{aligned}$$

As $p > 1$, this means (u_j) is a bounded sequence in a closed affine subspace of $\mathrm{W}^{1,p}(\Omega)$. As $\mathrm{W}^{1,p}(\Omega)$ is reflexive and separable, by Banach-Alaoglu theorem, (u_j) has a weakly convergent subsequence. Thus $\mathscr{F}$ is weakly coercive. $\square$

We next turn towards establishing sequential lower semi-continuity. A first result in this direction is due to Tonelli (for $n = 1$) and Serrin ($n > 1$). It shows that convexity implies sequential lower semi-continuity. For $n = 1$ (the one-dimensional case) or $m = 1$ (the scalar case), this is sharp. For $n > 1$, the sharp notion of convexity needed in order to imply sequential lower semi-continuity is called quasi-convexity. Quasi-convexity is still studied and far from being well-understood. We will return to study quasi-convexity in Chapter 6

Theorem 4.6. *Let $f \colon \Omega \times \mathbb{R}^{m \times n} \to [0, \infty)$ be Carathéodory and such that $f(x, \cdot)$ is convex for almost every $x \in \Omega$. Then $\mathscr{F}$ is weakly sequentially lower semi-continuous on $\mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$ for any $p \in (1, \infty)$.*

Proof. **Step 1.** We first prove that $\mathscr{F}$ is strongly sequentially lower semi-continuous. Suppose $u_j \to u$ in $\mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$. After passing to a subsequence, we may assume that $\mathrm{D}u_j \to \mathrm{D}u$ almost everywhere. Note $f(x, \mathrm{D}u_j) \geq 0$, so that by Fatou's theorem

$$\mathscr{F}[u] = \int_\Omega f(x, \mathrm{D}u)\,\mathrm{d}x \leq \liminf_{j\to\infty} \int f(x, \mathrm{D}u_j) = \liminf_{j\to\infty} \mathscr{F}[u_j].$$

As this holds for all subsequences, we deduce that $\mathscr{F}[u] \leq \liminf_{j\to\infty} \mathscr{F}[u_j]$ (see Discussion Sheet 2 for more details on this).

Step 2. Suppose $(u_j) \subset \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$ and $u \in \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$ are such that $u_j \rightharpoonup u$ in $\mathrm{W}^{1,p}(\Omega)$. Then there exists a subsequence, such that $\mathscr{F}[u_j] \to \liminf_{j\to\infty} \mathscr{F}[u_j] = m$. By Mazur's lemma, there exist

$$v_j = \sum_{n=j}^{N(j)} \theta_n^j u_n, \quad \theta_n^j \in [0,1], \quad \sum_{n=j}^{N(j)} \theta_n^j = 1$$

such that $v_j \to u$ in $\mathrm{W}^{1,p}(\Omega)$. On the one hand, as $f(x, \cdot)$ is convex for almost every $x \in \Omega$,

$$\begin{aligned}\mathscr{F}[v_j] &= \int_\Omega f\left(x, \sum_{n=j}^{N(j)} \theta_n^j \mathrm{D}u_n\right)\mathrm{d}x \\ &\leq \sum_{n=j}^{N(j)} \theta_n^j \int_\Omega f(x, \mathrm{D}u_n)\,\mathrm{d}x = \sum_{n=j}^{N(j)} \theta_n^j \mathscr{F}[u_n] \to m\end{aligned}$$

On the other hand, by Step 1,

$$\mathscr{F}[u] \leq \liminf_{j\to\infty} \mathscr{F}[v_j].$$

This concludes the proof. □

Combining Proposition 4.5 and Theorem 4.6 with Theorem 4.3 we obtain the following existence result:

Theorem 4.7. *Let $p \in (1, \infty)$. Let $f \colon \Omega \times \mathbb{R}^{m\times n} \to [0, \infty)$ be Carathéodory such that*

- *f is p-coercive: $f(x, z) \geq \lambda |z|^p$ for almost every $x \in \Omega$ and every $z \in \mathbb{R}^{m\times n}$.*
- *$f(x, \cdot)$ is convex for almost every $x \in \Omega$.*

Then $\mathscr{F}$ has a minimiser over $\mathrm{W}^{1,p}_g(\Omega, \mathbb{R}^m)$ where $g \in \mathrm{W}^{1-\frac{1}{p},p}(\partial\Omega, \mathbb{R}^m)$.

We now establish that convexity is the sharp assumption to obtain weak sequential lower semi-continuity in the scalar or one-dimensional case when f is independent of x.

Proposition 4.8. *Let $\mathscr{F} \colon \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m) \to \mathbb{R}$ be such that $f \equiv f(z) \colon \mathbb{R}^{m\times n} \to \mathbb{R}$ is continuous. If $\mathscr{F}$ is sequentially weakly lower semi-continuous on $\mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$ and $m = 1$ or $n = 1$, then f is convex.*

Proof. We focus on the case $m = 1$, $n \geq 1$. The case $n = 1$ can be dealt with similarly. Fix $a, b \in \mathbb{R}^n$, $a \neq b$ and $\theta \in (0,1)$. Write $v = \theta a + (1-\theta)b$. Introduce

$$u_j(x) = v \cdot x + \frac{1}{j}\phi_0(jx \cdot n - \lfloor jx \cdot n \rfloor)$$

where

$$\phi_0(x) = \begin{cases} -j(1-\theta)t & \text{if } t \in (0,\theta) \\ \theta t - \theta & \text{if } t \in [\theta, 1). \end{cases}$$

Note that

$$\mathrm{D}u_j = \begin{cases} a & \text{if } jx \cdot n - \lfloor jx \cdot n \rfloor \in (0,\theta), \\ b & \text{if } jx \cdot n - \lfloor jx \cdot n \rfloor \in [\theta, 1). \end{cases}$$

Further as ϕ_0 is bounded, $\frac{1}{j}\phi_0(jx \cdot n - \lfloor jx \cdot n \rfloor) \to 0$ uniformly in $[0,1]$. In particular, we deduce that $u_j \rightharpoonup v$ in $\mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$. By weak sequential lower semi-continuity,

$$|\Omega| f(v) = \mathscr{F}[x \cdot v] \leq \liminf_{j \to \infty} \mathscr{F}[u_j] = |\Omega| \left(\theta f(a) + (1-\theta) f(b)\right).$$

Thus f is convex. □

Finally, to close this section, we remark that the uniqueness part of Theorem 2.6 has a corresponding statement in this setting:

Proposition 4.9. *Let $p \in [1, \infty)$. Consider $\mathscr{F} \colon \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m) \to \mathbb{R}$ with $f \colon \Omega \times \mathbb{R}^{m \times n} \to \Omega$ Carathéodory. If $f(x, \cdot)$ is strictly convex for almost every $x \in \Omega$, then a minimiser $\overline{u} \in \mathrm{W}^{1,p}_g(\Omega, \mathbb{R}^m)$ of $\mathscr{F}$, where $g \in \mathrm{W}^{1-\frac{1}{p},p}(\partial\Omega, \mathbb{R}^m)$, is unique if it exists.*

Proof. Assume $u \neq v$ are minimisers of $\mathscr{F}$ in $\mathrm{W}^{1,p}_g(\Omega)$. Setting $w = \frac{u+v}{2}$, we find

$$\mathscr{F}[w] < \frac{1}{2}\mathscr{F}[u] + \frac{1}{2}\mathscr{F}[v] = \min_{u \in \mathrm{W}^{1,p}_g(\Omega, \mathbb{R}^m)} \mathscr{F}[u].$$

Thus $u = v$. □

4.4 Existence for integrands with u-dependence

We now wish to study functionals of the form

$$\mathscr{F}[u] = \int_\Omega f(x, u, \mathrm{D}u) \, \mathrm{d}x.$$

We will make the coercivity assumption

$$\lambda |z|^p + c_1 |y|^q + c_2 \leq f(x, y, z) \tag{4.3}$$

for some $\lambda > 0$, $c_1, c_2 \in \mathbb{R}$, $p > q \geq 1$ and almost every $x \in \Omega$, all $(y, z) \in \mathbb{R}^m \times \mathbb{R}^{n \times m}$. Thus $\mathscr{F}$ is weakly coercive using (the proof of) Proposition 4.5. If we try to prove sequential weak lower semi-continuity of $\mathscr{F}$ using the proof of Theorem 4.6, we encounter the expression

$$\int_\Omega f\left(x, \sum_{n=j}^{N(j)} \theta_n^j u_n, \sum_{n=j}^{N(j)} \theta_n^j \mathrm{D}u_n\right) \mathrm{d}x.$$

Whereas before, we were able to pull the sums outside of f using convexity, without assuming convexity of $f(x, \cdot, \cdot)$, we cannot do so anymore. Nevertheless, sequential weak lower semi-continuity does hold.

Theorem 4.10. *Suppose $f\colon \Omega \times \mathbb{R}^m \times \mathbb{R}^{n\times m} \to [0,\infty)$ is Carathéodory. Assume $f(x,y,\cdot)$ is convex for every $(x,y) \in \Omega \times \mathbb{R}^m$. Then for every $p \in (,\infty)$, $\mathscr{F}$ is sequentially weakly lower semi-continuous over $\mathrm{W}^{1,p}(\Omega,\mathbb{R}^m)$.*

Before completing the proof of Theorem 4.10, we note that as a consequence of it and of the coercivity of $\mathscr{F}$, we obtain an existence theorem.

Theorem 4.11. *Suppose the assumptions of the previous theorem hold. Assume in addition that (4.3) holds for some $\lambda > 0$ and $p \in (1,\infty)$. Then there exists at least one minimiser of the problem $\min_{u\in X} \mathscr{F}[u]$ with $X = \mathrm{W}^{1,p}_g(\Omega,\mathbb{R}^m)$, where $g \in \mathrm{W}^{1-\frac{1}{p},p}(\partial\Omega,\mathbb{R}^m)$.*

Proof of Theorem 4.10. We will prove the theorem under the additional assumption that $f \in C^1$ and that there exists $C \geq 0$ such that for all $(x,y,z) \in \overline{\Omega} \times \mathbb{R}^m \times \mathbb{R}^{m\times n}$ it holds

$$\lambda|z|^p \leq f(x,y,z) \tag{4.4}$$

$$|\partial_y f(x,y,z)| + |\partial_z f(x,y,z)| \leq C\left(1 + |y|^{p-1} + |z|^{p-1}\right) \tag{4.5}$$

The proof without this assumption is considerably more technical, we will go some way towards proving the full theorem in the problem class.

Let $(u_i) \subset \mathrm{W}^{1,p}_g(\Omega,\mathbb{R}^m)$ and $u \in \mathrm{W}^{1,p}_g(\Omega,\mathbb{R}^m)$ be such that $u_i \rightharpoonup u$ weakly in $\mathrm{W}^{1,p}(\Omega,\mathbb{R}^m)$. Using the convexity of $f(x,y,\cdot)$ and the C^1-regularity of f, we find

$$f(x,u_i,\mathrm{D}u_i) \geq f(x,u,\mathrm{D}u) + \partial_y f(x,u,\mathrm{D}u)(u_i - u) + \langle \partial_z f(x,u,\mathrm{D}u), \mathrm{D}u_i - \mathrm{D}u\rangle. \tag{4.6}$$

We now want to integrate this expression. In order to do this, we need to check that

$$\partial_y f(x,u,\mathrm{D}u)(u_i - u) + \langle \partial_z f(x,u,\mathrm{D}u), \mathrm{D}u_i - \mathrm{D}u\rangle \in \mathrm{L}^1(\Omega).$$

Indeed, using (4.4) and Hölder's inequality,

$$\begin{aligned}\int_\Omega |\partial_y f(x,u,\mathrm{D}u)||u_i - u|\,\mathrm{d}x \leq & \int \left(1 + |u|^{p-1} + |\mathrm{D}u|^{p-1})^{\frac{p}{p-1}}\,\mathrm{d}x\right)^{\frac{p-1}{p}} \|u_i - u\|_{L^p(\Omega)} \\ \leq & c\left(1 + \|u\|^p_{L^p(\Omega)} + \|u_i\|^p_{L^p(\Omega)}\right) < \infty.\end{aligned}$$

The other term is estimated similarly. In particular, integrating (4.6) over Ω,

$$\mathscr{F}[u_i] \geq \mathscr{F}[u] + \int_\Omega \partial_y f(x,u,\mathrm{D}u)(u_i - u)\,\mathrm{d}x + \int_\Omega \langle \partial_z f(x,u,\mathrm{D}u), \mathrm{D}u_i - \mathrm{D}u\rangle\,\mathrm{d}x.$$

However as $u_i \rightharpoonup u$ in $\mathrm{W}^{1,p}(\Omega)$,

$$\lim_{i\to\infty} \int_\Omega \partial_y f(x,u,\mathrm{D}u)(u_i - u)\,\mathrm{d}x = \lim_{i\to\infty} \int_\Omega \langle \partial_z f(x,u,\mathrm{D}u), \mathrm{D}u_i - \mathrm{D}u\rangle\,\mathrm{d}x = 0.$$

We deduce

$$\liminf_{i\to\infty} \mathscr{F}[u_i] \geq \mathscr{F}[u]$$

as desired. □

Example 4.12 (Linearised elasticity)**.** *With Theorem 4.10 we can handle a large class of problems. In particular, it is a useful tool to establish an existence theory also for* *problems that are coercive in the sense of* (4.3)*, but not of the form* $\mathscr{F}(x, u, \mathrm{D}u)$*. Such examples arise in linear elasticity theory.*

Suppose we are given a body in a reference configuration Ω*. We now deform* Ω *using an elastic deformation* $y \colon \Omega \to y(\Omega)$*. Since we are thinking of elastic deformations,* y *should be a differentiable bijection that is orientation preserving, i.e.* $\det \mathrm{D}y(x) > 0$ *for* $x \in \Omega$*. It is natural to describe the energy in terms of the deformation* $u(x) = y(x) - x$*. We expect the energy associated to the configuration* $y(\Omega)$ *to be preserved under rigid body motions of* Ω*. A commonly used energy with this property is the Green-St. Venant stress tensor*

$$G(u) = \frac{1}{2}\left(\mathrm{D}u + (\mathrm{D}u)^T + (\mathrm{D}u)^T\mathrm{D}u\right).$$

Note that $\mathscr{E}(u) = \mathrm{D}u + (\mathrm{D}u)^T$ *is the symmetric part of the gradient. If we assume* $\mathrm{D}u$ *is small, then*

$$G(u) \sim \frac{1}{2}\mathscr{E}(u).$$

Thus, in linearised elasticity a good basic model is given by looking for minimisers of the problem

$$\frac{1}{2}\int \langle \mathscr{E}(u), C(x)\mathscr{E}(u)\rangle \,\mathrm{d}x.$$

Here $C(x)$ *is a symmetric positive-definite fourth order tensor, known as the elasticity tensor. If the medium is homogeneous and isotropic,* $C(x) \equiv C$*, and it is possible to show that the energy above reduces to minimising*

$$\int_\Omega \mu|\mathscr{E}(u)|^2 + \frac{1}{2}\left(\kappa - \frac{2}{3}\mu\right)|\mathrm{tr}\,\mathscr{E}(u)|^2 \,\mathrm{d}x.$$

Here κ *is the bulk constant, while* μ *is called shear constant. Both are material constants describing properties of the medium.*

Consider hence the minimisation problem

$$\min_{u\in X} \mathscr{F}[u] = \min_{u \in X} \frac{1}{2}\int_\Omega 2\mu|\mathscr{E}(u)|^2 + \left(\kappa - \frac{2}{3}\mu\right)|\mathrm{tr}\,\mathscr{E}(u)|^2 - b\cdot u \,\mathrm{d}x$$
$$X = \{u \in \mathrm{W}^{1,2}(\Omega, \mathbb{R}^2) \colon u\big|_{\partial\Omega} = g, g \in \mathrm{W}^{\frac{1}{2},2}(\partial\Omega, \mathbb{R}^m)\}.$$

Here $b \in \mathrm{L}^2(\Omega)$ *describes an external force we apply to the body. Note that* $\mathscr{F}$ *has quadratic growth and is convex in* $\mathscr{E}(u)$*. Thus, in order to obtain existence of minimisers, using Theorem 4.10 it suffices to show* $\mathscr{F}$ *is coercive.*

We will establish coercivity under the assumption $g = 0$ *and* $\kappa - \frac{2}{3}\mu \geq 0$ *for simplicity. Note that by a direct, pointwise calculation for any* ϕ*,*

$$2\langle \mathscr{E}(\phi) \colon \mathscr{E}(\phi)\rangle - \langle \mathrm{D}\phi, \mathrm{D}\phi\rangle = \mathrm{div}\,\left(\mathrm{D}\phi\phi - \mathrm{div}\,(\phi)\phi\right) + (\mathrm{div}\,\phi)^2.$$

Integrating this identity over Ω *and applying the divergence theorem, we find*

$$\begin{aligned} 2\|\mathscr{E}(\phi)\|^2_{\mathrm{L}^2(\Omega)} - 2\|\mathrm{D}\phi\|^2_{\mathrm{L}^2(\Omega)} &= \int_\Omega \mathrm{div}\,\left(\mathrm{D}\phi\phi - \mathrm{div}\,(\phi)\phi\right) + (\mathrm{div}\,\phi)^2 \,\mathrm{d}x \\ &\geq \int_\Omega (\mathrm{div}\,\phi)^2 \,\mathrm{d}x = 0 \end{aligned}$$

This inequality is known as Korn's inequality. Thus, using Hölder's inequality, we deduce

$$\mathscr{F}[u] \geq \mu\|\mathscr{E}(u)\|^2_{\mathrm{L}^2(\Omega)} - \|b\|_{\mathrm{L}^2(\Omega)}\|u\|_{\mathrm{L}^2(\Omega)}$$

$$
\begin{aligned}
&\geq \mu \|\mathrm{D}u\|^2_{\mathrm{L}^2(\Omega)} - C(\delta)\|b\|^2_{\mathrm{L}^2(\Omega)} - \delta\|u\|^2_{\mathrm{L}^2(\Omega)} \\
&\geq \frac{\mu}{2}\|\mathrm{D}u\|_{\mathrm{L}^2(\Omega)} - C\|b\|_{\mathrm{L}^2(\Omega)}.
\end{aligned}
$$

To obtain the last line, we chose δ sufficiently small. In particular, arguing as for Proposition 4.5, we see that $\mathscr{F}$ is coercive and hence existence of minimisers follows.

Theorem 4.10 is essentially sharp if $m = 1$ or $n = 1$. If $m > 1$ or $n < 1$, the convexity assumption can be weakened as we already discussed in relation to Tonelli's lower semi-continuity theorem and will discuss in more detail in the future. We collect below a number of examples that show that the assumption $p \in (1, \infty)$, as well as the assumptions $\lambda > 0$ and $p < q$ in (4.3) cannot be relaxed in general.

Example 4.13 ($\boldsymbol{p = 1}$)**.** *Let $n = 1$, $f(x, y, z) = \sqrt{y^2 + z^2}$. Consider*

$$
\inf_{u \in X} \int_0^1 f(x, u, u')\,\mathrm{d}x \qquad \textit{where } X = \{u \in \mathrm{W}^{1,1}((0,1))\colon u(0) = 0,\ u(1) = 1\}.
$$

Note

$$
\mathscr{F}[u] \geq \int_0^1 u'\,\mathrm{d}x = 1.
$$

Consider

$$
u_i = \begin{cases} 0 & \textit{if } x \in [0, 1 - 1/i] \\ 1 + i(x-1) & \textit{if } x \in [1 - 1/i, 1]. \end{cases}
$$

Note that then

$$
\mathscr{F}[u_i] = \int_{1-1/i}^1 \sqrt{(1 + i(x-1))^2 + i^2}\,\mathrm{d}x \leq \frac{1}{i}\sqrt{1 + i^2} \to 1.
$$

Hence, if the problem admits a minimiser u, it must satisfy

$$
1 = \mathscr{F}[u] = \int_0^1 \sqrt{u^2 + (u')^2}\,\mathrm{d}x \geq \int u'\,\mathrm{d}x = 1.
$$

In particular, $u' = 1$ almost everywhere in $(0,1)$ and $u = 0$ almost everywhere in $(0,1)$, which is a contradiction.

Example 4.14 (Weierstrass example, $\boldsymbol{\lambda > 0}$)**.** *Let $n = 1$ and consider $f(x, y, z) = xz^2$. We study*

$$
\inf_{u \in X} \int_0^1 f(x, u, u')\,\mathrm{d}x \qquad \textit{where } X = \{u \in \mathrm{W}^{1,1}((0,1))\colon u(0) = 0,\ u(1) = 1\}.
$$

We saw on a problem sheet that this problem does not have a solution in the space $Y = X \cap C^1([0,1])$ and that the infimum over this space is 0. Clearly

$$
0 \leq \inf_{u \in X} \int_0^1 f(x, u, u')\,\mathrm{d}x \leq \inf_{u \in X \cap C^1([0,1])} \int_0^1 f(x, u, u')\,\mathrm{d}x = 0.
$$

In particular, any possible minimiser must satisfy $u' = 0$ almost everywhere in $(0,1)$, which gives a contradiction.

Example 4.15 ($\boldsymbol{p = q}$)**.** *We have already seen in a problem class that if $n = 1$ and $\lambda > \pi$ with $f(x, y, z) = \frac{1}{2}(z^2 - \lambda^2 y^2)$,*

$$\inf_{u \in \mathrm{W}_0^{1,2}(\Omega)} \int_0^1 f(x, u, u')\,\mathrm{d}x = -\infty.$$

Thus this problem admits no solution.

Example 4.16 (convexity if $\boldsymbol{n = 1}$)**.** *Consider with $n = 1$, $f(y, z) = (z^2 - 1)^2 + y^4$ and the problem*

$$\inf_{u \in \mathrm{W}_0^{1,4}((0,1))} \int_0^1 f(u, u')\,\mathrm{d}x = m$$

Assume that we have shown that $m = 0$. Then if u is a minimiser, $u = 0$ almost everywhere in $(0, 1)$ and $|u'| = 1$ almost everywhere in $(0, 1)$. But by Sobolev embededing, functions in $\mathrm{W}^{1,4}((0, 1))$ are continuous, so that this provides a contradiction.

This also suggests how to build competitors (u_i) with $\lim \int_0^1 f(u_i, u_i')\,\mathrm{d}x = 0$. By direct calculation the family (u_i) defined on each interval $[k/i, (k + 1)/i]$ for $0 \le k \le i - 1$ via

$$u_i(x) = \begin{cases} x - \frac{k}{i} & \text{if } x \in \left[\frac{2k}{2i}, \frac{2k+1}{2i}\right] \\ -x + \frac{k+1}{i} & \text{if } x \in \left(\frac{2k+1}{2i}, \frac{2(k+1)}{2i}\right] \end{cases}$$

has the desired property.

4.5 Integral side constraints

In this section, we have so far only considered Dirichlet boundary conditions. However in Section 1, we introduced a number of other side conditions. The most important of these was an integral side condition of the form $\int h(x, u)\,\mathrm{d}x = 0$ for some function h. We now adapt the direct method to this setting, beginning with an abstract existence theorem in the framework of weak convergence.

Theorem 4.17. *Let X be a Banach space or a closed affine subset of a Banach space. Suppose $\mathscr{F}, \mathscr{H} \colon X \to \mathbb{R} \cup \{+\infty\}$. Assume*

(wH1) *$\mathscr{F}$ is weakly coercive: For any $\Lambda > 0$, $\{u \in X \colon \mathscr{F}[u] \le \Lambda\}$ is sequentially weakly pre-compact.*

(wH2) *$\mathscr{F}$ is sequentially weakly lower semi-continuous.*

(wH3) *$\mathscr{H}$ is weakly continuous: Whenever $(u_j) \subset X$ with $u_j \rightharpoonup u$ in X, then $\mathscr{H}[u_j] \to \mathscr{H}[u]$.*

Assume there is $u_0 \in X$ with $\mathscr{H}[u_0] = 0$. Then the problem

$$\min_{\{u \in X \colon \mathscr{H}[u] = 0\}} \mathscr{F}[u]$$

admits at least one solution.

Proof. The proof is exactly as the proof of Theorem 4.3, except that we need to take the subsequence (u_j) to additionally satisfy $\mathscr{H}[u_j] = 0$. In particular, this will imply, with $u_j \rightharpoonup u$ weakly in X, $\mathscr{H}[u] = 0$. $\square$

The following lemma gives the required weak continuity for an integral side constraint under reasonable assumptions on h.

Lemma 4.18. *Let $h\colon \Omega \times \mathbb{R}^m \to \mathbb{R}$ be Carathéodory. Suppose for some $p \in [1, \infty)$ and $C > 0$,*

$$|h(x, y)| \leq C(1 + |y|^q) \tag{4.7}$$

for almost every $x \in \Omega$ and every $y \in \mathbb{R}^m$, where $q \in \left[1, \frac{np}{n-p}\right)$ if $p < n$ and $q < \infty$ if $p \geq n$. Then $\mathscr{H}[u] = \int_\Omega h(x, u)\, \mathrm{d}x$ is weakly continuous.

Proof. We deal with the case $p < n$. The argument in case $p \geq n$ is similar, but easier. Suppose $u_j \rightharpoonup u$ in $\mathrm{W}^{1,p}(\Omega)$. Up to passing to a non-relabeled subsequence, we may assume $u_j \to u$ in $\mathrm{L}^q(\Omega)$ and almost everywhere in Ω. By (4.7), we find for almost every $x \in \Omega$,

$$h(x, y) + C(1 + |y|^q) \geq 0,$$

so that by Fatou's lemma,

$$\liminf_{j\to\infty} \mathscr{H}[u_j] + \int_\Omega C(1 + |u_j|^q)\, \mathrm{d}x \geq \mathscr{H}[u] + \int_\Omega C(1 + |u|^q)\, \mathrm{d}x$$

Noting that we may replace $h \to -h$ in the argument above, combined with the fact that $u_j \to u$ in $\mathrm{L}^q(\Omega)$, we deduce

$$\liminf_{j\to\infty} \mathscr{H}[u_j] \geq \mathscr{H}[u], \qquad \liminf_{j\to\infty} -\mathscr{H}[u_j] \geq -\mathscr{H}[u].$$

However, note that this implies $\lim_{j\to\infty} \mathscr{H}[u_j] = \mathscr{H}[u]$. As this holds for a subsequence of any subsequence of (u_j), the lemma is proven. $\square$

Remark 4.19. *Note that the previous lemma captures in particular the side constraint $\int_\Omega u = 0$, which is commonly imposed when considering Neumann boundary conditions.*

Combining Theorem 4.17 with Proposition 4.5, Theorem 4.6 and Lemma 4.18, we obtain an existence result.

Theorem 4.20. *Let $f\colon \Omega \to \mathbb{R}^{m\times n} \to [0, \infty)$ and $h\colon \Omega \times \mathbb{R}^m \to \mathbb{R}$ be Carathéodory and such that*

- *There is $\lambda > 0$ such that for almost every $x \in \Omega$ and every $z \in \mathbb{R}^{m\times n}$, $f(x, z) \geq \lambda |z|^p$.*
- *For almost every $x \in \Omega$, $f(x, \cdot)$ is convex.*
- *For almost every $x \in \Omega$, $z \in \mathbb{R}^{m\times n}$, $|h(x, y)| \leq C(1 + |y|^q)$ for some $C > 0$ and $q \in \left[1, \frac{np}{n-p}\right)$ if $p < n$, any $q < \infty$ if $p \geq n$.*

Let $g \in \mathrm{W}^{1-\frac{1}{p},p}(\partial\Omega, \mathbb{R}^m)$. Assume there exists $u_0 \in \mathrm{W}^{1,p}_g(\Omega, \mathbb{R}^m)$ with

$$\mathscr{H}[u_0] = \int_\Omega h(x, u_0)\, \mathrm{d}x = 0.$$

Then there exists at least one solution to the problem

$$\min_{\{u\in \mathrm{W}^{1,p}_g(\Omega,\mathbb{R}^m)\colon\, \mathscr{H}[u]=0\}} \int_\Omega f(x, \mathrm{D}u)\, \mathrm{d}x.$$

5 Regularity

Our existence results in the previous section raise the question of whether it is possible to return from the regime of minimisation in $\mathrm{W}^{1,p}(\Omega)$ to the regime of classical solutions of the Euler-Lagrange equation which we studied in Section 2. In other words, we are asking whether a priori regularity estimates for minimisers in Sobolev spaces are available. In general, there are two natural directions of regularity questions: concerning regularity of u and concerning regularity of derivatives of u, i.e. of $\mathrm{D}u$. Studying properties of $\mathrm{D}u$ requires use of the Euler-Lagrange equation (in its weak form) and thus falls into the regime of techniques applicable to elliptic PDEs. This setting will be covered in other courses. While in certain situations, it is still the case that properties of the Euler-Lagrange equation need to be combined with minimality, we focus here on aiming for $C^{0,\alpha}$-regularity of u for some $\alpha > 0$ in the scalar case $m = 1$, which can be obtained using minimality only.

Throughout this section Ω will always be a Lipschitz domain.

5.1 Lavrentiev phenomenon

Before proving Hölder-regularity of minimisers, we study the question of whether at least the minimal value of energies agrees in Sobolev spaces and smooth spaces. In other words, does it hold that

$$\inf_{u \in \mathrm{W}^{1,p}(\Omega)} \mathscr{F}[u] = \inf_{u \in C^\infty(\Omega)} \mathscr{F}[u].$$

Note that even if the answer to this question is positive, we should in general not expect both infima to be attained- there are counterexamples that show that minimisers may fail to be smooth! This question is not just a mathematical question, but has direct implications for numerical methods. A standard approach to solving energy minimisation problems numerically is to apply a finite element approach to the Euler-Lagrange equation. Thus, we hope to approximate our solution by solutions that live in smooth (usually at least piecewise affine) function spaces. If the minimal value over Sobolev functions and over the approximation spaces do not agree, then this approach cannot be used. In this case, more advanced and difficult methods, such as using non-conformal methods need to be applied.

Rephrasing the question in a slightly more abstract framework, we consider $X \subset Y$, where X is a dense subspace of a Banach space Y. Given $\mathscr{F} \colon Y \to \mathbb{R} \cup \{+\infty\}$, we ask whether it is true that

$$\inf_{u \in X} \mathscr{F}[u] = \inf_{u \in Y} \mathscr{F}[u]. \tag{5.1}$$

Note that if $\mathscr{F}$ is strongly continuous in Y, then the above identity does indeed hold. This observation is at the heart of the following result:

Theorem 5.1. *Suppose $f \colon \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times n} \to \mathbb{R}$ is Carathéodory and satisfies for almost every $x \in \Omega$ and $(y, z) \in \mathbb{R}^m \times \mathbb{R}^{m \times n}$,*

$$|f(x, y, z)| \leq C(1 + |y|^p + |z|^p)$$

for some $C > 0$ and $p \in [1, \infty)$. Then $\mathscr{F}[u] = \int_\Omega f(x, u, \mathrm{D}u)\,\mathrm{d}x$ is strongly continuous in $\mathrm{W}^{1,p}(\Omega)$. Consequently,

$$\inf_{u \in \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)} \mathscr{F}[u] = \inf_{u \in C^\infty(\Omega, \mathbb{R}^m)} \mathscr{F}[u].$$

Remark 5.2. *Theorem 5.1 can be adapted to include problems with boundary conditions.*

Proof of Theorem 5.1. Suppose $u_j \to u$ in $\mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$. Up to passing to a subsequence, we may assume that $u_j \to u$ and $\mathrm{D}u_j \to \mathrm{D}u$ almost everywhere in Ω. Further

$$\mathscr{F}[u] \leq \int_\Omega C(1 + |u_j|^p + |\mathrm{D}u_j|^p)\,\mathrm{d}x \leq C|\Omega| + C \sup_j \|u_j\|^p_{\mathrm{W}^{1,p}(\Omega)} < \infty.$$

Thus, by a version of the dominated convergence theorem,

$$\lim_{j\to\infty} \mathscr{F}[u_j] = \mathscr{F}[u].$$

By the usual subsequence of subsequences argument, we conclude the proof. □

In light of Theorem 5.1 and Theorem 4.11, if $f(x, y, z)$ satisfies p-growth and a p-coercivity condition, we obtain existence of minimisers and (5.1) holds with $Y = \mathrm{W}^{1,p}(\Omega)$, $X = C^\infty(\Omega)$. The following example shows that in general, (5.1) need not hold.

Example 5.3 (Mania's example (1934))**.** *Consider the minimisation problem*

$$\min_{u(0)=0,u(1)=1} \mathscr{F}[u] = \min_{u(0)=0,u(1)=1} \int_0^1 (u^3 - t)^2 (u')^6\,\mathrm{d}t.$$

Note that in $\mathrm{W}^{1,1}((0,1))$, $u(t) = t^{\frac{1}{3}}$ *is clearly a solution and*

$$\inf_{\{u\in\mathrm{W}^{1,1}((0,1))\colon u(0)=0,u(1)=1\}} \mathscr{F}[u] = 0.$$

We will show that

$$\inf_{\{u\in\mathrm{W}^{1,\infty}((0,1))\colon u(0)=0,u(1)=1\}} \mathscr{F}[u] > 0.$$

Note that if $u \in \mathrm{W}^{1,\infty}((0,1))$, *then* u *is Lipschitz. In particular, there is* $\tau \in (0,1)$ *such that for* $t \in [0,\tau]$

$$u(t) \leq h(t) = \frac{t^{\frac{1}{3}}}{2}, \qquad u(\tau) = h(\tau).$$

In particular, we can estimate

$$\mathscr{F}[u] \geq \int_0^\tau (u^3 - t)^2 (u')^6\,\mathrm{d}t \geq \frac{7^2}{8^2} \int_0^\tau t^2 (u')^6\,\mathrm{d}t.$$

Using Hölder's inequality, we further find,

$$\frac{\tau^{\frac{1}{3}}}{2} = \int_0^\tau u'\,\mathrm{d}t = \int_0^\tau t^{-\frac{1}{3}} t^{\frac{1}{3}} u'\ dt \leq \frac{5^{\frac{5}{6}}}{3^{\frac{5}{6}}} \tau^{\frac{1}{2}} \left(\int t^2 (u')^6\,\mathrm{d}t \right)^{\frac{1}{6}}.$$

Combining estimates, we deduce

$$\mathscr{F}[u] \geq \frac{7^2 3^5}{8^2 5^2 2^6 \tau} \geq \frac{7^2 3^5}{8^2 5^2 2^6} > 0.$$

Note that $\mathscr{F}$ *is not coercive and one may wonder whether this causes the energies to not match. However Ball and Mizel showed in 1985 that one can adapt the above construction and obtain a coercive integrand. To be precise they showed that if*

$$\mathscr{F}[u] = \int_{-1}^1 (t^4 - u^6)^2 |u'|^2 6 + \varepsilon |u'|^2\,\mathrm{d}t$$

for ε *sufficiently small, then this integrand is coercive and nevertheless,*

$$\inf_{\{u\in\mathrm{W}^{1,1}((-1,1))\colon u(-1)=-1,u(1)=1\}} \mathscr{F}[u] < \inf_{\{u\in\mathrm{W}^{1,\infty}((-1,1))\colon u(-1)=-1,u(1)=1\}} \mathscr{F}[u]$$

Example 5.3 shows that as soon as the space Y in which we have coercivity does not match the space in which we have strong continuity of our integrand X, we should expect that

$$\inf_Y \mathscr{F}[u] < \inf_X \mathscr{F}[u]$$

can occur. This phenomenon is known as Lavrentiev's phenomenon. One way of correcting this problem is to relax our notion of what we mean with $\mathscr{F}[u]$. We illustrate this example thinking of an integrand $f(x,z)\colon \Omega\times\mathbb{R}^{m\times n}\to\mathbb{R}$ satisfying p-coercivity and q-growth with $1<p\leq q<\infty$: For some $\mu, C>0$,

$$f(x,z)\geq \mu|z|^p \qquad |f(x,z)|\leq 1+|z|^q \tag{5.2}$$

for almost every $x\in\Omega$ and every $z\in\mathbb{R}^{m\times n}$. We denote $X=\mathrm{W}^{1,q}(\Omega)\cap\mathrm{W}^{1,p}_0(\Omega)$, $Y=\mathrm{W}^{1,p}_0(\Omega)$. Note that $X\subset Y$ is dense. Following the framework of Buttazzo and Mizel, introduce the relaxed functional

$$\overline{\mathscr{F}}_X(u)=\inf\{\liminf_{j\to\infty}\mathscr{F}[u_j]\colon (u_j)\subset X, u_j\rightharpoonup u \text{ weakly in } Y\}.$$

Note that if $x\in X$, $\overline{\mathscr{F}}_X[u]=\mathscr{F}[u]$, by strong continuity of $\mathscr{F}$ on X (e.g. Theorem 5.1). Further if $\mathscr{G}$ is sequentially weakly lower semi-continuous (swlsc) on Y, then whenever $(u_j\subset X)$, $u_j\rightharpoonup u$ weakly in Y, it holds that $\mathscr{G}[u]\leq\liminf_{j\to\infty}\mathscr{G}[u_j]$. In particular, we see that $\overline{\mathscr{F}}_X$ is the sequential weakly lower semi-continuous envelope of $\mathscr{F}$,

$$\overline{\mathscr{F}}_X=\sup\{\mathscr{G}\colon X\to\mathbb{R}\colon \mathscr{G} \text{ is swlsc on } Y, \mathscr{G}\leq\mathscr{F} \text{ on } X\}.$$

The relaxed functional is convex if $\mathscr{F}$ is.

Lemma 5.4. *If $\mathscr{F}\colon Y\to\mathbb{R}$ is (strictly) convex, then so is $\mathscr{F}_X\colon Y\to\mathbb{R}$.*

Proof. Let $u,v\in Y$ and $\lambda\in[0,1]$. Then

$$\begin{aligned}
&\mathscr{F}_X(\lambda u+(1-\lambda)v)\\
\leq&\inf\{\liminf_{j\to\infty}\mathscr{F}[\lambda u_j+(1-\lambda)v_j]\colon (u_j),(v_j)\subset X, u_j\rightharpoonup u, v_j\rightharpoonup v \text{ weakly in } Y\}\\
\leq&\inf\{\lambda\liminf_{j\to\infty}\mathscr{F}[u_j+(1-\lambda)\liminf_{j\to\infty}\mathscr{F}[v_j]\colon (u_j),(v_j)\subset X, u_j\rightharpoonup u, v_j\rightharpoonup v \text{ weakly in } Y\}\\
=&\lambda\overline{\mathscr{F}}_X[u]+(1-\lambda)\overline{\mathscr{F}}_X[v].
\end{aligned}$$

□

Using this fact in combination with Theorem 4.3, we see that $\mathscr{F}_X$ has at least one minimiser and if $\mathscr{F}_X$ is strictly convex, this minimiser is unique. For the following, we restrict to the case $p\geq 2$ and additionally impose that f is strongly p-convex, in the sense that for almost every $x\in\Omega$ and $z_1,z_2\in\mathbb{R}^{m\times n}$, $\lambda\in[0,1]$,

$$f(\lambda z_1+(1-\lambda)z_2)\leq\lambda f(z_1)+(1-\lambda)f(z_2)-\lambda(1-\lambda)C|z_1-z_2|^p \tag{5.3}$$

for some $C>0$. Our restriction is caused due to the fact that if $p\leq 2$, only affine functions satisfy the above inequality and hence a technical adaption is required.

A second natural approach to correct the problem of Lavrentiev's phenomenon in this set-up would be to perturb the energy in such a way as to naturally enforce minimisers to lie in X. This is usually achieved by introducing

$$\mathscr{F}_\varepsilon[u]=\int_\Omega f(x,\mathrm{D}u)\,\mathrm{d}x+\varepsilon\int_\Omega|\mathrm{D}u_\varepsilon|^q\,\mathrm{d}x.$$

Note that from Theorem 4.11, we obtain existence of a minimiser $u_\varepsilon \in \mathrm{W}^{1,q}(\Omega)$ of the problem

$$\min_{u\in X} \mathscr{F}_\varepsilon[u].$$

We might hope that as $\varepsilon \to 0$, $\mathscr{F}_\varepsilon[u_\varepsilon] \to \inf_{u\in Y} \mathscr{F}[u]$ and even $u_\varepsilon \to u$ in Y, where u is a minimiser of $\min_{u\in Y} \mathscr{F}[u]$. This is not quite true and in fact, convergence to the relaxed functional and its minimiser holds.

Theorem 5.5. *Suppose $p \geq 2$. Assume $f\colon \Omega \times \mathbb{R}^{m\times n}$ satisfies (5.2) and $f(x,\cdot)$ is strictly convex for almost every $x \in \Omega$. Then*

$$\mathscr{F}_\varepsilon[u_\varepsilon] \to \overline{\mathscr{F}}_X[u]$$

where u_ε solves $\min_{u\in X} \mathscr{F}_\varepsilon[u]$ and u is the unique minimiser of $\overline{\mathscr{F}}_X$ over Y. Moreover if, in addition f satisfies (5.3), $u_\varepsilon \to u$ in Y.

Proof. Note that (u_ε) is bounded in $\mathrm{W}_0^{1,p}(\Omega,\mathbb{R}^m)$ due to (5.2). In particular any subsequence of (u_ε) admits a weakly converging subsequence in $\mathrm{W}_0^{1,p}(\Omega,\mathbb{R}^m)$. Thus, we deduce,

$$\overline{\mathscr{F}}_X[u] \leq \liminf_{\varepsilon\to 0} \mathscr{F}[u_\varepsilon] \leq \liminf_{\varepsilon\to 0} \mathscr{F}_\varepsilon[u_\varepsilon].$$

On the other hand, for any $v \in X$,

$$\limsup_{\varepsilon\to 0} \mathscr{F}_\varepsilon(u_\varepsilon) \leq \limsup_{\varepsilon\to 0} \mathscr{F}_\varepsilon(v) = \mathscr{F}(v) = \overline{\mathscr{F}}_X(v)$$

By definition of $\overline{\mathscr{F}}_X$, this inequality extends to $v \in Y$. In particular,

$$\limsup \mathscr{F}_\varepsilon[u_\varepsilon] \leq \overline{\mathscr{F}}_X[u].$$

In particular, if $u_\varepsilon \rightharpoonup v$ weakly in $\mathrm{W}^{1,p}(\Omega)$ as $p \to \infty$, then v is a minimiser of $\overline{\mathscr{F}}_X$. Due to strict convexity of $\overline{\mathscr{F}}_X$, $v = u$.

Arguing as for Lemma 5.4 and using (5.3), we deduce

$$\overline{\mathscr{F}}_X\left[\frac{u+u_\varepsilon}{2}\right] + \frac{C}{4}\|\mathrm{D}u_\varepsilon - \mathrm{D}u\|^p_{\mathrm{L}^p(\Omega)} \leq \frac{1}{2}\overline{\mathscr{F}}_X[u] + \frac{1}{2}\overline{\mathscr{F}}_X[u_\varepsilon] \to \mathscr{F}_X[u].$$

By minimality of u, we deduce $\mathrm{D}u_\varepsilon \to \mathrm{D}u$ in $\mathrm{L}^p(\Omega)$. Using Poincarés inequality, we conclude the proof. $\square$

5.2 Quasiminimality

We now introduce the set-up in which we will prove $C^{0,\alpha}$-regularity results. Since our approach will be local in nature, it is convenient to work with a notion of minimisers that is adapted to this:

Definition 5.6. *Let $f\colon \Omega\times\mathbb{R}^m\times\mathbb{R}^{m\times n} \to \mathbb{R}$ be Carathéodory. $u \in \mathrm{W}^{1,p}_{\mathrm{loc}}(\Omega,\mathbb{R}^m)$ is a local minimum of $\mathscr{F}[u] = \int_\Omega f(x,u,\mathrm{D}u)\,\mathrm{d}x$ if for any $\phi \in \mathrm{W}^{1,p}(\Omega,\mathbb{R}^m)$ with $K = \operatorname{supp}\phi \Subset \Omega$,*

$$\mathscr{F}[u,K] = \int_K f(x,u,\mathrm{D}u)\,\mathrm{d}x \leq \mathscr{F}[u+\phi,K].$$

Remark 5.7. *Under suitable growth conditions local minimisers are local weak solutions of the Euler-Lagrange equation.*

In fact, we can consider a slightly more general set-up of quasi-minimisers.

Definition 5.8. *Let $f\colon \Omega\times\mathbb{R}^m\times\mathbb{R}^{m\times n}\to\mathbb{R}$ be Carathéodory and $Q\geq 1$. A function $u\in \mathrm{W}^{1,p}_{\mathrm{loc}}(\Omega,\mathbb{R}^m)$ is a local Q-quasiminimiser of $\mathscr{F}[u]=\int_\Omega f(x,u,\mathrm{D}u)\,\mathrm{d}x$ if for any $\phi\in \mathrm{W}^{1,p}(\Omega,\mathbb{R}^m)$ with $K=\operatorname{supp}\phi\Subset\Omega$,*

$$\mathscr{F}[u,K]=\int_K f(x,u,\mathrm{D}u)\,\mathrm{d}x\leq Q\mathscr{F}[u+\phi,K].$$

Note that local minimisers are precisely local 1-quasiminimisers.

Example 5.9. *Consider $f\colon\mathbb{R}^{m\times n}\to\mathbb{R}$ Carathéodory with $\Lambda^{-1}|z|^p\leq f(z)\leq\Lambda|z|^p$ for some $p\in[1,\infty)$ and $\Lambda>0$. Then minimisers $u\in\mathrm{W}^{1,p}(\Omega,\mathbb{R}^m)$ of $\int_\Omega F(\mathrm{D}u)\,\mathrm{d}x$ are quasi-minimisers of $\int_\Omega|\mathrm{D}v|^p\,\mathrm{d}x$. Indeed, for any $v\in\mathrm{W}^{1,p}(\Omega,\mathbb{R}^m)$,*

$$\int_\Omega|\mathrm{D}u|^p\,\mathrm{d}x\leq\Lambda\int_\Omega f(\mathrm{D}u)\,\mathrm{d}x\leq\Lambda\int_\Omega f(\mathrm{D}v)\,\mathrm{d}x\leq\Lambda^2\int_\Omega|\mathrm{D}v|^p\,\mathrm{d}x.$$

Another reason for introducing quasi-minimisers is that solutions of the Euler-Lagrange equation turn out, under certain assumptions to be quasi-minimisers. We provide an example of this.

Definition 5.10. *If for all $\phi\in\mathrm{W}^{1,p}_0(\Omega,\mathbb{R})$,*

$$\int A^i(x,u,\mathrm{D}u)\mathrm{D}_i\phi-B(x,u,\mathrm{D}u)\phi\,\mathrm{d}x=0,$$

then u is a weak solution of

$$\operatorname{div}A(x,u,\mathrm{D}u)-B(x,u,\mathrm{D}u)=0. \tag{5.4}$$

We impose the following ellipticity and growth conditions on A,

$$A^i(x,y,z)z_i\geq|z|^p-a_1(x),\qquad|A(x,y,z)|\leq\Lambda|z|^{p-1}+a_2(x) \tag{5.5}$$

for almost every $x\in\Omega$ and every $(y,z)\in\mathbb{R}\times\mathbb{R}^n$ and some $\Lambda>0$, $0\leq a_1\in\mathrm{L}^1(\Omega)$ and $0\leq a_2\in\mathrm{L}^{p'}(\Omega)$.

Theorem 5.11. *Let u be a bounded weak solution of (5.4) where B satisfies the growth assumption*

$$|B(x,y,z)|\leq\Lambda|z|^p+a_3$$

for almost every $x\in\Omega$, $(y,z)\in\mathbb{R}\times\mathbb{R}^n$, some $\Lambda>0$ and $0\leq a_3\in\mathrm{L}^1(\Omega)$. Then u is a quasi-minimiser of

$$\mathscr{H}[u]=\int_\Omega|\mathrm{D}u|^p+a(x)\,\mathrm{d}x,$$

where $a=a_1+a_2^{p'}+a_3$.

Proof. Set $M=\|u\|_{\mathrm{L}^\infty(\Omega)}$. Let $v\in\mathrm{W}^{1,p}(\Omega)$ with $K=\operatorname{supp}\ u-v\Subset\Omega$ and $|v|\leq M$. Set $\phi=(u-v)^+e^{\lambda(u-v)}$ where $\lambda>0$ will be determined at a later stage and $u^+=\min(0,u)$. Write $S=\operatorname{supp}\ \phi$. Note that on S, $(u-v)^+=(u-v)$. Then testing (5.4) with ϕ we find

$$\int_S A^i\mathrm{D}_iu[1+\lambda(u-v)]e^{\lambda(u-v)}\,\mathrm{d}x=\int_S A^i\mathrm{D}_iv[1+\lambda(u-v)]e^{\lambda(u-v)}\,\mathrm{d}x$$

$$+\int_S B(u-v)e^{\lambda(u-v)}\,\mathrm{d}x.$$

Using our coercivity and growth assumptions on A and B, we deduce, noting $u-v \geq 0$ on S,

$$\begin{aligned}
&\int_S |\mathrm{D}u|^p[1+\lambda(u-v)]e^{\lambda(u-v)}\,\mathrm{d}x \\
&\leq \int_S a_1[1+\lambda(u-v)]e^{\lambda(u-v)}\,\mathrm{d}x + \int_S (\Lambda|\mathrm{D}u|^{p-1}+a_2)[1+\lambda(u-v)]e^{\lambda(u-v)}|\mathrm{D}v|\,\mathrm{d}x \\
&\qquad + \int_S \Lambda(|\mathrm{D}u|^p + a_3)(u-v)e^{\lambda(u-v)}\,\mathrm{d}x \\
&\leq \int_S c(M,\lambda)a_1 + c(M,\lambda)(\Lambda|\mathrm{D}u|^{p-1}+a_2)|\mathrm{D}v| + a_3 c(M,\Lambda) \\
&\qquad + \Lambda|\mathrm{D}u|^p(u-v)e^{\lambda(u-v)}\,\mathrm{d}x.
\end{aligned}$$

To obtain the last line, we used that $|u|, |v| \leq M$. Choosing $\lambda = 2\Lambda$, we may absorb the last term on the left-hand side. Applying Young's inequality with ε to the second term on the right-hand side, we deduce after adding a to both sides,

$$\int_S |\mathrm{D}u|^p + a\,\mathrm{d}x \leq c(M,\Lambda)\int_S |\mathrm{D}v|^p + a\,\mathrm{d}x \tag{5.6}$$

Repeating the same argument with $\tilde{\phi} = (v-u)^+ e^{\lambda(v-u)}$, we deduce

$$\int_{\operatorname{supp}\,\tilde{\phi}} |\mathrm{D}u|^p + a\,\mathrm{d}x \leq c(M,\Lambda)\int_{\operatorname{supp}\,\tilde{\phi}} |\mathrm{D}v|^p + a\,\mathrm{d}x. \tag{5.7}$$

Adding (5.6) and (5.7) gives the desired estimate

$$\int_\Omega |\mathrm{D}u|^p + a\,\mathrm{d}x \leq c(M,\Lambda)\int_\Omega |\mathrm{D}v|^p + a\,\mathrm{d}x.$$

for any $v \in \mathrm{W}^{1,p}(\Omega)$ for which $|v| \leq M$. If this does not hold, we consider $\tilde{v} = \min(M, \max(v, -M))$. Note that $|\mathrm{D}\tilde{v}| \leq |\mathrm{D}v|$ and $|\tilde{v}| \leq M$. Thus

$$\mathscr{H}[u] \leq c(m,\Lambda)\mathscr{H}[\tilde{v}] \leq c(m,\Lambda)\mathscr{H}[v],$$

which is exactly the desired quasi-minimality of u. $\square$

Remark 5.12. *The boundedness assumption in Theorem 5.11 is necessary. We will show later that in the scalar case $m = 1$, quasi-minimisers are $C^{0,\alpha}$ and in particular bounded. However, solutions to the Euler-Lagrange equation need not be bounded as the following example due to Frehse shows: It can be checked that the function $u = 12\log\log|x|^{-1}$ is a weak solution to the Euler-Lagrange equation for the functional*

$$\mathscr{F}[u] = \int_D 1 + \frac{1}{1+e^u\log(|x|)^{-12}}|\mathrm{D}u|^2\,\mathrm{d}x,$$

where $D \subset \mathbb{R}^2$ is the disc of radius e^{-1}, but u is evidently not bounded.

We close this section by mentioning two further examples where quasi-minimality naturally arises.

Example 5.13 (Quasi-regular maps)**.** *In complex analysis, a central role is played by quasi-conformal maps, that is maps which map circles into ellipses with bounded eccentricity. An important generalisation are so-called quasi-regular maps.*

Definition 5.14. *$u\colon \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ is quasi-regular if there exists $A > 0$ such that*

$$|\mathrm{D}u|^n \leq A \det(\mathrm{D}u).$$

Quasi-regular maps naturally fall into our framework of quasi-minimality.

Theorem 5.15. *A quasi-regular map $u \in \mathrm{W}^{1,n}(\Omega, \mathbb{R}^n)$ is a quasi-minimiser of the functional $u \to \int_\Omega |\mathrm{D}u|^n \,\mathrm{d}x$.*

Proof. Suppose u is quasi-regular and let $\phi \in \mathrm{W}^{1,n}(\Omega, \mathbb{R}^n)$ with $K = \operatorname{supp} \phi \Subset \Omega$. Then using the divergence theorem,

$$\int_K \det(\mathrm{D}u) \,\mathrm{d}x = \int_K \det(\mathrm{D}u + \mathrm{D}\phi) \,\mathrm{d}x \leq C \int_K |\mathrm{D}(u + \phi)|^n \,\mathrm{d}x.$$

To obtain the last inequality, note that the determinant is a sum of products of n components of the partial derivative of u. The result now follows directly from the definition of quasi-regularity. ☐

Example 5.16 (Obstacle problems)**.** *Let $f\colon \Omega \times \mathbb{R}^{m\times n}$ with $|z|^p \leq f(x,z) \leq \Lambda(1+|z|^p)$ for almost every $x \in \Omega$, all $z \in \mathbb{R}^{m\times n}$ and some $\Lambda > 0$. Let $\psi \in \mathrm{W}^{1,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$. Suppose $u \in \mathrm{W}^{1,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ is such that $u \geq \psi$ almost everywhere in Ω and whenever for some $w \in \mathrm{W}^{1,p}_{\mathrm{loc}}(\Omega)$, $K = \operatorname{supp}(u - w) \Subset \Omega$ and $w \geq \psi$ almost everywhere in Ω, we have*

$$\mathscr{F}[u, K] = \int_\Omega f(x, \mathrm{D}u) \,\mathrm{d}x \leq \mathscr{F}[w, K].$$

In other words, u is a local minimiser of the problem $\min_{u \in \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)} \int f(x, \mathrm{D}u) \,\mathrm{d}x$ under the additional obstacle constraint $u \geq \psi$.

Let $v \in \mathrm{W}^{1,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ with $K = \operatorname{supp}\ (u - v) \Subset \Omega$. Then, denote $\Sigma = \{x \in \Omega\colon v \geq \psi\}$ and set $w = \max(v, \psi) \in \mathrm{W}^{1,p}(\Omega)$. Noting $w \geq \psi$ almost everywhere in Ω, we find

$$\mathscr{F}[u, K] \leq \mathscr{F}[w, K] = \mathscr{F}[w, K \cap \Sigma] + \mathscr{F}[\psi, K \setminus \Sigma] \leq \mathscr{F}[w, K] + \mathscr{F}[\psi, K].$$

In particular, with $\gamma = f(x, \mathrm{D}\psi)$,

$$\int_K f(x, \mathrm{D}u) + \gamma(x) \,\mathrm{d}x \leq 2 \int f(x, \mathrm{D}v) + \gamma(x) \,\mathrm{d}x.$$

Thus u is a 2-quasi-minimiser of the functional $u \to \int_\Omega f(x, \mathrm{D}u) + \gamma(x) \,\mathrm{d}x$. Note in particular, that this is a functional satisfying the growth conditions

$$|z|^p \leq f(x,z) + \gamma(x) \leq \Lambda(1 + |z|^p) + \Lambda(1 + |\mathrm{D}\psi|^p)$$

and $\Lambda(1 + |\mathrm{D}\psi|^p) \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$, so that this modified energy essentially remains in the same class of energies.

5.3 Caccioppoli's inequality

The aim of this section is to prove a version of Cacciopolli's inequality. This is the basic inequality we have available for elliptic problems and in the case of systems, it is one of very few tools that exist. In combination with Sobolev embedding, it allows to control the gradient in L^p by the gradient on a bigger ball in L^q where $q < p$. Hence, this type of inequality is also known as reverse Hölder inequality.

Theorem 5.17. *Assume $u \in \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$ is a Q-quasi-minimiser for the functional $\mathscr{F}[u,\Omega] = \int_\Omega f(x,u,\mathrm{D}u)\,\mathrm{d}x$, where we assume that*

$$\begin{aligned} f(x,y,z) &\geq |z|^p - \theta(x,u)^p, \\ |f(x,y,z)| &\leq \Lambda(|z|^p + \theta(x,u)^p), \end{aligned}$$

where $\theta(x,u)^p = b(x)|u|^\gamma + a(x)$ with $1 \leq \gamma < p^*$ and $0 \leq b \in \mathrm{L}^{\frac{p^*}{p^*-\gamma}}(\Omega)$, $0 \leq a \in \mathrm{L}^1(\Omega)$. Then there exists R_0 depending only on u such that for $R < R_0$ and $Q_R \Subset \Omega$,

$$\int_{Q_{R/2}} (|\mathrm{D}u|^p + |u|^{p^*})\,\mathrm{d}x \leq c\left(\frac{1}{R^p}\int_{Q_R} |u-u_R|^p\,\mathrm{d}x + |Q_R|\left(\fint_{Q_R} |u|\,\mathrm{d}x\right)^{p^*} + \int_{Q_R} g\,\mathrm{d}x\right).$$

Here $u_s = \fint_{Q_s} u\,\mathrm{d}\ x$ and $g = a + b^{\frac{p^*}{p^*-\gamma}}$. Moreover, with $m = \frac{p}{p+n} < 1$,

$$\fint_{Q_{R/2}} (|\mathrm{D}u|^p + |u|^{p^*})\,\mathrm{d}x \leq c\left(\left(\fint_{Q_R} (|\mathrm{D}u|^p + |u|^{p^*})^m\,\mathrm{d}x\right)^{\frac{1}{m}} + \fint_{Q_R} g\,\mathrm{d}x\right).$$

Proof. The moreover part is a direct consequence of employing Poincarés inequality and Hölder's inequality to the first inequality. Thus, we focus on the first inequality. For simplicity, we only describe the case $a = b = 0$, leaving the general case for the exercise class.

Let $Q_R \Subset \Omega$, $R/2 < t < s \leq R$ and take η to be a cut-off such that $\eta = 1$ in Q_t, $0 \leq \eta \leq 1$, $\operatorname{supp}\eta \subset Q_s$ and $|\mathrm{D}\eta| \leq \frac{2}{s-t}$. Set $\phi = \eta(u-u_s)$. Then,

$$\int_{Q_R} |\mathrm{D}\phi|^p\,\mathrm{d}x \leq \int_{Q_R} f(x,u,\mathrm{D}\phi) - f(x,u,\mathrm{D}u)\,\mathrm{d}x + \int_{Q_R} f(x,u,\mathrm{D}u)\,\mathrm{d}x.$$

Setting $v = u - \phi = u_s + (1-\eta)(u-u_s)$, we use quasi-minimality to deduce

$$\begin{aligned} \int_{Q_R} f(x,u,\mathrm{D}u)\,\mathrm{d}x \leq& Q\int_{Q_R} f(x,v,\mathrm{D}v)\,\mathrm{d}x \lesssim \int_{Q_R} |\mathrm{D}v|^p\,\mathrm{d}x \\ &\lesssim \int_{Q_R} (1-\eta)^p|\mathrm{D}u|^p + |u-u_s|^p\frac{1}{(s-t)^p}\,\mathrm{d}x. \end{aligned}$$

Recalling that $\eta = 1$ in Q_t, we have

$$\int_{Q_R} f(x,u,\mathrm{D}u)\,\mathrm{d}x \leq c\left(\int_{Q_s\setminus Q_t} |\mathrm{D}u|^p\,\mathrm{d}x + \frac{1}{(s-t)^p}\int_{Q_s} |u-u_s|^p\,\mathrm{d}x\right).$$

Further, note that $\mathrm{D}\phi = \mathrm{D}u$ in $Q_s \setminus Q_t$, so that

$$\begin{aligned} \int_{Q_s} f(x,u,\mathrm{D}\phi) - f(x,u,\mathrm{D}u)\,\mathrm{d}x &\lesssim \int_{Q_s\setminus Q_t} |\mathrm{D}\phi|^p + |\mathrm{D}u|^p\,\mathrm{d}x \\ &\lesssim \int_{Q_s\setminus Q_t} |\mathrm{D}v|^p + |\mathrm{D}u|^p\,\mathrm{d}x. \end{aligned}$$

Estimating the $\mathrm{D}v$ term as in our estimate for $\int_{Q_R} f(x,u,\mathrm{D}u)$, we have shown

$$\int_{Q_t} |\mathrm{D}u|^p\,\mathrm{d}x \leq c\int_{Q_s\setminus Q_t} |\mathrm{D}u|^p\,\mathrm{d}x + \frac{c}{(s-t)^p}\int_{Q_s} |u-u_s|^p\,\mathrm{d}x.$$

Adding the term $c\int_{Q_t}|\mathrm{D}u|^p\,\mathrm{d}x$ to both sides and re-arranging, we obtain

$$\int_{Q_t}|\mathrm{D}u|^p\,\mathrm{d}x \le \frac{c}{1+c}\int_{Q_s}|\mathrm{D}u|^p\,\mathrm{d}x + \frac{c}{(c+1)(s-t)^p}\int_{Q_s}|u-u_s|^p\,\mathrm{d}x.$$

Using an iteration argument (which we consider in Lemma 5.18, right after this proof) we conclude

$$\int_{Q_{R/2}}|\mathrm{D}u|^p\,\mathrm{d}x \lesssim \frac{1}{R^p}\int_{Q_R}|u-u_s|^p\,\mathrm{d}x.$$

□

We now turn to proving the iteration inequality, we employed in the proof of Theorem 5.17.

Lemma 5.18. *Let $Z(t)$ be a bounded non-negative function in $[\rho,R]$. Assume for any $\rho\le t<s\le R$,*

$$Z(t)\le\theta Z(s)+A(s-t)^{-\alpha}+B(s-t)^{-\beta}+C$$

for some $\theta\in[0,1)$, $A,B,C\ge 0$ and $\alpha>\beta>0$. Then

$$Z(\rho)\le c(\alpha,\theta)\left(A(R-\rho)^{-\alpha}+B(R-\rho)^{-\beta}+C\right).$$

Proof. Set $t_0=\rho$ and define $\{t_i\}$ by requiring

$$t_{i+1}-t_i=(1-\lambda)\lambda^i(R-\rho)$$

for some $0<\lambda<1$. Note $t_i\to R$ as $i\to\infty$. By induction, we find

$$Z(\rho)\le\theta^k Z(t_k)+\left(\frac{A}{(1-\lambda)^\alpha(R-\rho)^\alpha}+\frac{B}{(1-\lambda)^\beta(R-\rho)^\beta}+C\right)\sum_{i=0}^{k-1}\theta^i\lambda^{-i\alpha}.$$

Choosing λ sufficiently small that $\theta\lambda^{-\alpha}<1$, we may let $k\to\infty$ to prove the lemma. □

5.4 Gehring's inequality

Setting $f=|\mathrm{D}u|+|u|^{p^*}$, we may rewrite the outcome of Theorem 5.17 as

$$⨍_{Q_R} f\,\mathrm{d}x \lesssim \left(⨍_{Q_{2R}} f^m\,\mathrm{d}x\right)^{\frac{1}{m}} + ⨍_{Q_{2R}} g\,\mathrm{d}x.$$

It is an observation due to Gehring, that such reverse Hölder-type inequalities are not sharp in the exponents. Instead, they self-improve, that is, the L^1-norm of f on the left-hand side may be replaced by $\mathrm{L}^{1+\delta}$ for some $\delta>0$. Returning to the setting of Theorem 5.17, this tells us that quasi-minimisers are $\mathrm{W}^{1,p+\delta}$-regular for some $\delta>0$.

Before proceeding with the proof of this, we require a central result in Calderon-Zygmund theory. Given a function f, it allows to decompose a domain into cubes in such a way that on each cube the average of f is controlled, while outside of the cubes f is bounded.

Theorem 5.19. *Let Q_0 be a cube in $\mathbb{R}^n$ and $0 \le g \in \mathrm{L}^1(Q_0)$. Let L be such that*

$$\fint_{Q_0} g \,\mathrm{d}x \le L.$$

Then there exists a countable family of cubes $\{Q_i\}$ of pairwise disjoint cubes $Q_i \subset Q_0$ with faces parallel to Q_0 such that

$$L \le \fint_{Q_i} g \,\mathrm{d}x \le 2^n L,$$
$$g \le L \text{ almost everywhere in } Q \setminus \cup Q_i.$$

Proof. We call a cube $Q \subset Q_0$ final if $\fint_Q g \,\mathrm{d}x > L$. Divide Q_0 into 2^n equal subcubes, each with side length one half of that of Q_0. If any of these cubes are not final, subdivide them again. Continuing in this way, by Zorn's lemma, we obtain a countable family $\{Q_i\}$ of final cubes. Note that obviously $\fint_{Q_i} g \,\mathrm{d}x \ge L$. However, each Q_i is contained in some Q of twice the side-length, where Q is not final. Thus

$$\fint_{Q_i} g \,\mathrm{d}x \le 2^n \fint_Q g \,\mathrm{d}x \le 2^n L.$$

If $x \in Q \setminus \cup Q_i$, then there exists a decreasing family of cubes $\tilde{Q}_i$, none of which is final, such that $x \in \cap \tilde{Q}_i$. Then

$$\fint_{\tilde{Q}_i} g \,\mathrm{d}x \le L.$$

Since almost every $x \in \Omega$ is a Lebesgue-point, we may pass to the limit for almost every x and conclude $g(x) \le L$. $\square$

Having Theorem 5.19 at hand, we can proceed to prove Gehring's estimate. By a re-scaling, it suffices to consider $Q = Q_1(x_0)$. Set $d(x) = d(x, \partial Q)$ and define the concentric shells

$$C_k = \left\{ x \in Q \colon \frac{3}{4} 2^{-k-1} \le d(x) \le \frac{3}{4} 2^{-k} \right\}$$

Each shell C_k can be divided into a finite family $\mathscr{G}_k$ of equal cubes of side $\delta_k = \frac{3}{4}2^{-k-1}$. Note that $Q = \cup_{G \in \mathscr{G}_k} G \cup Q_{1/4}$. If P is a cube, we denote by $\tilde{P}$, the cube concentric to P, but with twice the side length. Then our main assumption is

$$\fint_P f \,\mathrm{d}x \le c \left(\left(\fint_{\tilde{P}} f^m \right)^{\frac{1}{m}} + \fint_{\tilde{P}} g \right) \tag{5.8}$$

Setting $F = d(x)^n f$ and $G = d(x)^n g$ and noting that if $P \subset C_k$, $\tilde{P} \Subset Q$, we find for such P,

$$\fint_P F \,\mathrm{d}x \le c \left(\left(\fint_{\tilde{P}} F^m \,\mathrm{d}x \right)^{\frac{1}{m}} + \fint_{\tilde{P}} G \,\mathrm{d}x \right). \tag{5.9}$$

The key observation is contained in the next lemma, where we estimate the contribution to $\fint_Q F$ from the level sets where F is large.

Lemma 5.20. *For every $t > t_0 = \fint_Q f \,\mathrm{d}x$, setting*

$$\Phi_t = \{x \in Q \colon F(x) > t\}, \qquad \Gamma_t = \{x \in Q \colon G(x) > t\},$$

it holds that

$$\int_{\Phi_t} F \,\mathrm{d}x \le c \left(t^{1-m} \int_{\Phi_t} F^m \,\mathrm{d}x + \int_{\Gamma_t} G \,\mathrm{d}x \right).$$

Proof. Let $s = \lambda t$ for some $\lambda > 0$ to be determined. For $P \in \mathscr{G}_k$, we find

$$s > \lambda \fint_Q f \,\mathrm{d}x \ge \lambda \frac{|P|}{|Q|} \fint_P f \,\mathrm{d}x \ge \lambda 4^{-n} \fint_P F \,\mathrm{d}x \ge \fint_P F \,\mathrm{d}x,$$

if we ensure $\lambda \ge 4^n$. The same inequality evidently also holds for $P = Q_{1/4}$.

Applying Theorem 5.19 to each such cube P we obtain a countable family $\{Q_i\}$ of disjoint subcubes of Q such that

$$s < \fint_{Q_j} F \,\mathrm{d}x \le 2^n s,$$
$$F(x) \le s \text{ in } Q \setminus \cup Q_i.$$

From (5.9) we learn that either

$$\fint_{Q_i} F \,\mathrm{d}x \le 2c \left(\fint_{\tilde{Q}_i} F^m \,\mathrm{d}x \right)^{\frac{1}{m}}$$

or

$$\fint_{Q_i} F \,\mathrm{d}x \le 2c \fint_{\tilde{Q}_i} G \,\mathrm{d}x.$$

In the first case,

$$s \le 2c \left(\fint_{\tilde{Q}_i} F^m \,\mathrm{d}x \right)^{\frac{1}{m}}.$$

Moreover,

$$\int_{\tilde{Q}_i} F^m \,\mathrm{d}x \le \int_{\tilde{Q}_i \cap \Phi_t} F^m \,\mathrm{d}x + t^m |\tilde{Q}_i|.$$

Thus, we deduce, combining estimates,

$$|\tilde{Q}_i| \le 2(2c)^m s^{-m} \int_{\tilde{Q}_i \cap \Phi_t} F^m \,\mathrm{d}x.$$

A similar argument in the second case, allows us to deduce, that in any case,

$$|\tilde{Q}_i| \le \frac{c}{s} \left(t^{1-m} \int_{\tilde{Q}_j \cap \Phi_t} F^m \,\mathrm{d}x + \int_{\tilde{Q}_j \cap \Gamma_t} G \,\mathrm{d}x \right).$$

We now consider $\int_{\Phi_s} F \,\mathrm{d}x$,

$$\int_{\Phi_s} F \,\mathrm{d}x \le \sum_{j=1}^{\infty} \int_{Q_j} F \,\mathrm{d}x \le 2^n s \sum_{j=1}^{\infty} |Q_j| \le 2^n s |\cup_{j=1}^{\infty} \tilde{Q}_j|.$$

Note, that if we naively estimate the size of $\cup\tilde{Q}_j$, this is not summable. Thus, we need to choose a good covering of $\tilde{Q}_j$. By means of Vitali covering lemma, we obtain a countable subfamily $\{\Pi_i\} \subset \{Q_i\}$ such that

$$\cup\tilde{Q}_i \subset \cup\hat{\Pi}_i.$$

where $\hat{\Pi}_i$ denotes the cube concentric with Π_i but with quintuple side length. Now

$$|\cup\tilde{Q}_i| \leq 5^n \sum_{i=1}^{\infty} |\Pi_i|.$$

Inserting this into our estimate we have obtained

$$\int_{\Phi_s} F \,\mathrm{d}x \leq c\left(t^{1-m}\int_{\Phi_t} F^m \,\mathrm{d}x + \int_{\Gamma_t} G \,\mathrm{d}x\right).$$

As also,

$$\int_{\Phi_t\setminus\Phi_s} F \,\mathrm{d}x \leq s^{1-n}\int_{\Phi_t} F^m \,\mathrm{d}x \leq ct^{1-m}\int_{\Phi_t} F^m \,\mathrm{d}x,$$

this concludes the proof. □

We record without proof an elementary lemma, which follows from a layer-cake type argument:

Lemma 5.21. *Let $h \geq m \geq 0$ and $F \in \mathrm{L}^h(Q)$. Set $\phi(t) = \int_{\Phi_t} F^m \,\mathrm{d}x$. Then*

$$\int_{\Phi_\tau} F^h \,\mathrm{d}x = -\int_{\tau}^{\infty} t^{h-m} \,\mathrm{d}\phi(t).$$

In light of Lemma 5.21, the outcome of Lemma 5.20 may be written as

$$-\int_{t}^{\infty} \tau^{1-m} \,\mathrm{d}\phi(\tau) \leq A(t^{1-m}\phi(t) + \omega(t)), \tag{5.10}$$

where $\omega(t) = \int_{\Gamma_t} G \,\mathrm{d}x$. Gehring's famous result applies in this setting:

Theorem 5.22 (Gehring's lemma)**.** *Assume $\phi(t)$ is a decreasing function in $[a, +\infty)$, infinitessimal for $t \to \infty$ and satisfying* (5.10) *with $m < 1$ for every $t \geq a$. Then there is $r > 1$ such that*

$$-\int_{a}^{\infty} u^{r-m} \,\mathrm{d}\phi(u) \leq -2a^{r-1}\int_{a}^{\infty} u^{1-m} \,\mathrm{d}\phi(u) - 2A\int_{a}^{\infty} u^{r-1} \,\mathrm{d}\omega(u).$$

Proof. We first assume that $\phi(s) = 0$ and $\omega(s) = 0$ for $s \geq k-1$. For $q > 0$, set $I_q(s) = -\int_s^k u^q \,\mathrm{d}\phi(u)$, $I_q = I_q(a)$ and $\Omega_q = -\int_a^k u^q \,\mathrm{d}\omega(u)$.

Using an integration by parts, we have

$$I_{r-m} = -\int_{a}^{k} u^{r-1}u^{1-m} \,\mathrm{d}\phi(u) = a^{r-1}I_{1-m} + (r-1)\int_{a}^{k} u^{r-2}I_{1-m}(u) \,\mathrm{d}u.$$

The last integral can be estimated using (5.10) and an integration by parts, giving

$$I_{r-m} \leq a^{r-1}I_{1-m} + A(r-1)\left(\int_{a}^{k} u^{r-m-1}\phi(u) \,\mathrm{d}u + \int_{a}^{k} u^{r-2}\omega(u) \,\mathrm{d}u\right)$$

$$
\begin{aligned}
=&a^{r-1}I_{1-m}+A(r-1)\left(\frac{I_{r-m}}{r-m}-\frac{a^{r-m}}{r-m}\phi(a)+\int_a^k u^{r-2}\omega(u)\,\mathrm{d}u\right)\\
\leq&a^{r-1}I_{1-m}+\frac{A(r-1)I_{r-m}}{r-m}+A\Omega_{r-1}.
\end{aligned}
$$

To obtain the last line, we estimated the ω-term using the same integration by parts argument we employed to estimate the u^{r-m-1}-term. If we take r sufficiently close to 1 that $A(r-1)\leq\frac{r-m}{2}$, this concludes the proof.

We now turn to the general case. Note that

$$
-\int_k^T s^{1-m}\,\mathrm{d}\phi 9s)\geq -k^{1-m}\int_k^T\,\mathrm{d}\phi(s)=-k^{1-m}\left(\phi(k)-\phi(T)\right).
$$

Letting $T\to\infty$, we deduce

$$
-\int_k^\infty s^{1-m}\,\mathrm{d}\phi(s)\geq -k^{1-m}\phi(k).
$$

Set $\phi_k(t)=1_{t\leq k}\phi(t)$ (and the same for $\omega_k(t)$). Then for $t\leq k$,

$$
\begin{aligned}
-\int_t^\infty s^{1-m}\,\mathrm{d}\phi_k(s)=&-\int_t^k s^{1-m}\,\mathrm{d}\phi(s)+k^{1-m}\phi(k)\\
\leq&-\int_t^\infty s^{1-m}\,\mathrm{d}\phi(s)\leq A(t^{1-m}\phi_k(t)+\omega_k(t)).
\end{aligned}
$$

Note this inequality obviously still holds if $t>k$. Thus, we may estimate

$$
\begin{aligned}
-\int_a^\infty s^{r-m}\,\mathrm{d}\phi_k(s)\leq&-2a^{r-1}-\int_a^\infty s^{1-m}\,\mathrm{d}\phi_k(s)-2A\int_a^\infty s^{r-1}\,\mathrm{d}\omega_k(s)\\
\leq&-2a^{r-1}-\int_a^\infty s^{1-m}\,\mathrm{d}\phi(s)-2A\int_a^\infty s^{r-1}\,\mathrm{d}\omega(s).
\end{aligned}
$$

Letting $k\to\infty$, the result follows. $\square$

Returning to f and g, Theorem 5.22 translates to the following statement:

Theorem 5.23. *Let $f\in\mathrm{L}^1(Q_R)$ and assume that for every cube $Q\subset\tilde{Q}\Subset Q_R$, we have*

$$
\fint_Q f\,\mathrm{d}x\leq c\left(\left(\fint_{\tilde{Q}} f^m\,\mathrm{d}x\right)^{\frac{1}{m}}+\fint_{\tilde{Q}} g\,\mathrm{d}x\right)
$$

for some $0<m<1$. Assume $g\in\mathrm{L}^s(Q_R)$ for some $s>1$. Then there is $r>1$ such that $f\in\mathrm{L}^r(Q_{R/2})$ and

$$
\fint_{Q_{R/2}} f^r\,\mathrm{d}x\leq c\left(\left(\fint_{Q_R} f\,\mathrm{d}x\right)^r+\fint_{Q_R} g^r\,\mathrm{d}x\right).
$$

Combining Theorem 5.17 and Theorem 5.23, we obtain the following regularity statement.

Theorem 5.24. *Let $u\colon\Omega\to\mathbb{R}^m$ be a quasi-minimiser for the functional*

$$
\mathscr{F}[u,\Omega]=\int_\Omega F(x,u,\mathrm{D}u)\,\mathrm{d}x
$$

and assume that the hypothesis of Theorem 5.17 are satisfied, in particular

$$|z|^p - \theta(x,u)^p \leq F(x,y,z) \leq c(|z|^p + \theta(x,u)^p)$$

where $\theta(x,u)^p = a(x) + |u|^\gamma b$, $1 \leq \gamma \leq p^*$, $0 \leq a \in \mathrm{L}^s$ *for some* $s > 1$ *and* $0 \leq b \in \mathrm{L}^\sigma$ *for some* $\sigma > \frac{p^*}{p^*-\gamma}$.

Then there is $r > 1$ *such that for every* $Q_r \subset Q_{2R} \Subset \Omega$,

$$\fint_{Q_R} (|\mathrm{D}u|^p + |u|^{p^*})^r \,\mathrm{d}x \leq c\left(\left(\fint_{Q_{2R}} |\mathrm{D}u|^p + |u|^{p^*} \,\mathrm{d}x\right)^r + \fint_{Q_{2R}} g^r \,\mathrm{d}x\right).$$

where $g = a + b^{\frac{p^*}{p^*-\gamma}}$.

5.5 Hölder regularity

Our next goal is to show Hölder regularity of quasi-minimisers in the scalar setting $m = 1$. The key tool will once again be a Caccioppoli inequality. The main difference is that we want to incorporate information about the level sets of u. It will be useful to denote

$$A(k,R) = \{x \in Q_R \colon u > k\}, \qquad B(k,R) = \{x \in Q_R \colon u \leq k\}.$$

Note that for almost every R, $|Q_R| = |A(k,R)| + |B(k,R)|$. We will always choose R such that this equality holds.

Theorem 5.25. *Assume* $u \in \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$ *is a Q-quasi-minimiser for the functional* $\mathscr{F}[u,\Omega] = \int_\Omega f(x,u,\mathrm{D}u)$, *where we assume that*

$$f(x,y,z) \geq |z|^p - \theta(x,u)^p,$$
$$|f(x,y,z)| \leq \Lambda(|z|^p + \theta(x,u)^p).$$

Here $\theta(x,u)^p = b(x)|u|^\gamma + a(x)$ *with* $1 \leq \gamma < p^*$ *and* $0 \leq b \in \mathrm{L}^\sigma(\Omega)$, $0 \leq a \in \mathrm{L}^s(\Omega)$ *and for some* $\varepsilon > 0$, $\frac{1}{s} = \frac{p}{n} - \varepsilon$, $\frac{1}{\sigma} = 1 - \frac{\gamma}{p^*} - \varepsilon$. *Then there exists a radius* $R_0 = R_0(|u|_{\mathrm{L}^{p^*}(\Omega)}, |b|_{\mathrm{L}^\sigma(\Omega)})$ *such that for all* $x_0 \in \Omega$, $0 < \rho < R < \min(R_0, d(x_0, \partial\Omega)$ *and* $k \geq 0$,

$$\int_{A(k,\rho)} |\mathrm{D}u|^p \,\mathrm{d}x \leq \frac{C}{(R-\rho)^p} \int_{A(k,R)} (u-k)^p \,\mathrm{d}x + c(|a|_{\mathrm{L}^s(\Omega)} + k^p R^{-n\varepsilon})|A(k,R)|^{1-\frac{p}{n}+\varepsilon}.$$

Proof. We consider only the case $a = b = 0$. The full argument is an exercise on Sheet 4. Let $\eta \in C_0^\infty(Q_R)$ such that $\eta = 1$ on Q_ρ and $|\mathrm{D}\eta| \leq \frac{2}{R-\rho}$. Set $v = u - \eta(u-k)^+$. Note that $v = u$ on $Q_R \setminus A(k,R)$ and $v = u$ on $\partial\Omega$. Using quasi-minimality, we obtain

$$\mathscr{F}[u, A(k,R)] \leq Q\mathscr{F}[v, A(k,R)].$$

Using our growth conditions and noting that on $A(k,R)$

$$|\mathrm{D}v| \leq (1-\eta)|\mathrm{D}u| + |\mathrm{D}\eta|(u-k),$$

we get

$$\int_{A(k,R)} |\mathrm{D}u|^p \,\mathrm{d}x \lesssim \int_{A(k,R)} (1-\eta)^p |\mathrm{D}u|^p \,\mathrm{d}x + \frac{1}{(R-\rho)^p} \int_{A(k,R)} (u-k)^p \,\mathrm{d}x$$

$$\leq \int_{A(k,R)\setminus A(k,\rho)} |\mathrm{D}u|^p \,\mathrm{d}x + \frac{1}{(R-\rho)^p} \int_{A(k,R)} (u-k)^p \,\mathrm{d}x.$$

Re-arranging and using the hole-filling trick, we already employed in the proof of Theorem 5.17, we deduce for some $\theta < 1$,

$$\int_{A(k,\rho)} |\mathrm{D}u|^p \,\mathrm{d}x \leq \theta \int_{A(k,R)} |\mathrm{D}u|^p \,\mathrm{d}x + \frac{C(\theta)}{(R-\rho)^p} \int_{A(k,R)} (u-k)^p \,\mathrm{d}x.$$

Applying Lemma 5.18, we deduce the desired inequality. □

It is a remarkable observation due to de Giorgi, that the inequality in Theorem 5.25 contains essentially all the information about u. This inspires the introduction of the so-called De-Giorgi classes, functions satisfying precisely this estimate.

Definition 5.26. *Let $u \in \mathrm{W}^{1,p}_{\mathrm{loc}}(\Omega)$. We say $u \in DG_p^+ = DG_p^+(\Omega, H, \xi, \varepsilon, R_0, k_0)$ if for all concentric cubes $Q_\rho \subset Q_R \Subset \Omega$, $R < R_0$ and $k \geq k_0 \geq 0$, we have*

$$\int_{A(k,\rho)} |\mathrm{D}u|^p \,\mathrm{d}x \leq \frac{H}{(R-\rho)^p} \int_{A(k,R)} (u-k)^p \,\mathrm{d}x + H(\xi^p + k^p R^{-n\varepsilon}) |A(k,R)|^{1-\frac{p}{n}+\varepsilon}.$$

We say $u \in DG_p^-$ if $-u \in DG_p^+$ and $u \in DG_p$ if $u \in DG_p^+ \cap DG_p^-$.

Remark 5.27. *Given $u \in DG_p^+$, for our proofs it will be convenient to simplify the defining inequality by introducing variants of u. Setting $v = u + \xi R^\beta$ where $\beta = \frac{n\varepsilon}{p}$ and $h = k + \xi R^\beta$, we see that*

$$\int_{A(h,\rho)} |\mathrm{D}v|^p \,\mathrm{d}x \leq \frac{H}{(R-\rho)^p} \int_{A(h,R)} (v-h)^p \,\mathrm{d}x + H h^p R^{-n\varepsilon} |A(h,R)|^{1-\frac{p}{n}+\varepsilon}.$$

Moreover, we can normalise to the situation where $R = 1$ by setting for $s < r < R$, $y = Rx$, $s = \sigma R$, $t = \tau R$. With $w(x) = v(y)$, it holds that

$$\int_{A(h,\sigma)} |\mathrm{D}w|^p \,\mathrm{d}x \leq \frac{H}{(\tau-\sigma)^p} \int_{A(h,\tau)} (w-h)^p \,\mathrm{d}x + H h^p \tau^{-n\varepsilon} |A(h,\tau)|^{1-\frac{p}{n}+\varepsilon}.$$

If $\tau \geq \frac{1}{2}$, we deduce for some $H_1 > 0$,

$$\int_{A(h,\sigma)} |\mathrm{D}w|^p \,\mathrm{d}x \leq \frac{H_1}{(\tau-\sigma)^p} \int_{A(h,\tau)} (w-h)^p \,\mathrm{d}x + H_1 h^p |A(h,\tau)|^{1-\frac{p}{n}+\varepsilon}.$$

Our first result is to show that if $u \in DG_p^+$, then u is bounded above.

Theorem 5.28. *Let $u \in DG_p^+$. Then u is locally bounded above in Ω. For $x_0 \in \Omega$, $R \leq \min(R_0, d(x_0, \partial\Omega))$, it holds that*

$$\sup_{Q_{R/2}} u \lesssim \left(\fint_{Q_R} u_+^p \right)^{\frac{1}{p}} + k_0 + \xi R^\beta.$$

Proof. In light of Remark 5.27, it suffices to prove the theorem for $R = 1$ assuming that for $h \geq h_0$,

$$\int_{A(h,\sigma)} |\mathrm{D}u|^p \,\mathrm{d}x \leq \frac{H}{(\tau-\sigma)^p} \int_{A(h,\tau)} (u-h)^p \,\mathrm{d}x + H h^p |A(h,\tau)|^{1-\frac{p}{n}+\varepsilon}. \qquad (5.11)$$

For $\frac{1}{2} \leq \sigma < \tau \leq 1$, let $\eta \in C_0^\infty(Q_{\frac{\sigma+\tau}{2}})$ with $\eta = 1$ on Q_σ and $|\mathrm{D}\eta| \leq \frac{4}{\tau-\sigma}$. Setting $\xi = \eta(w-k)^+$, $k \geq h_0$, we obtain using Hölder's inequality and Sobolev embedding

$$\int_{A(k,\sigma)} (u-k)^p \,\mathrm{d}x$$

$$
\begin{aligned}
&\leq \left(\int_{A(k,\sigma)} \xi^{p^*}\,\mathrm{d}x\right)^{\frac{p}{p^*}} |A(k,\tau)|^{1-\frac{p}{p^*}} \\
&\lesssim \int |\mathrm{D}\xi|^p\,\mathrm{d}x |A(k,\tau)|^{\frac{p}{n}} \\
&\lesssim \left(\int_{A(k,\frac{\sigma+\tau}{2})} |\mathrm{D}u|^p\,\mathrm{d}x + \frac{1}{(\tau-\sigma)^p}\int_{A(k,\frac{\sigma+\tau}{2})} (u-k)^p\,\mathrm{d}x\right) |A(k,\tau)|^{\frac{p}{n}}.
\end{aligned}
$$

Employing (5.11), we deduce

$$
\int_{A(k,\sigma)} (u-k)^p\,\mathrm{d}x \lesssim \frac{|A(k,\tau)|^{\frac{p}{n}}}{(\tau-\sigma)^p}\int_{A(k,\frac{\sigma+\tau}{2})} (u-k)^p\,\mathrm{d}x + k^p|A(k,\tau)|^{1+\varepsilon}. \tag{5.12}
$$

Now note that if $h<k$,

$$
\int_{A(h,\tau)} (u-h)^p \geq (k-h)^p|A(k,\tau)|
$$

and

$$
\int_{A(k,\tau)} (u-k)^p\,\mathrm{d}x \leq \int_{A(k,\tau)} (u-h)^p\,\mathrm{d}x \leq \int_{A(h,\tau)} (u-h)^p\,\mathrm{d}x.
$$

Inserting these two inequalities in (5.11), noting $\varepsilon \leq \frac{p}{n}$ and $|A(k,\tau)| \leq |Q_1|$, we obtain

$$
\int_{A(k,\sigma)} (u-k)^p\,\mathrm{d}x \lesssim \left(\int_{A(h,\tau)} (u-h)^p\,\mathrm{d}x\right)^{1+\varepsilon} \frac{1}{(k-h)^{p\varepsilon}}\left(\frac{1}{(\tau-\sigma)^p} + \frac{k^p}{(k-h)^p}\right).
$$

Let $d \geq h_0$ be a constant to be determined. Setting $\Phi_i = d^{-p}\int_{A(k_i,\sigma_i)}(u-k_i)^p\,\mathrm{d}x$ with $k_i = 2d(1-2^{-i-1})$ and $\sigma_i = \frac{1}{2}(1+2^{-i})$, we have shown

$$
\Phi_{i+1} \lesssim 2^{ip(1+\varepsilon)}\Phi_i^{1+\varepsilon}.
$$

Applying Lemma 5.29 (which we prove directly after this proof) with $d = h_0 + c\left(\int_{Q_1} u_+^p\,\mathrm{d}x\right)^{\frac{1}{p}}$, we deduce

$$
\int_{A(k_i,\sigma)} (u-k_i)^p\,\mathrm{d}x \to 0
$$

as $k_i \to 2d$. In particular, this shows $\left|A\left(2d,\frac{1}{2}\right)\right| = 0$, in other words

$$
\sup_{Q_{1/2}} u \leq 2d.
$$

Recalling the definition of d this concludes the proof. $\square$

It remains to prove the iteration argument needed in the proof above.

Lemma 5.29. *Let $\alpha>0$ and let $\{x_i\}$ be a sequence of real positive numbers such that*

$$
x_{i+1} \leq CB^i x_i^{1+\alpha}
$$

for some $C>0$ and $B>1$. If $x_0 \leq C^{-\frac{1}{\alpha}}B^{-\frac{1}{\alpha^2}}$, we have

$$
x_i \leq B^{-\frac{1}{\alpha}}x_0.
$$

In particular, $x_i \to 0$ as $i\to\infty$.

Proof. see Exercise sheet 4. □

If $u \in DG_p$, we can apply Theorem 5.28 to u and $-u$ and cover $\Sigma \Subset \Omega$ with a finite number of cubes to obtain the following L^∞-bound:

Theorem 5.30. *If $u \in DG_p$, $\Sigma \Subset \Omega$, then*

$$\sup_\Sigma |u| \lesssim \left(\int_\Omega |u|^p \,\mathrm{d}x \right)^{\frac{1}{p}} + k_0 + \xi.$$

Combining Theorem 5.30 with Theorem 5.25, we obtain a L^∞-bound for quasi-minimisers.

Theorem 5.31. *Let $u \in \mathrm{W}^{1,p}(\Omega)$ be a quasi-minimiser of $\mathscr{F}$, where $\mathscr{F}$ satisfies the assumptions of Theorem 5.25. Then u is locally bounded in Ω. Moreover there exists $c = c(\|u\|_{\mathrm{W}^{1,p}(\Omega)}, \|b\|_{\mathrm{L}^\sigma(\Omega)})$ such that for every $\rho < R < \min(R_0, d(x_0, \partial\Omega))$,*

$$\sup_{Q_\rho} |u| \le \left(\frac{1}{(R-\rho)^n} \int_{Q_R} |u|^p \,\mathrm{d}x \right)^{\frac{1}{p}} + \|a\|_{\mathrm{L}^s(\Omega)}^{\frac{1}{p}} R^\beta.$$

Remark 5.32. *In Theorem 5.31, the L^p-norm can be replaced by a L^q-norm for any $q > 1$. If $p < q$, this is a direct consequence of Hölder's inequality. If $q < p$, set for $\rho \le \sigma < \tau \le R$, $U_\tau = \sup_{Q_\tau} |u|$. Further set $\xi = \|a\|_{\mathrm{L}^s(\Omega)}$. Then,*

$$\begin{aligned} U_\sigma &\lesssim \left(\frac{1}{(\tau-\sigma)^n} \int_{Q_\tau} |u|^p \,\mathrm{d}x \right)^{\frac{1}{p}} + \xi R^\beta \\ &\le \frac{1}{(\tau-\sigma)^n} \left(\int_{Q_\tau} |u|^q \,\mathrm{d}x \right)^{\frac{1}{p}} U_\tau^{1-\frac{q}{p}} + \xi R^\beta \end{aligned}$$

Applying Hölder's inequality, we obtain

$$U_\sigma \le \frac{1}{2} U_\tau + \left(\frac{c}{(\tau-\sigma)^n} \int_{Q_R} |u|^q \,\mathrm{d}x \right)^{\frac{1}{q}} + c\xi R^\beta.$$

Applying Lemma 5.18, we deduce the desired inequality:

$$\sup_{Q_\rho} |u| \lesssim \left(\frac{1}{(R-\rho)^n} \int_{Q_R} |u|^q \,\mathrm{d}x \right)^{\frac{1}{q}} + c\|a\|_{\mathrm{L}^s(\Omega)}^{\frac{1}{p}} R^\beta.$$

We can now use the boundedness information we have obtained on quasi-minimisers to improve on the previous arguments. In particular, we may replace the growth conditions with

$$|z|^p - \alpha(x, M) \le F(x, y, z) \le \Lambda(M)|z|^p + \alpha(x, M),$$

for some function α and $M \ge \sup|u|$. Using this growth condition in the proof of Theorem 5.25, we see that a quasiminimiser u of the energy functional corresponding to such a F satisfies the following inequality:

$$\int_{A(k,\rho)} |\mathrm{D}u|^p \,\mathrm{d}x \le \frac{H}{(R-\rho)^p} \int_{A(k,R)} (u-k)^p \,\mathrm{d}x + H\xi^p |A(k,R)|^{1-\frac{p}{n}+\varepsilon}. \tag{5.13}$$

Decreasing ε if necessary, we may always assume $\varepsilon \le \frac{p}{n}$. Moreover, the same inequality is satisfied by $-u$. Comparing to the outcome of Theorem 5.25, we see that we no longer require the term involving k^p. However, H and ξ now depend on M. Instead of Theorem 5.28, we now obtain the following improved statement:

Proposition 5.33. *Let u be bounded and satisfying* (5.13) *for every $k \in \mathbb{R}$. Then if $|k_0| + \sup|u| \le M$, we have*

$$\sup_{Q_{R/2}} u \lesssim \left(\frac{1}{R^n} \int_{A(k_0,R)} (u - k_0)^p \,\mathrm{d}x \right)^{\frac{1}{p}} \left(\frac{|A(k_0,R)|}{R^n} \right)^{\frac{\alpha}{p}},$$

where $\alpha > 0$ solves $\alpha^2 + \alpha = \varepsilon$.

Proof. Without loss of generality, we assume $k_0 = 0$. Otherwise, we replace u by $u - k_0$. Repeating the proof of Theorem 5.28, we arrive at

$$\int_{A(k,\rho)} (u-k)^p \,\mathrm{d}x \lesssim \frac{|A(k,r)|^{\frac{p}{n}}}{(r-\rho)^p} \int_{A(k,r)} (u-k)^p \,\mathrm{d}x + \xi^p |A(k,R)|^{1+\varepsilon}$$

for every $\rho < r \le R$. Setting $U(k,t) = \int_{A(k,t)} (u-k)^p \,\mathrm{d}x$, we moreover find for $h < k$, $\rho \le R$,

$$|A(k,\rho)| \le (k-h)^{-p} U(h,r). \tag{5.14}$$

In particular, we deduce

$$U(k,\rho) \lesssim (r-\rho)^{-p} U(h,r) |A(h,r)|^{\frac{p}{n}} + \xi^p (k-h)^{-p} U(k,r) |A(k,r)|^{\varepsilon}. \tag{5.15}$$

Multiplying the left-hand side of (5.15) with the left-hand side of (5.14) raised to the power α and the right-hand side of (5.15) with the right-hand side of (5.14) raised to the power α, we obtain, after setting $\phi(k,t) = U(k,t)|A(k,t)|^{\alpha}$,

$$\phi(k,\rho) \lesssim \left(\left(\frac{r}{r-\rho} \right)^p + \left(\frac{\xi r^\beta}{k-h} \right)^p \right) \frac{r^{-n\varepsilon}}{(k-h)^{p\alpha}} \phi(h,r)^{1+\alpha}.$$

This inequality holds for $\rho < r \le R$ and $h < k$. Letting $d \ge \xi R^\beta$ be a constant, we will determine later and setting $k_i = d(1 - 2^{-i})$, $r_i = \frac{R}{2}(1 + 2^{-i})$, $\phi_i = \phi(k_i, r_i)$, we have shown, after simplifying

$$\phi_{i+1} \le c d^{-p\alpha} 2^{pi(1+\alpha)} R^{-n\varepsilon} \phi_i^{1+\alpha}.$$

Applying Lemma 5.29 with $d \ge c R^{-\frac{n\varepsilon}{\alpha p}} \phi_0^{\frac{1}{p}}$, we conclude that

$$\phi\left(d, \frac{R}{2}\right) = 0.$$

Noting that $d = \xi R^\beta + c R^{-\frac{n\varepsilon}{\alpha p}} \phi_0^{\frac{1}{p}}$ is a valid choice of d, we can translate this to obtain the desired inequality. $\square$

In order to take advantage of Proposition 5.33, we need to estimate $|A(k,R)|$ carefully, in particular, when k is close to $\sup u$. This is achieved in the following Lemma.

Lemma 5.34. *Suppose u is bounded and satisfies* (5.13)*. Consider the constant k_0 defined by $2k_0 = M(2R) + m(2R) = \sup_{Q_{2R}} u + \inf_{Q_{2R}} u$. Assume $|A(k_0,R)| \le \gamma |Q_R|$ for some $\gamma < 1$. If $\operatorname{osc}(u,2R) \ge 2^{\nu+1} \xi R^\beta$ for some integer ν, then*

$$|A(k_\nu, R)| \le c R^{-\frac{n}{p}\frac{p-1}{n-1}} |Q_R|.$$

Here $k_\nu = M(2R) - 2^{-\nu-1} \operatorname{osc}(u,2R)$.

Proof. Let $k_0 < h < k$. Set

$$v = \begin{cases} k-h & \text{if } u \geq k \\ u-h & \text{if } h < h < k \\ 0 & \text{if } u \leq h. \end{cases}$$

Note that $v = 0$ on $Q_R \setminus A(k_0, R)$ which is a set of positive measure by assumption. Thus, Sobolev's embedding applies and gives

$$\left(\int_{Q_R} v^{\frac{n}{n-1}} \, \mathrm{d}x \right)^{1-\frac{1}{n}} \leq c \int_\Delta |\mathrm{D}v| \; dx = c \int_\Delta |\mathrm{D}u| \, \mathrm{d}x,$$

where we denote $\Delta = A(h, R) \setminus A(k, R)$. In particular, using Hölder's inequality, we deduce

$$(k-h)|A(k,R)|^{1-\frac{1}{n}} \leq \left(\int_{Q_R} v^{\frac{n}{n-1}} \, \mathrm{d}x \right)^{1-\frac{1}{n}} \leq c|\Delta|^{1-\frac{1}{p}} \left(\int_{A(k,R)} |\mathrm{D}u|^p \, \mathrm{d}x \right)^{\frac{1}{p}}.$$

However, by (5.13),

$$\begin{aligned} \int_{A(k,R)} |\mathrm{D}u|^p \, \mathrm{d}x \lesssim & \frac{1}{R^p} \int_{A(h,2R)} (u-h)^p \, \mathrm{d}x + \xi^p |A(k,2R)|^{1-\frac{p}{n}+\varepsilon} \\ \lesssim & R^{n-p}(M(2R)-h)^p + \xi^p R^{n-p+n\varepsilon}. \end{aligned}$$

Noting that for $h \leq k_\nu$, $M(2R) - h \geq M(2R) - k_\nu \geq \xi R^\beta$, we combine inequalities to see

$$(k-h)|A(k,R)|^{1-\frac{1}{n}} \lesssim |\Delta|^{1-\frac{1}{p}} R^{\frac{n-p}{p}} (M(2R) - h).$$

Set for $i \leq \nu$, $k_i = M(2R) - 2^{-i-1}\mathrm{osc}(u, 2R)$, $h = k_{i-1}$. Then, after raising to the power $\frac{p}{p-1}$, the inequality above gives

$$|A(k_\nu, R)|^{\frac{p}{n}\frac{n-1}{p-1}} \leq |A(k_i, R)|^{\frac{p}{n}\frac{n-1}{p-1}} \leq cR^{\frac{n-p}{p-1}} |A(k_i, R) \setminus A(k_{i-1}, R)|.$$

Summing from $i = 1$ to ν, this gives

$$\nu |A(k_\nu, R)|^{\frac{p}{n}\frac{n-1}{p-1}} \lesssim R^{\frac{n-p}{p-1}} |A(k_0, R)| \lesssim R^{\frac{n-p}{p-1}} R^n \sim R^{\frac{p}{n}\frac{n-1}{p-1}}.$$

□

We can now finally prove Hölder regularity of functions u satisfying (5.13), for which $-u$ also satisfies (5.13). In particular, this provides Hölder-regularity of quasi-minima as claimed.

Theorem 5.35. *Let u be a bounded function satisfying* (5.13) *with $p > 1$ for every $k \in \mathbb{R}$, so that $-u$ also satisfies* (5.13). *Then u is (locally) Hölder-continuous in Ω.*

Proof. We set as in Lemma 5.34 $2k_0 = M(2R) + m(2R)$. Without loss of generality, we may assume $|A(k_0, R)| \leq \frac{1}{2}|Q_R|$, as otherwise the inequality is satisfied by $-u$. Set $k_\nu = M(2R) - 2^{-\nu-1}\mathrm{osc}(u, 2k)$. Note $k_\nu > k_0$. Applying Lemma 5.33 to k_ν, we find

$$\sup_{Q_{R/2}} u - k_\nu \lesssim \left(R^{-n} \int_{A(k_\nu, R)} (u-k_\nu)^p \, \mathrm{d}x \right)^{\frac{1}{p}} \left(\frac{|A(k_\nu, R)|}{R^n} \right)^{\frac{\alpha}{p}} + \xi R^\beta$$

$$\lesssim \sup_{Q_R}(u-k_\nu)\left(\frac{|A(k_\nu,R)|}{R^n}\right)^{1+\alpha} p+\xi R^\beta.$$

Choose now ν such that $c\nu^{-\frac{n}{p}\frac{n-1}{p-1}} \leq \frac{1}{2}$. Then if $\operatorname{osc}(u,2R) \geq 2^{\nu+1}\xi R^\beta$ due to Lemma 5.34, we find

$$M(R/2)-k_\nu \leq \frac{1}{2}(M(2R)-k_\nu)+c\xi R^\beta.$$

In particular,

$$\operatorname{osc}(u,R/2) \leq \left(1-\frac{1}{2^{\nu+2}}\right)\operatorname{osc}(u,2R)+c\xi R^\beta.$$

Thus, in any case

$$\operatorname{osc}(u,R/2) \leq \left(1-\frac{1}{2^{\nu+2}}\right)\operatorname{osc}(u,2R)+c2^\nu\xi R^\beta.$$

Applying an iteration lemma, Lemma 5.36, which we state after this proof, we conclude

$$\operatorname{osc}(u,\rho) \lesssim \left(\frac{\rho}{R}\right)^\beta \operatorname{osc}(u,R)+\xi\rho^\beta$$

for $\rho < R < \min(R_0, d(x_0,\partial\Omega))$. This gives the required Hölder-regularity of u. □

Lemma 5.36. *Let $\phi(t) > 0$, and assume there exists q and $\tau \in (0,1)$ such that for every $R < R_0$, some $0 < \beta < \delta$ and $t \in (\tau^{k+1}R, \tau^k R)$,*

$$\phi(\tau R) \leq \tau^\delta \phi(R)+BR^\beta \tag{5.16}$$

$$\phi(t) \leq q\phi(\tau^k R) \tag{5.17}$$

Then for every $\rho < R \leq R_0$,

$$\phi(\rho) \leq C\left(\left(\frac{\rho}{R}\right)^\beta \phi(R)+B\rho^\beta\right).$$

Proof. see Problem Sheet 4. We remark however that the second assumption is satisfied with $q=1$ if ϕ is non-decreasing. □

5.6 Harnack's inequality

We close our considerations on regularity theory by observing that the Hölder-continuity, we proved in the previous section, allows us to deduce relatively easily an important property known as Harnack's inequality.

Theorem 5.37. *Let $u \in DG_P^-(\Omega)$ with $k_0 = 0$. Suppose $u > 0$ in Ω and $\rho < \frac{R_0}{2}$ is such that $Q_{6\rho} \subset \Omega$. Then*

$$\sup_{Q_\rho} u \leq C\inf_{Q_\rho} u+\xi\rho^\alpha.$$

The key new ingredient in the proof will be a careful analysis of the infimum of a positive function satisfying DG_p^-, similar to Lemma 5.34.

Lemma 5.38. *Let $0 < u \in DG_p^-(Q)$ with $k_0 = 0$ and $Q = Q_1$. There exists $\gamma_0 > 0$ such that if $|B(\theta, 1)| \leq \gamma_0|Q|$ for some $\theta > 0$, then*

$$\inf_{Q_{1/2}} u \geq \frac{\theta}{2}.$$

Proof. For $h < k < \theta$, we set

$$v = \begin{cases} 0 & \text{if } u \geq k \\ k-h & \text{if } h < u < k \\ k-h & \text{if } u \leq h. \end{cases}$$

Note that for $\frac{1}{2} \leq \rho \leq 1$, $v = 0$ in $Q_\rho \setminus B(k,\rho)$. Moreover,

$$|Q_\rho \setminus B(k,\rho)| \geq 2^{-n}|Q| - |B(\theta,1)| \geq (2^{-n} - \gamma_0)|Q_\rho|.$$

Thus, if $\gamma_0 \leq 2^{-n-1}$, we may apply Sobolev embedding to obtain

$$\begin{aligned}(k-h)|B(k,\rho)|^{1-\frac{1}{n}} \leq \left(\int_{Q_\rho} v^{\frac{n}{n-1}}\,\mathrm{d}x\right)^{1-\frac{1}{n}} &\leq c\int_\Delta |\mathrm{D}v|\,\mathrm{d}x \\ &\leq c|\Delta|^{1-\frac{1}{p}}\left(\int_{B(k,\rho}|\mathrm{D}u|^p\,\mathrm{d}x\right)^{\frac{1}{p}}, \end{aligned} \tag{5.18}$$

where $\Delta = B(k,\rho) \setminus B(h,\rho)$. As $u \in DG_p^-$ and $u > 0$, we have

$$\int_{B(k,\rho)} |\mathrm{D}u|^p\,\mathrm{d}x \leq \frac{c}{(R-\rho)^p}\int_{B(k,R)} (k-u)^p\,\mathrm{d}x + ck^p|B(k,R)|^{1-\frac{p}{n}+\varepsilon} \tag{5.19}$$

$$\leq \frac{c}{(R-\rho)^p}k^p|B(k,R)|^{1-\frac{p}{n}+\varepsilon}. \tag{5.20}$$

Combining inequalities, we have

$$(k-h)|B(h,\rho)|^{1-\frac{1}{n}} \leq \frac{ck}{R-\rho}|B(k,R)|^{1-\frac{1}{n}+\frac{\varepsilon}{p}}.$$

Set $r_i = \frac{1}{2}(1+2^{-i})$, $k_i = \frac{\theta}{2}(1+2^{-i}$ and $B_i = |B(k_i, r_i)$. Then the inequality above reads

$$2^{-i-1}B_{i+1}^{1-\frac{1}{n}} \leq c2^{i+1}B_i^{1-\frac{1}{n}+\frac{\varepsilon}{p}} \Leftrightarrow B_{i+1} \leq C4^{\frac{ni}{n-1}}B_i^{1+\alpha},$$

with $\alpha = \frac{\varepsilon n}{p(n+1)}$. Using Lemma 5.29, we find $u \geq \frac{\theta}{2}$ in $Q_{1/2}$ provided

$$B_0 = |B(\theta,1)| \leq C^{-\frac{1}{\alpha}}4^{-\frac{n}{(n-1)\alpha^2)}} = \gamma_1|Q|.$$

Choosing $\gamma_0 = \min(\gamma_1, 2^{-n-1})$ concludes the proof. □

Lemma 5.38 can be upgraded into the stronger statement:

Lemma 5.39. *Let $0 < u \in DG_p^-(Q_2)$ with $k_0 = 0$. For every $\gamma \in (0,1)$ there exists $\lambda(\gamma) > 0$ such that if $|B(\theta,1)| \leq \gamma|Q_1|$ for some $\theta > 0$, then*

$$\inf_{Q_{1/2}} u \geq \lambda(\gamma)\theta.$$

Proof. Setting $\rho = 1$ and $R = 2$ in (5.19), we find

$$\int_{B(k,1)} |\mathrm{D}u|^p \,\mathrm{d}x \leq k^p.$$

Combined with (5.18) this gives

$$(k-h)^{\frac{p}{p-1}} |B(h,1)|^{\frac{p}{n}\frac{n-1}{p-1}} \leq ck^{\frac{p}{p-1}}(|B(k,1)| - |B(h,1)|)$$

Now set $k_i = \theta 2^{-i}$ for $i \leq \nu$, which we will determine at a later stage. Writing $b_i = |B(k_1, 1)|$, we deduce

$$(\theta 2^{-i-1})^{\frac{p}{p-1}} b_\nu^{\frac{p}{n}\frac{n-1}{p-1}} \leq c(\theta 2^{-i})^{\frac{p}{p-1}}(b_i - b_{i+1}).$$

Simplifying and summing gives

$$(\nu+1) b_\nu^{\frac{p}{n}\frac{n-1}{p-1}} \leq c|Q| = c|Q|^{\frac{p}{n}\frac{n-1}{p-1}}.$$

Thus

$$b_\nu \leq \left(\frac{c}{\nu+1}\right)^{\frac{n}{p}\frac{p-1}{n-1}} |Q|.$$

In particular, for ν sufficiently large, we may apply Lemma 5.38 in order to conclude $u \geq \theta 2^{-\nu-1}$ in $Q_{1/2}$. $\square$

We now prove Theorem 5.37.

Proof of Theorem 5.37. We only consider the case $\xi = 0$ and $R = 1$. The general case can be re-covered by rescaling and considering $u + \xi R^\beta$. We first prove that

$$u(x_0) \leq c \inf_{Q(x_0,R)} u \Leftrightarrow v = \frac{u(x)}{u(x_0)} \geq c > 0. \tag{5.21}$$

It is easy to check that $v \in DG_p(\Omega)$ with the same constants. Thus, due to Theorem 5.35, we obtain

$$\mathrm{osc}_{Q(x,\rho)} v \leq c\,\mathrm{osc}_{Q(x,R)} v \left(\frac{\rho}{R}\right)^\beta \leq c\|v\|_{\mathrm{L}^\infty(Q(x,R))} \left(\frac{\rho}{R}\right)^\beta, \tag{5.22}$$

for $x \in \Omega$ and $\rho < R < \frac{1}{2} d(x, \partial\Omega)$. For δ to be determined later, set $k_\tau = (1-\tau)^{-\delta}$. Choose τ_0 to be the largest value of τ such that $\|v\|_{\mathrm{L}^\infty(Q(x_0,\tau)} \geq k_\tau$. Note that since the left-hand side is bounded and the right-hand side diverges, $\tau_0 \in [0,1)$. Let $\overline{x} \in \overline{Q(x_0, \tau_0)}$ be such that $v(\overline{x}) = \|v\|_{\mathrm{L}^\infty(Q(x_0,\tau_0)} \geq (1-\tau_0)^{-\delta}$. Then

$$\|v\|_{\mathrm{L}^\infty(Q(\overline{x}, \frac{1-\tau_0}{2})} \leq \|v\|_{\mathrm{L}^\infty(Q(x_0, \frac{1+\tau_0}{2})} < k_{\frac{1+\tau_0}{2}} = 2^\delta (1-\tau_0)^{-\delta}.$$

Applying (5.22) with $\rho = \varepsilon R$ and $R = \frac{1-\tau_0}{2}$, we deduce

$$\mathrm{osc}_{Q(\overline{x}, \frac{1-\tau_0}{2}\varepsilon)} v \leq c\|v\|_{\mathrm{L}^\infty(Q(\overline{x}, \frac{1-\tau_0}{2})} \leq c2^\delta (1-\tau_0)\varepsilon^\beta.$$

In particular,

$$v(x) \geq v(\overline{x}) - \mathrm{osc}_{Q(\overline{x}, \frac{1-\tau_0}{2}\varepsilon)} v \geq (1-\tau_0)^{-\delta}(1 - c2^\delta \varepsilon^\beta)$$

for any $x \in Q(\overline{x}, \frac{1-\tau_0}{2}\varepsilon)$. Choosing $\varepsilon = c^{-1}2^{-\delta-1}$, this gives for $x \in Q(\overline{x}, \frac{1-\tau_0}{2}\varepsilon)$,

$$v(x) \geq \frac{1}{2}(1-\tau_0)^{-\delta}.$$

Applying (a re-scaled version of) Lemma 5.39 with $\gamma = 0$ and $\theta = \frac{1}{2}(1-\tau_0)^{-\delta}$, we obtain $\mu(0,2)$ such that for $x \in Q(\overline{x}, (1-\tau)_0\varepsilon)$,

$$v(x) \geq \frac{\mu(0,2)}{2}(1-\tau_0)^{-\delta}.$$

Iterating this argument we find for for any integer ν, $x \in Q(\overline{x}, 2^{\nu-1}(1-\tau_0)\varepsilon)$,

$$v(x) \geq \frac{\mu(0,2)^\nu}{2}(1-\tau_0)^{-\delta}.$$

Choose now ν such that $2 \leq 2^{\nu-1}(1-\tau_0)\varepsilon < 4$. Then

$$v(x) \geq \frac{1}{2}\left(\frac{8}{\varepsilon}\right)^{\log_2(\mu)}(1-\tau_0)^{-\delta-\log_2(\mu)}$$

for $x \in Q(\overline{x}, 2) \supset Q(x_0, 1)$. Choosing $\delta = -\log_2(\mu)$, we have $\varepsilon = \frac{\mu}{2c}$ and deduce

$$v(x) \geq \frac{1}{2}\left(\frac{16c}{\mu}\right)^{\log_2(\mu)}$$

for $x \in Q(x_0, 1)$.

For ρ sufficiently small and $Q_\rho \subset \Omega$, let $x_0 \in \overline{Q}_\rho$ be such that $u(x_0) = \sup_{Q_\rho} u$. Taking $R = 3\rho$ in (5.21), we find

$$\sup_{Q_\rho} u \leq c \inf_{Q(x_0,R)} u \leq c \inf_{Q_\rho} u.$$

□

Theorem 5.37 is of particular interest when $\xi = 0$ in the defining equation of DG_p^+. We denote the corresponding function classes by DGO_p^+, DGO_p^- and DGO_p, respectively. Note in particular that due to Theorem 5.25 this happens when u is a quasi-minimum of

$$\int_\Omega f(x, u, \mathrm{D}u)\,\mathrm{d}x$$

with $|z|^p \leq f(x,y,z) \leq \Lambda|z|^p$. A covering argument in combination with Theorem 5.37 then gives

Theorem 5.40. *Let Ω be bounded, connected and open. Suppose $\Sigma \Subset \Omega$. Let $0 < u \in DGO_p(\Omega)$. Then there exists $C(\Sigma, \Omega) > 0$ such that*

$$\sup_\Sigma u \leq C \inf_\Sigma u.$$

Theorem 5.40 is a strong property as the following two examples illustrate: First, we prove a strong maximum principle for $u \in DGO_p(\Omega)$.

Theorem 5.41. *Suppose Ω is connected, bounded and $u \in DGO_p(\Omega)$. If u has an interior minimum in Ω, then u is constant in Ω.*

Proof. Note that for any $\lambda \in \mathbb{R}$, $u + \lambda \in DGO_p(\Omega)$. Thus, we may assume that $\inf_\Omega u = u(x_0) = 0$ for some $x_0 \in \operatorname{int}(\Omega)$. Assume for a contradiction that there is $x \in \Omega$ with $u(x) > 0$. We may find a finite sequence of balls B_i such that $x_0 \in B_1$ and $x \in B_k$ with $B_i \cap B_{i+1} \neq \emptyset$. Due to Theorem 5.40, we deduce

$$\sup_{B_{i+1}} u \leq c \inf_{B_{i+1} u} \leq c \inf_{B_{i+1} \cap B_i} u \leq c \sup_{B_i} .$$

As $\sup_{B_1} u \leq c \inf_{B_1} = 0$, we deduce $u(x) = 0$, which is a contradiction. □

A second consequence is a Liouville theorem.

Theorem 5.42. *Suppose $u \in DGO_p(\mathbb{R}^n)$ is such that for some c, $u \geq c$ in $\mathbb{R}^n$. Then u is constant.*

Proof. Set $\lambda = \inf_{\mathbb{R}^n} u > -\infty$. Considering $u - \lambda$, we may assume without loss of generality, that $\lambda = 0$. Using Theorem 5.41, we see that if u is not constant, then $u > 0$ in $\mathbb{R}^n$. By Theorem 5.37, for any $R > 0$,

$$\sup_{Q_{R/2}} u \leq c \inf_{Q_{R/2}} u.$$

The right-hand side tends to 0 as $R \to \infty$, so that we deduce $u = 0$ in $\mathbb{R}^n$. □

6 Young measures and weak lower semi-continuity

We now return to a question we mentioned in Section 4. There we saw that in order to prove existence, we required coercivity and lower-semicontiuity properties. In Theorem 4.11, convexity of $f(x, \cdot)$ was crucial in order to derive lower-semicontinuity of the energy. Moreover, we saw that in the one-dimensional, as well as in the scalar, setting, this convexity condition is essentially sharp. In the full vectorial setting this is generally not the case and convexity can be replaced by a weaker notion. The main result of this section will be an essentially sharp lower-semicontinuity statement in this setting. The result will involve a notion called quasi-convexity, which is naturally related to a general tool, known as Young measures, which we explore first.

6.1 Young measures

Young measures are an useful and very versatile tool when attempting to identify the limits under weak convergence. In the following example, we illustrate a typical example, where such a question arises.

Example 6.1. *Consider $\mathscr{F}[u] = \int_\Omega f(x, v_j)\,\mathrm{d}x$, where $v_j \rightharpoonup v$ in $\mathrm{L}^2(\Omega)$. Note that then $(\mathscr{F}[v_j])$ is bounded and hence converges up to a subsequence. We can rephrase identifying this limit as understanding the limit under weak* convergence of the functions $F_j(x) = f(x, v_j)$, which are a bounded sequence in $\mathrm{L}^1(\Omega)$. A natural first guess would be that*

$$f(x, v_j) \overset{*}{\rightharpoonup} f(x, v).$$

However, it is easy to see that in general this is not the case. Consider $\Omega = (0, 1)$ and define for $\theta \in (0, 1)$ and $a \neq b$,

$$v_j = \begin{cases} a & \text{if } jx - \lfloor jx \rfloor \in [0, \theta) \\ b & \text{if } jx - \lfloor jx \rfloor \in [\theta, 1]. \end{cases}$$

Suppose f is a smooth, bounded function such that $f(x,a)=\alpha$, $f(x,b)=\beta$. Note that

$$v_j \overset{*}{\rightharpoonup} \theta a + (1-\theta)b$$
$$f(x,v_j) \overset{*}{\rightharpoonup} \theta\alpha + (1-\theta)\beta = F(x).$$

However, in general $\theta\alpha+(1-\theta)\beta \neq f(x,\theta a+(1-\theta)b)$.

Nevertheless, we can write

$$F(x) = \langle f(x,\cdot),\nu_x\rangle = \int f(x,z)\,\mathrm{d}\nu_x(z)$$

where we define

$$\nu_x = \theta\delta_a + (1-\theta)\delta_b.$$

Thus, the asymptotic distribution of values of $(f(x,v_j))$ is captured by the probability measure ν_x. The weak limit of $(f(x,v_j))$ is thus described by the collection of probability measures $(\nu_x)_x$. We will call $(\nu_x)_x$ the Young measure generated by (v_j).*

Our first task is to prove the existence of the Young measure generated by a L^p-sequence (v_j). This fact is known as the fundamental theorem of Young measure theory.

Theorem 6.2. *Let $(v_j) \subset \mathrm{L}^p(\Omega,\mathbb{R}^N)$ be norm-bounded where $p\in[1,\infty]$. Then there exists a (non-relabeled) subsequence of (v_j) and a family of probability measures $(\nu_x)_{x\in\Omega} \subset \mathscr{M}^1(\mathbb{R}^n)$ called the L^p-Young measure generated by (v_j) such that*

- (i) *$(\nu_x)_x$ is weakly*-measureable*
- (ii) *If $p\in[1,\infty)$, then $\int_\Omega\int |z|^p\,\mathrm{d}\nu_x(z)<\infty$, while if $p=\infty$, there exists $K\subset\mathbb{R}^N$ compact such that $\operatorname{supp}\nu_x \subset K$ for almost every $x\in\Omega$.*
- (iii) *For all $f\colon \Omega\times\mathbb{R}^N\to\mathbb{R}$ Carathéodory such that $f(x,v_j)$ is uniformly L^1-bounded and equi-integrable,*

$$f(x,v_j) \rightharpoonup \left(x \to \int f(x,z)\,\mathrm{d}\nu_x(z)\right) \text{ in } \mathrm{L}^1(\Omega).$$

If $\nu=(\nu_x)_{x\in\Omega}$ satisfies (i) and (ii) in Theorem 6.2, we call ν a L^p-Young measure and write $\nu\in Y^p(\Omega,\mathbb{R}^N)$. When (v_j) generates $\nu=(\nu_x)_{x\in\Omega}$, in the sense that (iii) in Theorem 6.2 holds, we write $v_j \overset{Y}{\to} v$.

Recall that $(f(x,v_j))$ is equi-integrable if and only it is weakly pre-compact in $\mathrm{L}^1(\Omega)$. Absorbing a test-function for weak convergence in f, we note that (iii) can be re-written as

$$\int_\Omega f(x,v_j)\,\mathrm{d}x \to \int_\Omega\int f(x,z)\,\mathrm{d}\nu_x(z)\,\mathrm{d}x = \langle\langle f,\nu\rangle\rangle. \tag{6.1}$$

for all Carathéodory integrands $f(x,z)$ such that $(f(x,v_j))$ is uniformly L^1-bounded and equi-integrable. We refer to $\langle\langle f,\nu\rangle\rangle$ as the duality pairing between f and ν. The barycenter $[\nu]\in\mathrm{L}^p(\Omega,\mathbb{R}^N)$ of a Young-measure $\nu=(\nu_x)\in Y^p(\Omega,\mathbb{R}^N)$ is defined for $x\in\Omega$ as

$$[\nu](x) = [\nu_x] = \langle \mathrm{id},\nu_x\rangle = \int z\,\mathrm{d}\nu_x(z).$$

Finally, we introduce the elementary Young measures $\delta[v_j]=(\delta[v_j]_x)\in Y^p(\Omega,\mathbb{R}^N)$ by setting $\delta[v_j]_x=\delta_{v_j(x)}$. Note that $\delta[v_j]_x$ is only defined up to a set of $\mathscr{L}^n$-measure 0. This ambiguity will always be implicitly present.

We comment that the reason we only require (i) and (ii) to hold in order to define a Young mesaure is that for any parametrised measure $\nu = (\nu)_{x\in\Omega} \subset \mathscr{M}^1(\mathbb{R}^n)$ satisfying (i) and (ii) we can construct a norm-bounded sequence $(v_j) \subset \mathrm{L}^p(\Omega, \mathbb{R}^n)$ such that (iii) is satisfied. In order to see this, note that in order to obtain (iii), considering (6.1) it suffices to show that for $\phi \in S$, $\xi \in \Gamma$, where S and Γ are dense subsets of $\mathrm{L}^1(\Omega)$ and $C_0(\mathbb{R}^N)$ respectively,

$$\lim_{j\to\infty} \int_\Omega \xi\phi(v_j)\,\mathrm{d}x = \int_\Omega \xi(x) \int_{\mathbb{R}^N} \phi(z)\,\mathrm{d}\nu_x(z)\,\mathrm{d}x = \int_\Omega \xi(x)\overline{\phi}(x)\,\mathrm{d}x,$$

where $\overline{\phi}(x) = \int \phi(z)\,\mathrm{d}\nu_x(z)$. In fact, it even suffices to take S countable. Using a Vitali covering with balls centered around Lebesgue points of $\overline{\phi}$, it is not too difficult to see that there are points a_{ki} in Ω and $\varepsilon_{ki} > 0$ such that

$$\int_\Omega \xi(x)\overline{\phi}(x)\,\mathrm{d}x = \lim_{k\to\infty} \sum_i \overline{\phi}(a_{ki}) \int_{a_{ki}+\varepsilon_{ki}\Omega} \xi(x)\,\mathrm{d}x.$$

for all $\xi \in \mathrm{L}^1(\Omega)$, where

$$\Omega = \cup\,(a_{ki} + \varepsilon_{ki}\Omega) \cup N_k \quad |N_K| = 0.$$

In fact, as S is countable, we can ensure by working with points in the intersection of Lebesgue points of all $\phi \in S$, that the above identity holds for all $\phi \in S$. Thus, it suffices to consider the case where ν_x is a sum of Diracs and as we may ensure that the sets $(a_{ki} + \varepsilon_{ki}\Omega)$ are disjoint, even to the case $\nu_x = \delta_{x_0}$. However, in this case any L^p-approximation of the δ-function will give an appropriate sequence v_j.

Before proving Theorem 6.2, we require two measure-theoretic statements. The first is a disintegration argument.

Theorem 6.3. *Let $\Omega \subset \mathbb{R}^n$ open, $\mu \in \mathscr{M}^+(\Omega \times \mathbb{R}^N)$ be a positive Radon measure. Then there exists a weakly* measurable family $(\nu_x)_{x\in\Omega} \subset \mathscr{M}^1(\mathbb{R}^N)$ of probability measures such that with $\kappa \in \mathscr{M}^+(\Omega)$, where κ is defined by setting for $B \subset \Omega$ Borel*

$$\kappa(B) = \mu(B \times \mathbb{R}^N),$$

it holds that

$$\mu = \kappa(\,\mathrm{d}x) \otimes \nu_x.$$

Furthermore, $(\nu_x)_{x\in\Omega}$ is κ-essentially unique.

Proof. Given $\phi \in C_0(\mathbb{R}^N)$ define $\mu_\phi \in \mathscr{M}(\Omega)$ by setting for a Borel set B,

$$\mu_\phi(B) = \int_{B\times\mathbb{R}^N} \phi(z)\,\mathrm{d}\mu(x,z).$$

Note that

$$\mu_\phi(B) \leq \|\phi\|_{\mathrm{L}^\infty(\mathbb{R}^N)}\mu(B \times \mathbb{R}^N) = \|\phi\|_{\mathrm{L}^\infty(\mathbb{R}^N)}\kappa(B).$$

Thus, applying the Besicovitch differentiation theorem, there exists a κ-measurable map $h_\phi\colon \Omega \to \mathbb{R}$ with

$$|h_\phi| \leq \|\phi\|_{\mathrm{L}^\infty(\mathbb{R}^N)}, \qquad \mu_\phi = h_\phi\kappa.$$

Fix $\mathscr{D} \subset C_0(\mathbb{R}^N)$ a dense and countable subset. Then there exists a κ-negligible set $N \subset \Omega$ such that

$$h_{\phi_1}(x) + h_{\phi_2(x)} = h_{\phi_1+\phi_2}(x) \quad \forall x \in \Omega \setminus N, \phi_1, \phi_2 \in \mathscr{D}.$$

Set $T_x[\phi] = h_\phi(x)$ for $x \in \Omega \setminus N$, $\phi \in \mathscr{D}$. Note that

$$|T_x[\phi]| \le \|\phi\|_{\mathrm{L}^\infty(\mathbb{R}^N)}.$$

Thus T_x is a linear bounded operator on $\mathscr{D}$, which can be extended (as a linear bounded operator) to $C_0(\mathbb{R}^N)$. By the Riesz representation theorem for all $x \in \Omega \setminus N$ there exists $\nu_x \in \mathscr{M}(\mathbb{R}^N)$ with $|\nu_x|(\mathbb{R}^N) \le 1$ such that

$$T_x[\phi] = \int \phi(z)\,\mathrm{d}\nu_x(z) \qquad \forall \phi \in C_0(\mathbb{R}^N).$$

Setting $\nu_x = \delta_0$ at $x \in N$, for all $\phi \in \mathscr{D}$,

$$x \to \langle \phi, \nu_x \rangle = T_x[\phi] = h_\phi(x)$$

is a κ-measurable map. By approximation, the κ-measurability extends to $\phi \in C_0(\mathbb{R}^N)$. By a further approximation argument (using Theorem 6.4 below), the weak* measurability of ν_x follows. Now, for $\phi \in \mathscr{D}$ and $B \subset \Omega$ Borel,

$$\begin{aligned}
&\int_{\Omega \times \mathbb{R}^N} 1_B(x)\phi(z)\,\mathrm{d}\mu(x,z) = \mu_\phi(B) = \int_B h_\phi(x)\,\mathrm{d}\kappa(x) \\
&= \int_B \int_{\mathbb{R}^N} \phi(z)\,\mathrm{d}\nu_x(z)\,\mathrm{d}\kappa(x) = \int_\Omega \int_{\mathbb{R}^N} \phi(z) 1_B(x)\,\mathrm{d}\nu_x(z)\,\mathrm{d}\kappa(x).
\end{aligned}$$

This is the claim of the theorem in the case $f = 1_B \otimes \phi$. By an approximation argument, the same identity holds for $f \in C_0(\Omega \times \mathbb{R}^N)$ and also for $f = 1_{B \times \mathbb{R}^n}$, whenever B is a Borel set.

It remains to see that the ν_x are probability measures. Indeed,

$$\mu(B \times \mathbb{R}^N) = \int_B \nu_x(\mathbb{R}^N)\,\mathrm{d}\kappa(x) \le \int_B 1\,\mathrm{d}\kappa(x) = \mu(B \times \mathbb{R}^N).$$

Thus, $\nu_x(\mathbb{R}^N) = 1$ for κ-almost every $x \in \Omega$. The uniqueness claim follows immediately by applying the outcome of the theorem to $f = \phi \otimes \psi$ with $\phi \in C_0(\Omega)$, $\psi \in C_0(\mathbb{R}^N)$. □

Our next theorem concerns the fact that we can restrict $f(x,z)$ to large sets on which $f(x,z)$ is continuous.

Theorem 6.4. *Let $f : \Omega \times \mathbb{R}^N \to \mathbb{R}$ be Carathéodory such that for almost every $x \in \Omega$ $f(x,\cdot)$ is uniformly continuous. Then there exists an increasing sequence of compact sets $S_k \subset \Omega$ with $|\Omega \setminus S_k| \to 0$ such that $f\big|_{S_k \times \mathbb{R}^N}$ is continuous.*

Proof. Set $g_j(x) = \sup_j \left\{ |f(x,z_1) - f(x,z_2)| : z_1, z_2 \in \mathbb{R}^N, |z_1 - z_2| \le \frac{1}{j} \right\}$. As f is Cara-théodory and since $f(x,\cdot)$ is uniformly continuous for almost every $x \in \Omega$, $g_j \to 0$ pointwise almost everywhere. Let $n \in \mathbb{N}$. By Egorov's theorem there exists $K_0 \subset \Omega$ compact with $|\Omega \setminus K_0| \le (2n)^{-1}$ such that $g_j \to 0$ uniformly on K_0. Let (z_i) be dense in $\mathbb{R}^N$. By Lusin's theorem, there exist K_i compact such that $|\Omega \setminus K_i| \le (2^{i+1}n)^{-1}$ and $f(\cdot, z_i)$ is continuous in K_i. Set $S_n = K_0 \cap (\cap K_i)$. Note that $|\Omega \setminus S_n| \to 0$ as $n \to \infty$. Thus given $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|z_1 - z_2| \le 2\delta \quad \Rightarrow \quad |f(x,z_1) - f(x,z_2)| \le \varepsilon$$

for all $x \in S_n \subset K_0$. Given $(\overline{x}, \overline{z}) \in S_n \times \mathbb{R}^N$, pick z_i from (z_j) such that $|\overline{z} - z_i| \le \delta$. For this z_i, there exists $\eta > 0$ such that for all $y \in S_n \subset K_i$

$$|\overline{x} - y| \le \eta \quad \Rightarrow \quad |f(\overline{x}, z_i) - f(y, z_i)| \le \varepsilon.$$

Thus, if also $(x,z) \in S_n \times \mathbb{R}^N$ and $|\overline{x}-x| \le \eta$, $|\overline{z}-z| \le \delta$, noting that then $|z_i - z| \le 2\delta$, we conclude

$$\begin{aligned}|f(\overline{x},\overline{z})-f(x,z)| \le& |f(\overline{x},\overline{z})-f(\overline{x},z_i)|+|f(\overline{x},z_i)-f(x,z_i)|+|f(x,z_i)-f(x,z)|\\ \le& 3\varepsilon.\end{aligned}$$

Consequently $f\big|_{S_n\times\mathbb{R}^n}$ is continuous at $(\overline{x},\overline{z})$. □

The key ingredient in proving Theorem 6.2 will be the following compactness principle.

Proposition 6.5. *Let $p \in [1,\infty]$ and suppose $(\nu^j) \subset Y^p(\Omega,\mathbb{R}^N)$ is a sequence of L^p Young measures. If $p \in [1,\infty)$ assume*

$$\sup_j \langle\langle |\cdot|^p, \nu^j \rangle\rangle = \sup_j \int_\Omega \int_{\mathbb{R}^N} |z|^p \,\mathrm{d}\nu_x^j(z)\,\mathrm{d}x < \infty. \tag{6.2}$$

If $p=\infty$, assume that there exists K compact such that

$$\operatorname{supp} \nu_x^j \subset K \qquad \textit{for almost every } x \in \Omega \textit{ and all } j \in \mathbb{N}. \tag{6.3}$$

Then there exists a non-relabeled subsequence and $\nu \in Y^p(\Omega,\mathbb{R}^N)$ such that

$$\langle\langle f, \nu^j \rangle\rangle \to \langle\langle f, \nu \rangle\rangle \tag{6.4}$$

as $j \to \infty$ for all $f\colon \Omega \to \mathbb{R}^N$ Carathéodory for which $x \to \langle f(x), \nu_x^j \rangle$ is uniformly L^1-bounded and the equi-integrability condition

$$\sup_j \langle\langle |f(x,z)| 1\big|_{\{|f(x,z)|\ge h\}}, \nu^j \rangle\rangle \to 0 \tag{6.5}$$

as $h \to \infty$ holds. Moreover, if $p<\infty$,

$$\langle\langle |\cdot|^p, \nu \rangle\rangle \le \liminf_j \langle\langle |\cdot|^p, \nu^j \rangle\rangle,$$

while if $p=\infty$, for almost every $x \in \Omega$,

$$\operatorname{supp} \nu_x \subset K.$$

We write $\nu^j \overset{*}{\rightharpoonup} \nu$ if

$$\langle\langle f, \nu^j \rangle\rangle \to \langle\langle f, \nu \rangle\rangle$$

as $j \to \infty$ for all $f \in C_0(\Omega \times \mathbb{R}^N)$. Note in particular that (6.4) implies that $\nu^j \overset{*}{\rightharpoonup} \nu$.

Proof. We begin by proving the result in the case where $f \in C_0(\Omega \times \mathbb{R}^N)$. Define

$$\mu^j = \mathscr{L}_x^n \llcorner \Omega \otimes \nu_x^j.$$

To be precise, this is short-hand notation for the Radon measures μ^j defined through their action on $f \in C_0(\Omega \times \mathbb{R}^N)$ by setting

$$\langle f, \mu^j \rangle = \int_\Omega \int f(x,z)\,\mathrm{d}\nu_x^j(z)\,\mathrm{d}x.$$

As an example, if $\nu^j = \delta[v_j]$, then for all $f \in C_0(\Omega \times \mathbb{R}^N)$,

$$\langle f, \mu^j \rangle = \int_\Omega \int f(x,z)\,\mathrm{d}\delta_{v_j(x)}\,\mathrm{d}x = \int_\Omega f(x,v_j)\,\mathrm{d}x.$$

Note that every such μ^j is a positive measure satisfying

$$|\langle f, \mu^j \rangle| \leq |\Omega| \|f\|_{\mathrm{L}^\infty(\Omega \times \mathbb{R}^N)}.$$

In particular, (μ^j) is uniformly bounded in $C_0(\Omega \times \mathbb{R}^N)^*$. By Banach-Alaoglu, there exists a non-relabeled subsequence and $\mu \in C_0(\Omega \times \mathbb{R}^N)$ such that for all $f \in C_0(\Omega \times \mathbb{R}^N)$,

$$\langle f, \mu^j \rangle \to \langle f, \mu \rangle. \tag{6.6}$$

We want to show that $\mu = \mathscr{L}^n_x \llcorner \Omega \otimes \nu_x$ for a weakly* measurable parametrised family $\nu = (\nu_x)_{x \in \Omega} \subset \mathscr{M}^1(\mathbb{R}^N)$ of probability measures. For this, we first observe that due to (6.6), for all $U \subset \Omega$ open,

$$\mu(U \times \mathbb{R}^N) \leq \liminf_{j \to \infty} \mu^j(U \times \mathbb{R}^N) = |U|. \tag{6.7}$$

Further for all $K \subset \Omega$ compact and $R > 0$,

$$\begin{aligned} \mu(K \times \overline{B(0,R)}) \geq& \limsup_{j \to \infty} \mu^j(K \times \overline{B(0,R)}) \\ =& \limsup_{j \to \infty} \int_K \int_{|z| \leq R} 1 d\nu_x^j(z) \,\mathrm{d}x \\ \geq& |K| - \frac{1}{R^p} \sup_j \langle\langle |\cdot|^p, \nu^j \rangle\rangle. \end{aligned}$$

Letting $R \to \infty$ and employing (6.3) (or (6.4) if $p = \infty$), by the inner regularity of Radon measures,

$$\mu(K \times \mathbb{R}^N) \geq |K|. \tag{6.8}$$

In particular, (6.7) and (6.8) justify the application of Theorem 6.3 which gives that

$$\mu = \mathscr{L}^n_x \llcorner \Omega \otimes \nu_x,$$

for some family $(\nu_x)_{x \in \Omega}$ of weak*-measurable probability measures. In particular, for $f \in C_0(\Omega \times \mathbb{R}^N)$,

$$\lim_{j \to \infty} \langle\langle f, \nu^j \rangle\rangle = \lim_{j \to \infty} \langle f, \mu^j \rangle = \langle f, \mu \rangle = \langle\langle f, \nu \rangle\rangle.$$

Our next goal is to remove the continuity assumption on f. Thus assume f is Carathéodory and such that there exists $K \subset \mathbb{R}^N$ compact such that we have $\operatorname{supp} f \subset \Omega \times K$. As a consequence of this assumption for almost every $x \in \Omega$, $f(x, \cdot)$ is uniformly continuous. Due to Theorem 6.4 there exist $S_k \Subset \Omega$ with $|\Omega \setminus S_k| \to 0$ such that $f\big|_{S_k \times \mathbb{R}^N}$ is continuous. Let $f_k \in C_0(\Omega \times \mathbb{R}^N)$ be an extension of $f\big|_{S_k \times \mathbb{R}^N}$. We moreover assume that f_k are uniformly bounded in k. This can be achieved using the Tietze extension theorem, a cut-off and truncation. The details are straightforward.

Then $(\langle f_k(x, \cdot), \nu_x^j \rangle)_j$ is weakly pre-compact in L^1. Due to (6.6),

$$\langle f_k(x, \cdot), \nu_x^j \rangle \rightharpoonup \langle f_k(x, \cdot), \nu_x \rangle \quad \text{in } \mathrm{L}^1(\Omega)$$

The same convergence clearly holds in $\mathrm{L}^1(S_k)$. We now estimate

$$\int_\Omega |\langle f(x, \cdot), \nu_x^j \rangle - 1_{S_k} \langle f(x, \cdot), \nu_x^j \rangle| \,\mathrm{d}x$$

$$\leq \int_{\Omega\setminus S_k} |\langle f(x,\cdot), \nu_x^j\rangle| \,\mathrm{d}x \to 0$$

uniformly in j, since f is bounded. The same estimate holds with ν^j replaced by ν. Thus, we deduce that

$$\langle f(x,\cdot), \nu_x^j\rangle \rightharpoonup \langle f(x,\cdot), \nu_x\rangle \quad \text{in } \mathrm{L}^1(\Omega).$$

It remains to remove the boundedness assumption on the support of f. For this we note that, considering separately the positive and negative part of f, we may assume $f \geq 0$. For $h \in \mathbb{N}$, choose $\rho_h \in C_c^\infty(\mathbb{R}, [0,1])$ such that $\rho_h = 1$ on $B(0,h)$, $\operatorname{supp}\rho_h \subset B(0,2h)$ and set

$$f^h(x,z) = \rho_h(|z|^{\frac{p}{2}})\rho_h(f(x,z))f(x,z).$$

The main estimate is now the following:

$$\begin{aligned}
&E_{j,h} \\
=&\int_\Omega \int |f(x,\cdot) - f^h(x,\cdot)| \,\mathrm{d}\nu_x^j \,\mathrm{d}x \\
\leq& \int_\Omega \int \left| \Big(1 - \rho_h(|z|^{\frac{p}{2}})\rho_h(f(x,z))\Big)\right| |f(x,z)| \,\mathrm{d}\nu_x^j(z) \,\mathrm{d}x \\
\leq& \int \int_{\{(x,z)\in\Omega\times\mathbb{R}^N :\, |z|^{\frac{p}{2}} \geq h \text{ or } |f(x,z)|\geq h\}} |f(x,z)| \,\mathrm{d}\nu_x^j(z) \,\mathrm{d}x \\
\leq& \int_\Omega \int_{z:\, |z|^{\frac{p}{2}} \geq h, |f(x,z)|\leq h} h \,\mathrm{d}\nu_x^j(z) \,\mathrm{d}x + \int\int_{\{(x,z)\in\Omega\times\mathbb{R}^N :\, |f(x,z)|\geq h\}} |f(x,z)| \,\mathrm{d}\nu_x^j(z) \,\mathrm{d}x \\
\leq& \frac{1}{h} \int_\Omega \int_{z:\, |z|^{\frac{p}{2}} \geq h} h^2 \,\mathrm{d}\nu_x^j(z) \,\mathrm{d}x + \int\int_{\{(x,z)\in\Omega\times\mathbb{R}^N :\, |f(x,z)|\geq h\}} |f(x,z)| \,\mathrm{d}\nu_x^j(z) \,\mathrm{d}x \\
\leq& \frac{1}{h} \sup_j \langle\langle |\cdot|^p, \nu^j\rangle\rangle + \sup_j \langle\langle |f(x,z)| 1\big|_{|f(x,z)|\geq h}, \nu^j\rangle\rangle \to 0.
\end{aligned}$$

To obtain the convergence in the last line, we used (6.3) ((6.4) if $p = \infty$) and (6.5). In particular, this allows us to show

$$\begin{aligned}
\lim_j |\langle\langle f, \nu^j\rangle\rangle - \langle\langle f, \nu\rangle\rangle| \leq \limsup_j |\langle\langle f - f^h, \nu^j\rangle\rangle| + |\langle\langle f^h, \nu^j\rangle\rangle - \langle\langle f^h, \nu\rangle\rangle| \\
+ |\langle\langle f^h - f, \nu\rangle\rangle|.
\end{aligned}$$

By our arguments so far, we know that the second term tends to 0 as $j \to \infty$. The first is nothing but $E_{j,h}$. Thus, we deduce that

$$\lim_j |\langle\langle f, \nu^j\rangle\rangle - \langle\langle f, \nu\rangle\rangle| \leq \limsup_j |\langle f^h - f, \nu\rangle\rangle|.$$

But as $f \geq 0$ and $f^h \to f$ in pointwise, bounded convergence, the last term converges to 0. Thus, we have shown (6.4) also in this case.

It remains to prove the moreover part. If $p < \infty$, given $h \in \mathbb{N}$, set $|z|_h = \min(|z|, h)$. Then

$$\liminf_j \langle\langle |\cdot|^p, \nu^j\rangle\rangle \geq \lim_j \langle\langle |\cdot|_h^p, \nu^j\rangle\rangle = \langle\langle |\cdot|_h^p, \nu\rangle\rangle.$$

Letting $h \to \infty$ and using monotone convergence, this gives the desired result. If $p = \infty$, take $\phi \in C_0(\Omega)$ and $\psi \in C_0(\mathbb{R}^N)$ with $\operatorname{supp}\psi \cap K = \emptyset$. Then

$$\langle\langle \phi \otimes \psi, \nu\rangle\rangle = \lim_j \langle\langle \phi \otimes \psi, \nu^j\rangle\rangle = 0.$$

Thus $\operatorname{supp}\nu_x \subset K$ for almost every $x \in \Omega$ as we are able to choose ϕ and ψ arbitrarily. $\square$

Theorem 6.2 is almost a direct consequence of Proposition 6.5.

Proof of Theorem 6.2. The theorem follows directly after applying Proposition 6.5 to the elementary Young measures $(\delta[v_j])_j$. The boundedness assumptions are satisfied due to the L^p-bound on (v_j). The equi-integrability condition follows due to equi-integrability of $(f(x, v_j))_j$. $\square$

There is an alternative, more analytical, perspective and proof on Theorem 6.2. Consider the set $X = L^\infty_{w^*}(\Omega, \mathscr{M}(\mathbb{R}^N))$, the set of essentially bounded, weak* measurable functions. It is possible to identify $X = (L^1(\Omega, C_0(\mathbb{R}^N)))^*$. One then shows that $\nu^j = (x \to \nu^j_x)_j$ are uniformly bounded in X. By the Banach-Alaoglu theorem, we can then extract a weak* limit ν of $(\nu^j)_j$. ν inherits the property of being a collection of probability measures. The precise representation of limits if f is Carathéodory contained in Theorem 6.2 needs to be proven as before.

Without the equi-integrability condition (iii) in Theorem 6.2, we do not obtain convergence anymore. However, a lower semi-continuity property is retained.

Proposition 6.6. *Let $p \in [1, \infty]$ and $(v_j) \subset L^p(\Omega, \mathbb{R}^N)$ be a norm-bounded sequence generating a Young measure $\nu \in Y^p(\Omega, \mathbb{R}^N)$. Let $f\colon \Omega \times \mathbb{R}^N \to [0, \infty)$ be Carathéodory. Then*

$$\liminf_j \int_\Omega f(x, v_j(x))\,\mathrm{d}x = \liminf_j \langle\langle f, \delta[v_j]\rangle\rangle \geq \langle\langle f, \nu\rangle\rangle.$$

Proof. Set $f_h(x, z) = \min(f(x, z), h)$. Then Theorem 6.2 applies and gives

$$\int_\Omega f(x, v_j(x))\,\mathrm{d}x \geq \int_\Omega f_h(x, v_j(x))\,\mathrm{d}x \to \langle\langle f_h, \nu\rangle\rangle = \int_\Omega \int f_h(x, z)\,\mathrm{d}\nu_x(z)\,\mathrm{d}x.$$

Letting $h \to \infty$ and applying dominated convergence, the result follows. $\square$

Before calculating a number of examples, we record a lemma, which was essentially already noted in the proof of Theorem 6.2. However, this lemma is sufficiently useful in practice that it is worth stating separately.

Lemma 6.7. *There exists a countable family $\{\phi_k \otimes h_k\} \subset C_0(\Omega) \times C_0(\mathbb{R}^N)$ such that if $(v_j) \subset \mathrm{L}^p(\Omega, \mathbb{R}^N)$ is uniformly L^p-bounded and $\nu \in Y^p(\Omega, \mathbb{R}^N)$ is such that for all k,*

$$\lim_{j\to\infty} \int \phi_k(x) h_k(v_j(x))\,\mathrm{d}x = \int \phi_k(x)\langle h_k, \nu_x\rangle\,\mathrm{d}x,$$

then $v_j \xrightarrow{Y} \nu$.

Proof. In the proof of Theorem 6.2, we saw that the Young measure generated by a sequence $(v_j) \subset \mathrm{L}^p(\Omega, \mathbb{R}^N)$ is determined already by its behaviour on functions $f \in C_0(\Omega \times \mathbb{R}^N)$. Since there exists a countable dense subset

$$\{\phi_k \otimes h_k\} \subset C_0(\Omega) \times C_0(\mathbb{R}^N)$$

of $C_0(\Omega \times \mathbb{R}^N)$ this concludes the proof. $\square$

Example 6.8. *We calculate a number of examples for sequences generating Young measures.*

(i) *Let* $\Omega = (0,1)$. *Set* $u(x) = 1\big|_{(0,1/2]} - 1\big|_{(1/2,1)}$. *Extend* u *periodically to* $\mathbb{R}$. *Then the sequence given by setting* $u_j(x) = u(jx)$ *generates a Young measure* $\nu \in Y^\infty((0,1))$ *with*

$$\nu = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1.$$

The fact that $\nu \in Y^\infty((0,1))$ *is immediate. Thus, consider* $\phi \in C_0(0,1)$ *and* $h \in C_0(\mathbb{R})$. *Note that* ϕ *is uniformly continuous with modulus of continuity* w *and* h *is uniformly bounded. In particular,*

$$\begin{aligned}
\lim_{j\to\infty}\int_0^1 \phi(x)h(u_j(x))\,\mathrm{d}x &= \lim_{j\to\infty}\sum_{k=0}^{j-1}\int_{k/j}^{(k+1)/j} \phi(k/j)h(u_j(x))\,\mathrm{d}x + O(w(1/j)) \\
&= \lim_{j\to\infty}\sum_{k=0}^{j-1}\frac{1}{j}\phi(k/j)\int_0^1 h(u(y))\,\mathrm{d}y \\
&= \int_0^1 \phi(x)\,\mathrm{d}x\left(\frac{1}{2}h(-1) + \frac{1}{2}h(1)\right).
\end{aligned}$$

Using Lemma 6.7, we conclude $u_j \xrightarrow{Y} \nu$.

(ii) *Let* $\Omega = (0,1)$ *and set* $u_j(x) = \sin(2\pi jx)$. *Then* $u_j \xrightarrow{Y} \nu \in Y^\infty((0,1))$ *where*

$$\nu = \frac{1}{\pi\sqrt{1-y^2}}\mathscr{L}^1_y \llcorner (-1,1).$$

A formal proof of this is part of problem sheet 5.

(iii) *Let* $\Omega \subset \mathbb{R}^2$ *be a bounded Lipschitz domain. Suppose* $A, B \in \mathbb{R}^{2\times 2}$ *and* $a, b \in \mathbb{R}^2$ *are such that* $B - A = a \otimes b$. *Given* $\theta \in (0,1)$, *set for* $x \in \mathbb{R}^2$,

$$u(x) = Ax + \left(\int_0^{x\cdot b} \xi\right) a \qquad \text{where } \xi = 1_{\cup_{z\in\mathbb{Z}}[z,z+1-\theta)}.$$

Then (∇u_j) *generates* $\nu \in Y^\infty((0,1),\mathbb{R}^{2\times 2})$ *with*

$$\nu = \theta\delta_A + (1-\theta)\delta_B.$$

One of the reasons Young measures are useful is that they allows us to detect convergence properties of the sequence (v_j) generating a Young measure ν. We begin by noting that Young measure convergence implies weak convergence of (v_j) to the barycenter $[\nu]$.

Lemma 6.9. *Let* $p \in (1,\infty]$. *Suppose* $(v_j) \subset \mathrm{L}^p(\Omega,\mathbb{R}^N)$ *generates the Young measure* $\nu \in Y^p(\Omega,\mathbb{R}^N)$. *Then*

$$\begin{aligned}
v_j &\rightharpoonup [\nu](x) && \text{in } \mathrm{L}^p(\Omega,\mathbb{R}^N) \text{ if } p \in (1,\infty) \\
v_j &\overset{*}{\rightharpoonup} [v](x) && \text{in } \mathrm{L}^\infty(\Omega,\mathbb{R}^N) \text{ if } p = \infty.
\end{aligned}$$

Proof. Since $p > 1$, as (v_j) is norm-bounded, it is weakly pre-compact. In particular, it suffices to identify the limit of weak* converging subsequences. By the Dunford-Pettis theorem, any such sequence is L^1-equi-integrable. Thus, we can apply Theorem 6.2 with the choice $f(x,z) = z$. This gives exactly the desired conclusion. □

Further, we can use Young measure convergence in order to detect convergence in measure.

Lemma 6.10. *Let $p \in [1,\infty]$. Suppose $\nu \in Y^p(\Omega,\mathbb{R}^N)$ is generated by the norm-bounded sequence $(v_j) \subset \mathrm{L}^p(\Omega,\mathbb{R}^N)$. Let $K \subset \mathbb{R}^N$ be compact. Then*

$$d(v_j,K) \to 0 \text{ in measure} \quad \Leftrightarrow \quad \operatorname{supp}\nu_x \subset K \text{ for a.e.} x \in \Omega.$$

For $v \in \mathrm{L}^p(\Omega,\mathbb{R}^N)$,

$$v_j \to v \text{ in measure} \quad \Leftrightarrow \quad \nu_x = \delta_{v(x)} \text{ for a.e.} x \in \Omega.$$

Proof. Let $f\colon \Omega\times\mathbb{R}^N \to [0,1]$ be Carathéodory and $\delta \in (0,1)$. Then using Markov's inequality and noting that Theorem 6.2 applies to this choice of f,

$$\limsup_{j\to\infty}|\{x \in \Omega\colon f(x,v_j(x)) \geq \delta\}| \leq \lim_{j\to\infty}\frac{1}{\delta}\int_\Omega f(x,v_j)\,\mathrm{d}x = \frac{1}{\delta}\int_\Omega\int f(x,\cdot)\,\mathrm{d}\nu_x\,\mathrm{d}x.$$

However, as $f \leq 1$,

$$\int_\Omega\int f(x,\cdot)\,\mathrm{d}\nu_x\,\mathrm{d}x \leq \delta|\Omega| + \limsup_{j\to\infty}|\{x\in\Omega\colon f(x,v_j(x)) \geq \delta\}|.$$

As δ was arbitrary, we conclude that $f(x,v_j(x)) \to 0$ in measure if and only if we have that $\langle f(x,\cdot),\nu_x\rangle = 0$ for almost every $x \in \Omega$.

Set $f(x,z) = \frac{d(z,K)}{1+d(z,K)}$. Note $f\colon \Omega\times\mathbb{R}^N \to [0,1]$ is Carathéodory. Noting that $f(x,v_j) \to 0$ in measure if and only if $d(v_j,K) \to 0$ and $\langle f(x,\cdot),\nu_x\rangle = 0$ for almost every $x \in \Omega$ if and only if $\operatorname{supp}\nu_x \subset K$, we conclude the first part of the theorem.

To obtain the second part, consider the function $f(x,z) = \frac{|z-v(x)|}{1+|z-v(x)|}$ and argue similarly. □

6.2 Gradient Young measures

We now wish to specialise the theory of Young measures we have established so far to Young measures, which have a generating sequence that consists of gradients.

Definition 6.11. *We say $\nu \in Y^p(\Omega,\mathbb{R}^{m\times n})$ is a gradient Young measure if there exists $(u_j) \subset \mathrm{W}^{1,p}(\Omega,\mathbb{R}^m)$ such that $(\mathrm{D}u_j) \xrightarrow{Y} \nu$. We write $\nu \in GY^p(\Omega,\mathbb{R}^{m\times d})$. If $u \in \mathrm{W}^{1,p}(\Omega,\mathbb{R}^m)$ is such that $[\nu] = \mathrm{D}u$, we say u is an underlying deformation for ν.*

Note that if $\nu \in GY^p(\Omega,\mathbb{R}^{m\times d})$ not all sequences generating ν will consist of gradients. We are guaranteed only the existence of one such sequence. Thus, it is useful to be able to obtain generating sequences that share certain properties of the limiting Young measure.

Lemma 6.12. *Let $p \in (1,\infty]$ and $\nu \in GY^p(\Omega,\mathbb{R}^{m\times n})$. Suppose $u \in \mathrm{W}^{1,p}(\Omega,\mathbb{R}^m)$ is such that $[\nu] = \mathrm{D}u$. Then there exists $(u_j) \subset \mathrm{W}^{1,p}(\Omega,\mathbb{R}^m)$ such that we have $\operatorname{supp}(u_j - u) \Subset \Omega$ and $\mathrm{D}u_j \xrightarrow{Y} \nu$. If $p \in (1,\infty)$, we may additionally ensure that $(\mathrm{D}u_j)$ is equi-integrable.*

Proof. The proof is divided in several steps.
Step 1. As Ω is Lipschitz, we can extend a generating sequence $(\mathrm{D}v_j)$ for ν to all of $\mathbb{R}^n$. Thus, we will always assume that $(v_j) \subset \mathrm{W}^{1,p}(\mathbb{R}^n,\mathbb{R}^m)$ with the uniform bound $\sup_j\|v_j\|_{\mathrm{W}^{1,p}(\mathbb{R}^n,\mathbb{R}^m)} < \infty$. Consider the sequence

$$V_j = M(|v_j| + |\mathrm{D}v_j|),$$

where M is the maximal function. Since (V_j) is uniformly bounded in $\mathrm{L}^p(\Omega, \mathbb{R}^{m\times n})$, we may hence extract a (non-relabeled) sub-sequence generating a Young-measure $\mu \in Y^p(\Omega, \mathbb{R}^{m\times n})$.

Step 2. We first show the claim regarding equi-integrability. Introduce for $h > 0$ the truncation

$$\tau_h s = \begin{cases} s & \text{if } |s| \leq h \\ h\frac{s}{|s|} & \text{if } |s| > h. \end{cases}$$

Note that $(\tau_h V_j)$ is uniformly bounded in $\mathrm{L}^\infty(\Omega, \mathbb{R}^{m\times n})$. For $\phi \in \mathrm{L}^\infty(\Omega)$, we find

$$\begin{aligned} \lim_{h\to\infty} \lim_{j\to\infty} \int_\Omega \phi |\tau_h V_j|^p \,\mathrm{d}x = &\lim_{h\to\infty} \int_\Omega \phi \int |\tau_h s|^p \,\mathrm{d}\mu_x(s) \,\mathrm{d}x \\ =&\langle\langle \phi \otimes |\cdot|^p, \mu \rangle\rangle. \end{aligned} \tag{6.9}$$

To obtain the last identity, we used monotone convergence. Now choose natural numbers $j(k) > j(k-1)$ such that

$$\left| \lim_{j\to\infty} \int_\Omega |\tau_k V_j|^p \,\mathrm{d}x - \int_\Omega |\tau_k V_n|^p \,\mathrm{d}x \right| \leq \frac{1}{k}$$

for all $n \geq j(k)$. For $l \leq k$ and $\psi \in \mathrm{L}^\infty(\Omega)$, we find

$$\int_\Omega \psi |\tau_k V_{j(k)}(x)|^p \,\mathrm{d}x \leq \|\psi\|_{\mathrm{L}^\infty(\Omega)} \int_\Omega |\tau_k V_{j(k)}(x)|^p \,\mathrm{d}x - \int (\|\psi\|_{\mathrm{L}^\infty(\Omega)} - \psi) |\tau_l V_{j(k)}|^p \,\mathrm{d}x.$$

Using (6.9) with $\phi = 1_\Omega$, we deduce

$$\begin{aligned} &\limsup_{k\to\infty} \int \psi |\tau_k V_{j(k)}(x)|^p \,\mathrm{d}x \\ \leq &\|\psi\|_{\mathrm{L}^\infty(\Omega)} \langle\langle 1_\Omega \otimes |\cdot|^p, \mu \rangle\rangle - \int_\Omega (\|\psi\|_{\mathrm{L}^\infty(\Omega)} - \psi) \int |\tau_l s|^p \,\mathrm{d}\mu_x(s) \,\mathrm{d}x. \end{aligned}$$

Letting first $l \to \infty$ and then using monotone convergence, we have shown

$$\limsup_{k\to\infty} \int \psi |\tau_k V_{j(k)}(x)|^p \,\mathrm{d}x \leq \langle\langle \psi \otimes |\cdot|^p, \mu \rangle\rangle.$$

Applying the same argument to $-\psi$, we deduce that

$$|\tau_k V_{j(k)}(x)| \rightharpoonup (x \to \langle |\cdot|^p, \mu \rangle) \quad \text{in } \mathrm{L}^1(\Omega).$$

Thus, by the Dunford-Pettis Theorem, $(W_k) = (\tau_k V_{j(k)})$ is a sequence which is uniformly L^p-bounded and L^p-equi-integrable.

By properties of the maximal function, W_k is Lipschitz with Lipschitz constant at most Ck on the set

$$S_k = \{x \in \Omega \colon V_{j(k)}(x) \leq k\}.$$

By the Kirszbraun theorem we may extend each $v_{j(k)}$ to a globally Lipschitz function $w_k \colon \mathbb{R}^n \to \mathbb{R}^m$ with Lipschitz constant at most Ck. On S_k, we may compute

$$|\mathrm{D}w_k| = |\mathrm{D}v_{j(k)}| \leq V_{j(k)} = W_k.$$

For $x \in \Omega \setminus S_k$, we find

$$|\mathrm{D}w_k| \leq Ck \leq CW_k.$$

In particular, as W_k is equi-integrable, so is $(\mathrm{D}w_k)$. Moreover, by Markov's inequality

$$|\Omega \setminus S_k| \leq \frac{\|V_k\|^p_{\mathrm{L}^p(\Omega)}}{k^p} \to 0$$

as $k \to \infty$. Thus for $\phi \in C_0(\Omega)$ and $h \in C_0(\mathbb{R}^{m\times n})$ we find

$$\int_\Omega |\phi(x)h(\mathrm{D}w_k(x)) - \phi(x)h(\mathrm{D}v_k(x))|\,\mathrm{d}x \leq \|\phi\|_{\mathrm{L}^\infty(\Omega)}\|h\|_{\mathrm{L}^\infty(\mathbb{R}^{m\times n}}|\Omega \setminus S_k| \to 0$$

as $k \to \infty$. Invoking Lemma 6.7, we deduce that $(\mathrm{D}w_k)$ generates the same Young measure as $(\mathrm{D}v_j)$.

Step 3. It remains to modify the boundary behaviour of (w_k). By Rellich-Kondrachov and Lemma 6.9 there exists a non-relabeled subsequence such that $w_k \to u$ in $\mathrm{L}^p(\Omega, \mathbb{R}^m)$. Let $(\rho_j) \subset C_0^\infty(\Omega, [0,1])$ with $G_j = \{x \in \Omega \colon \rho_j = 1\}$ and $|\Omega \setminus G_j| \to 0$ as $j \to \infty$. Set

$$u_{j,k} = \rho_j w_k + (1-\rho_j)u \in \mathrm{W}^{1,p}_u(\Omega, \mathbb{R}^m).$$

Note that $\mathrm{D}u_{j,k} = \rho_j \mathrm{D}w_k + (1-\rho_j)\mathrm{D}u + (w_k - u) \otimes \mathrm{D}\rho_j$. Let $\phi \in C_0(\Omega)$, $h \in C_0(\mathbb{R}^{m\times n})$. Then

$$\int_\Omega |\phi(x)h(\mathrm{D}w_k) - \phi(x)h(\mathrm{D}u_{j,k})|\,\mathrm{d}x \leq |\Omega \setminus G_j|\|\phi\|_{\mathrm{L}^\infty(\Omega)}\|h\|_{\mathrm{L}^\infty(\mathbb{R}^{m\times n})} \to 0$$

uniformly in k. In particular, due to Lemma 6.7, after passing to a suitable diagonal subsequence, we have proven the theorem. ☐

As we saw in our examples, it can occur that $(\nu_x) \in Y^p(\Omega, \mathbb{R}^{m\times n})$ is constant in x. In this case, we write $(\nu_x) = \nu$ and refer to ν as a homogeneous Young measure. It is remarkable (and useful) that in the case of gradient Young measures it does not matter on what domain we define ν.

Lemma 6.13. *Let $\nu \in GY^p(\Omega, \mathbb{R}^{m\times n})$ with $p \in [1,\infty]$. Suppose $[\nu] = \mathrm{D}u$ where $u \in \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$ is linear on $\partial\Omega$. Then, for any Lipschitz domain $D \subset \mathbb{R}^n$, there exists a homogeneous gradient Young measure $\overline{\nu} \in GY^p(D, \mathbb{R}^{m\times n})$ such that*

$$\int h \,\mathrm{d}\overline{\nu} = \fint_\Omega \int h\,\mathrm{d}\nu_x\,\mathrm{d}x$$

for all continuous $h \colon \mathbb{R}^{m\times n} \to \mathbb{R}$ with p-growth (no growth-condition if $p = \infty$).

Proof. We consider the case $p < \infty$. The case $p = \infty$ is easier. Due to Lemma 6.12, we may find $(u_j) \subset \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$ with $u_j\big|_{\partial\Omega} = Fx$ for some $F \in \mathbb{R}^{m\times n}$ and $\mathrm{D}u_j \xrightarrow{Y} \nu$. Using a Vitali covering argument, we can write

$$D = Z^j \cup \left(\cup \Omega(a^j_k, r^j_k)\right),$$

where $|Z^j| = 0$ and $\Omega(a^j_k, r^j_k) = a^j_k + r^j_k\Omega$ for some $a^j_k \in D$, $0 < r^j_k \leq \frac{1}{j}$. Set

$$v_j(y) = r^j_k u_j\left(\frac{y - a^j_k}{r^j_k}\right) + Fa^j_k \qquad \text{if } y \in \Omega(a^j_k, r^j_k).$$

Note by direct calculation $v_j \in \mathrm{W}^{1,p}(D, \mathbb{R}^m)$ and

$$\mathrm{D}v_j(y) = \mathrm{D}u_j\left(\frac{y - a^j_k}{r^j_k}\right) \qquad \text{for } y \in \Omega(a^j_k, r^j_k).$$

Using a change of variables, we calculate for $\phi \in C_0(D)$ and $h\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ continuous with p-growth,

$$\begin{aligned}\int_\Omega \phi(y)h(\mathrm{D}v_j)\,\mathrm{d}x &= \sum_{k=1}^{\infty} \int_{\Omega(a_k^j, r_k^j)} \phi(y)h\left(\mathrm{D}u_j\left(\frac{y-a_k^j}{r_k^j}\right)\right)\mathrm{d}y \\ &= \sum_{k=0}^{\infty} r_k^j \phi(a_k^j) \int_\Omega h(\mathrm{D}u_j)\,\mathrm{d}x + O(1/j)|\mathrm{D}| \\ &\to \frac{1}{|\Omega|}\int_D \phi(x)\,\mathrm{d}x \int_\Omega h(z)\,\mathrm{d}\nu_x(z)\,\mathrm{d}x.\end{aligned}$$

Applying this with the choice $\phi = 1$ and $h(z) = |z|^p$, we see that we have the uniform bound $\sup_j \|\mathrm{D}v_j\|^p_{\mathrm{L}^p(D,\mathbb{R}^m)} < \infty$. Thus, there exists $\overline{\nu} \in GY^p(D, \mathbb{R}^{m\times n})$ such that $\mathrm{D}v_j \xrightarrow{Y} \overline{\nu}$. By Lemma 6.12, we may additionally assume that $(\mathrm{D}u_j)$ (and hence also $(\mathrm{D}v_j)$) are L^p-equi-integrable. In particular, for ϕ and h as above,

$$\int_D \phi(y)h(z)\,\mathrm{d}\overline{\nu}_y(z)\,\mathrm{d}y = \int_D \phi(x) ⨍_\Omega \int h(x)\,\mathrm{d}\nu_x(z)\,\mathrm{d}x.$$

Thus $(\overline{\nu}_y)_y$ is constant in y. Choosing $\phi = 1$ gives the desired result. □

Lemma 6.13 is in fact a generalisation of the Riemann-Lebesgue lemma. Applying it to a fundamental Young measure, we recover the Riemann-Lebesgue lemma.

Lemma 6.14. *Suppose $u \in \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$ with $p \in [1, \infty]$ is linear on $\partial\Omega$. Then there exists a homogeneous gradient Young measure $\overline{\delta[\mathrm{Du}]} \in GY^P(\Omega, \mathbb{R}^{m\times n})$ such that*

$$\int h\,\mathrm{d}\overline{\delta[\mathrm{Du}]} = ⨍_\Omega h(\mathrm{D}u)\,\mathrm{d}x.$$

for all continuous $h\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ with p-growth.

6.3 Quasiconvexity

In Section 4, we studied integrands of the form

$$\mathscr{F}[u] = \int_\Omega f(x, u, \mathrm{D}u)\,\mathrm{d}x$$

where $u\colon \Omega \Subset \mathbb{R}^n \to \mathbb{R}^m$. In particular, we saw that if $n = 1$ or $m = 1$, then convexity is essentially the necessary and sufficient condition on $f(x, \cdot)$ in order to obtain weak sequential lower semicontinuity of the functional. This, in combination with a coercivity assumption led to the existence theory we developed so far, which culminated in Theorem 4.11. However, we already indicated there that if $m > 1$ or $n > 1$, then the convexity assumption can be relaxed. The purpose of this chapter is to study the relevant convexity notion, which is known as quasiconvexity.

One of the main examples driving the development of the theory is the integrand

$$\mathscr{F}[u] = \int_\Omega \det \mathrm{D}u\,\mathrm{d}x.$$

Here $\Omega \subset \mathbb{R}^n$ and $u \in \mathrm{W}^{1,p}_0(\Omega, \mathbb{R}^m)$ for $p \in [n, \infty)$. Using the divergence theorem, we see that

$$\int_\Omega \det \mathrm{D}u\,\mathrm{d}x = \int_\Omega \mathrm{d}u^1 \wedge \ldots \wedge \mathrm{d}u^n = \int_{\partial\Omega} u^1 \wedge \mathrm{d}u^2 \wedge \ldots \wedge \mathrm{d}u^n = 0.$$

In particular, $\mathscr{F}$ is constant on $\mathrm{W}_0^{1,p}(\Omega, \mathbb{R}^m)$ and hence trivially weakly sequentially lower semi-continuous. It will be an outcome of slightly more careful arguments in this section, that actually $\mathscr{F}$ is even weakly sequentially lower semi-continuous on $\mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$. It is however easy to see that det is not convex. For example consider

$$A = \begin{pmatrix} -1 & -2 \\ 2 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & -2 \\ 2 & -1 \end{pmatrix}, \qquad \frac{1}{2}A + \frac{1}{2}B = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}.$$

Then $\frac{1}{2}\det A + \frac{1}{2}\det B = 3 < 4 = \det\left(\frac{1}{2}A + \frac{1}{2}B\right)$.

A further motivation for relaxing the convexity assumption comes from mechanics. A typical property of a mechanical system is frame-invariance. Formally, we consider energy densities $f\colon \mathbb{R}^{n\times n} \to \mathbb{R}$ with the property that

$$f(Qz) = f(z) \qquad \forall Q \in SO(n),\, z \in \mathbb{R}^{n\times n}. \tag{6.10}$$

In many circumstances it is moreover natural to assume that a deformation that is purely compressing or purely expanding increases the energy. In other words, we want to assume that for any $\gamma \neq 1$,

$$f(\gamma \mathrm{Id}) > f(\mathrm{Id}). \tag{6.11}$$

However, an energy density satisfying (6.10) and (6.11) cannot be convex! For simplicity suppose $n = 2$ and define for $\gamma \in (0, 2\pi)$,

$$Q = \begin{pmatrix} \cos(\gamma) & -\sin(\gamma) \\ \sin(\gamma) & \cos(\gamma) \end{pmatrix} \in SO(2).$$

Then if f was convex, we would obtain using (6.10),

$$f(\cos\gamma \mathrm{Id}) \leq \frac{1}{2}(f(Q) + f(Q^t)) = f(\mathrm{Id}),$$

in contradiction to (6.11). Thus f cannot be convex.

In order to resolve this issue, that convexity is not a suitable assumption for problems in mechanics, Morrey introduced the concept of quasiconvexity.

Definition 6.15. *A locally bounded Borel measurable function $f\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ is quasiconvex if*

$$f(z) \leq \fint_{B_1(0)} f(z + \mathrm{D}\phi(y))\,\mathrm{d}y \qquad \forall z \in \mathbb{R}^{m\times n},\ \phi \in \mathrm{W}_0^{1,\infty}(B_1(0), \mathbb{R}^m). \tag{6.12}$$

As a further motivation for this definition of quasiconvexity, we note that affine transformations are minimisers for their own boundary conditions for quasiconvex integrands. Indeed, consider for $u \in \mathrm{W}^{1,\infty}(B_1(0)), \mathbb{R}^m)$,

$$\mathscr{F}[u] = \int_{B_1(0)} f(\mathrm{D}u(y))\,\mathrm{d}y$$

Suppose $a(x) = y_0 + Ax$ for some $y_0 \in \mathbb{R}^m$, $A \in \mathbb{R}^{m\times n}$. Then

$$\mathscr{F}[a] = \int_{B_1(0)} f(A)\,\mathrm{d}x \leq \int_{B_1(0)} f(A + \mathrm{D}\phi(x))\,\mathrm{d}x = \mathscr{F}[a + \phi]$$

for any $\phi \in \mathrm{W}_0^{1,\infty}(B_1(0), \mathbb{R}^m)$. We remark that we will see shortly that the definition of quasi-convexity does not depend on the domain of definition (i.e. $B_1(0)$). However, first we show that convex functions are quasiconvex.

Lemma 6.16. *All convex functions are quasiconvex.*

Proof. Let $A \in \mathbb{R}^{m\times n}$ and $V \in \mathrm{L}^1(B_1(0), \mathbb{R}^{m\times n})$ with $\int_{B_1(0)} V = 0$. Then by the Riesz representation theorem the requirement

$$\langle h, \mu\rangle = \fint_{B_1(0)} h(A + V(x))\,\mathrm{d}x \qquad \forall h \in C_0(\mathbb{R}^{m\times n})$$

defines $\mu \in \mathscr{M}^1(\mathbb{R}^{m\times n})$. Indeed, note that $|\langle h, \mu\rangle| \le \|h\|_{\mathrm{L}^\infty(\mathbb{R}^{m\times n})}$, $\langle h, \mu\rangle \ge 0$ for any $h \ge 0$ and $\langle 1, \mu\rangle = 1$. Consequently, $\mu \in C_0(\mathbb{R}^{m\times n})^*$. Finally, we calculate

$$[\mu] = \langle \mathrm{id}, \mu\rangle = A + \fint_{B_1(0)} V(x)\,\mathrm{d}x = A.$$

Hence, by Jensen's inequality, whenever h is convex,

$$h(A) = h([\mu]) \le \langle h, \mu\rangle = \fint_{B_1((0)} h(A + V(x))\,\mathrm{d}x.$$

Choosing $V = \mathrm{D}\phi$ for $\phi \in \mathrm{W}^{1,\infty}_0(B_1(0), \mathbb{R}^m)$ this concludes the proof. ☐

We record two further elementary properties of quasiconvex functions.

Lemma 6.17. (i) *In the definition of quasiconvexity* (6.12)*,* $B_1(0)$ *may be replaced by any bounded Lipschitz domain* $\Omega \subset \mathbb{R}^d$*.*

(ii) *If* h *has* p*-growth, then it suffices to test with* $\phi \in \mathrm{W}^{1,p}_0(\Omega, \mathbb{R}^m)$ *in* (6.12)*.*

Proof. In order to prove (i), we first prove the following claim: If $\tilde{\Omega}$ is a bounded Lipschitz domain and $\psi \in \mathrm{W}^{1,p}_0(\Omega, \mathbb{R}^m)$, then there exists $\tilde{\psi} \in \mathrm{W}^{1,p}_0(\tilde{\Omega}, \mathbb{R}^m)$ such that

$$\fint_\Omega h(A + \mathrm{D}\psi)\,\mathrm{d}x = \fint_{\tilde{\Omega}} h(A + \mathrm{D}\tilde{\psi})\,\mathrm{d}x \qquad \forall h\colon \mathbb{R}^{m\times n} \to \mathbb{R} \text{ measurable}$$

if one of the integrals exists and is finite. Note that (i) and the fact that the definition of quasiconvexity does not depend on the domain is an immediate consequence of the claim.

Using a Vitali covering argument we have seen a number of times already, write

$$\tilde{\Omega} = Z \cup \bigcup_{k=1}^\infty \Omega(a_k, r_k)$$

where $|Z| = 0$, $a_k \in \Omega$, $r_k > 0$ and $\Omega(a_k, r_k) = a_k + r_k\Omega$. Define

$$\tilde{\psi}(y) = r_k \psi\left(\frac{y - a_k}{r_k}\right) \qquad \text{if } y \in \Omega(a_k, r_k).$$

We calculate for $h\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ measurable,

$$\begin{aligned}\int_{\tilde{\Omega}} h(A + \mathrm{D}\tilde{\psi})\,\mathrm{d}y &= \sum_k \int_{\Omega(a_k, r_K)} h\left(A + \mathrm{D}\psi\left(\frac{y - a_k}{r_k}\right)\right)\mathrm{d}y \\ &= \sum_k r_k^d \int_\Omega h(A + \mathrm{D}\psi)\,\mathrm{d}x = \frac{|\tilde{\Omega}|}{|\Omega|}\int_\Omega h(A + \mathrm{D}\psi)\,\mathrm{d}x.\end{aligned}$$

We turn to (ii). Note that $\mathrm{W}^{1,\infty}_0(B_1(0), \mathbb{R}^m)$ is dense in $\mathrm{W}^{1,p}_0(B_1(0), \mathbb{R}^m)$. Further for h measurable with p-growth, the function $\phi \to \fint_{B_1(0)} h(A + \mathrm{D}\phi)\,\mathrm{d}x$ is well-defined and continuous in $\mathrm{W}^{1,p}(B_1(0), \mathbb{R}^m)$. Thus the claim follows by approximation. ☐

Checking quasiconvexity and working with it is in general difficult. Hence, it will be useful to introduce another notion of convexity, that is weaker than quasiconvexity, but easier to deal with.

Definition 6.18. *A locally bounded, measurable function $f\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ is rank-one convex if*

$$f(\theta A + (1-\theta)B) \leq \theta f(A) + (1-\theta) f(B) \tag{6.13}$$

for all $\theta \in (0,1)$ and $A, B \in \mathbb{R}^{m\times n}$ such that $\operatorname{rank}(A-B) \leq 1$.

Proposition 6.19. *If $h\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ is quasiconvex, then it is rank-one convex.*

Proof. Note that (6.13) trivially holds when $\operatorname{rank}(A-B) = 0$. Thus, take $A, B \in \mathbb{R}^{m\times n}$ with $B - A = a \otimes b$ for some $a \in \mathbb{R}^n \setminus \{0\}$ and $b \in S^{n-1}$. For convenience we work in the unit cube Q_n. Note that this is allowed due to Lemma 6.17.

Step 1. Set $F = \theta A + (1-\theta)B$ and define the laminate $u_j \in \mathrm{W}_0^{1,\infty}(Q_n, \mathbb{R}^m)$ via

$$u_j(x) = Fx + \frac{1}{j}\phi_0(jx \cdot b - \lfloor jx \cdot b \rfloor)a$$

for $x \in Q_n$. Here

$$\phi_0(t) = \begin{cases} -(1-\theta)t & \text{for } t \in [0,\theta] \\ \theta t - t & \text{for } t \in (\theta, 1]. \end{cases}$$

Note that

$$\mathrm{D}u_j = \begin{cases} F - (1-\theta) a \otimes b = A & \text{for } jx \cdot b - \lfloor jx \cdot b \rfloor \in [0,\theta] \\ F + \theta a \otimes b = B & \text{for } jx \cdot b - \lfloor jx \cdot b \rfloor \in (\theta, 1]. \end{cases}$$

In particular, for h locally bounded, measurable,

$$\lim_{j\to\infty} \fint_{Q_n} h(\mathrm{D}u_j)\,\mathrm{d}x = \theta h(A) + (1-\theta) h(B). \tag{6.14}$$

Further, since ϕ_0 is bounded, $u_j \overset{*}{\rightharpoonup} Fx$ in $\mathrm{W}^{1,\infty}(Q_n, \mathbb{R}^m)$.

Step 2. We need to replace (u_j) by a sequence with boundary value Fx, ensuring that (6.14) still remains valid. To this end, let $(\rho_j) \subset C_c^\infty(Q_n, [0,1])$ such that $|\Omega \setminus G_j| \to 0$ where $G_j = \{x \in Q_n \colon \rho_j(x) = 1\}$. Set

$$v_{j,k} = \rho_j u_k + (1-\rho_j) Fx \in \mathrm{W}_{Fx}^{1,\infty}(Q_n, \mathbb{R}^m).$$

Note

$$\mathrm{D}v_{j,k} = \rho_j \mathrm{D}u_k + (1-\rho_j)F + (u_k - Fx) \otimes \mathrm{D}\rho_j.$$

The last term converges to 0 uniformly in k due to Rellich-Kondrachov. In particular, for fixed j,

$$\limsup_{k\to\infty} \|\mathrm{D}v_{j,k}\|_{\mathrm{L}^\infty(Q_n)} \leq \limsup_{k\to\infty} \|\mathrm{D}u_k\|_{\mathrm{L}^\infty(Q_n)} + |F| < \infty.$$

Extracting a diagonal subsequence, we may thus ensure $\|\mathrm{D}v_{j,k(j)}\|_{\mathrm{L}^\infty(Q_n)} \leq C < \infty$ independently of j. Setting $v_j = v_{j,k(j)}$, we find as h is locally bounded

$$\begin{aligned} \lim_{j\to\infty} \int_{Q_n} |h(\mathrm{D}v_j) - h(\mathrm{D}u_{k(j)})|\,\mathrm{d}x &\leq \lim_{j\to\infty} \int_{Q_n\setminus G_j} |h(\mathrm{D}v_j)| + |h(\mathrm{D}u_{k(j)})|\,\mathrm{d}x \\ &\lesssim \lim_{j\to\infty} |Q_n \setminus G_j| = 0. \end{aligned}$$

□

It is clear that if $n = 1$ or $m = 1$ rank-one convexity is equivalent to convexity. Thus combining Lemma 6.16 and Proposition 6.19, if $n = 1$ or $m = 1$, convexity is equivalent to quasiconvexity. We will go on to prove that the determinant (and other examples) are quasiconvex, but not convex. The following example is illustrative of the situation:

Example 6.20 (Albibert-Dacorogna-Marcellini (1988))**.** *Suppose $m = n = 2$ and for $\gamma \in \mathbb{R}$ define*

$$h_\gamma(A) = |A|^2(|A|^2 - 2\gamma \det A).$$

Then

- *h_γ is convex if and only if $|\gamma| \leq \frac{2\sqrt{3}}{3} \approx 0.94$,*
- *h_γ is rank-one convex if and only if $|\gamma| \leq \frac{2}{\sqrt{3}} \approx 1.15$,*
- *h_γ is quasiconvex if and only if $|\gamma| \leq \gamma_{qc}$, $\gamma_{qc} \in (1, 2/\sqrt{3}]$.*

In particular, a major open problem is whether rank-one convexity is equivalent to quasiconvexity if $m = n = 2$. In the case where $m > 2$ or $n > 2$, rank-one convexity does not imply quasiconvexity.

We turn to proving some more properties of rank-one (and thus also of quasiconvex) functions.

Lemma 6.21. *Suppose $h\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ is rank-one convex with $h(z) \leq M(1 + |z|^p)$ for all $z \in \mathbb{R}^{m\times n}$, and some $M > 0$, $p \in [1, \infty)$. Then h has p-growth.*

Proof. Let $R > 0$ and $F_1 \in \mathbb{R}^{m\times n}$ such that $h(F_1) = \inf_{|z|\leq R} h(z)$ (should this not be well-defined, approximate h uniformly by continuous functions). Denote by $F_1, \ldots, F_{2^{mn}}$ the matrices obtained by flipping any number of indices in F_1. The key observation is that the two matrices which only differ at the (i, j)-entry lie on the rank-one line $\mathbb{R}(e_i \otimes e_j)$ and their average has vanishing (i, j)-entry. Thus, applying the definition of rank-one convexity mn-times, we find

$$h(0) \leq \frac{1}{2^{mn}} \sum_k h(F_k).$$

We deduce

$$2^{mn} h(0) \leq (2^{mn} - 1) \sup_{|z|\leq R} h(z) + \inf_{|z|\leq R} h(z).$$

It follows that if $|z| \leq R$,

$$-h(z) \leq M(2^{mn} - 1)(1 + R^p) - 2^{mn} h(0).$$

Choosing $R = |z|$, this shows

$$-h(z) \leq \tilde{M}(1 + |z|^p).$$

□

In fact, rank-one convex functions enjoy much better continuity properties than being locally bounded. They turn out to be locally Lipschitz.

Lemma 6.22. *If $h\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ is rank-one convex, then it is locally Lipschitz. If h has p-growth, then there is $C = C(m, n) > 0$ such that*

$$|h(A) - h(B)| \leq CM(1 + |A|^{p-1} + |B|^{p-1})|A - B|.$$

In particular, if h is rank-one convex with linear growth, then it is globally Lipschitz.

Proof. We will prove that

$$\mathrm{Lip}(h, B(F,r)) \leq \sqrt{\min(m,n)}\frac{\mathrm{osc}(h, \overline{B(F,6r)}}{3r}.$$

Let $A, B \in B(F,r)$ with $\mathrm{rank}(A-B) \leq 1$. Let M be the intersection of the ray from B through A with $\partial B(F,2r)$. Then, since h is convex along the line connecting A and B and difference quotients of convex functions are non-decreasing,

$$\frac{|h(A)-h(B)|}{|A-B|} \leq \frac{|h(M)-h(B)|}{|M-B|} \leq \frac{\mathrm{osc}(h, \overline{B(F,2r)})}{r} = \alpha(2r). \tag{6.15}$$

Now let A, B be two general matrices in $B(F,r)$. Using the singular value decomposition, we can find orthogonal matrices P and Q such that

$$B - A = \sum_{i=1}^{\min(m,n)} \sigma_i P(e_i \otimes e_i)Q^t,$$

where $\sigma_i \geq 0$ is the i-th singular value of $B-A$. Set

$$A_k = A + \sum_{i=1}^{k-1} \sigma_i P(e_i \otimes e_i)Q^t.$$

Then

$$|A_k - F| \leq |A-F| + \sqrt{\sum_{i=1}^{k-1}\sigma_i^2} \leq |A-F| + |B-A| < 3r.$$

Further

$$\sum_{k=1}^{\min(m,n)} |A_k - A_{k+1}|^2 = \sum_{k=1}^{\min(m,n)} \sigma_k^2 = |A-B|^2.$$

In particular, applying (6.15) to $A_k, A_{k+1} \in B(F,3r)$, we may estimate

$$\begin{aligned}|h(A)-h(B)| \leq \sum_{k=1}^{\min(m,n)} |h(A_k)-h(A_{k+1})| &\leq \alpha(6r)\sum_{k=1}^{\min(m,n)} |A_k - A_{k+1}| \\ &\leq \alpha(6r)\sqrt{\min(m,n)}\left(\sum_{k=1}^{\min(m,n)} |A_k - A_{k+1}|^2\right)^{\frac{1}{2}} \\ &= \alpha(6r)\sqrt{\min(m,n)}|A-B|.\end{aligned}$$

If h has p-growth, then $\mathrm{osc}(h, \overline{B(0,R)}) \leq M(1+R^p)$. Setting $F=0$ and $r = \max(|A|,|B|)$, this gives the desired quantitative estimate. □

We now return to the example we considered at the start of this section and prove that the determinant is quasi-convex map. In fact, we will show that all minors are quasiconvex. This will follow from the fact that for these quantities the integral functional only depends on the boundary values. A function with this property is called a null-Lagrangian.

Definition 6.23. *Define ordered multi-indices by setting*

$$P(m,r) = \{(i_1,\dots,i_r) \in \{1,\dots,m\}^r : i_1 < i_2 < \dots < i_r,\ r \in 1,\dots,\min(m,n)\}.$$

Let $I \in P(m,r)$ and $J \in P(n,r)$ be ordered multi-indices. The minor $M\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ is given by setting for $A \in \mathbb{R}^{m\times n}$,

$$M(A) = M^I_J(A) = \det(A^I_J),$$

where A^I_J is the matrix obtained by considering the I-rows and J-columns of A. We call r the rank of M.

Lemma 6.24. *Let $r \in \{1,\dots,\min(m,n)\}$. Suppose $M\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ is a minor of rank r. For $p \in [n,\infty]$, if $u, v \in \mathrm{W}^{1,p}(\Omega,\mathbb{R}^m)$ with $u - v \in \mathrm{W}^{1,p}_0(\Omega,\mathbb{R}^m)$, then*

$$\int_\Omega M(\mathrm{D}u)\,\mathrm{d}x = \int_\Omega M(\mathrm{D}v)\,\mathrm{d}x.$$

In particular, all minors are null-Lagrangians.

Proof. By a standard approximation and cut-off argument, we may assume without loss of generality that u, v are smooth and $\operatorname{supp}(u-v) \Subset \Omega$. Note $|M(A)| \lesssim |A|^r$. Thus, M is strongly continuous in $\mathrm{W}^{1,p}$ for any $p \geq r$.

After re-ordering coordinates, we may assume without loss of generality that M is a principal minor, that is

$$M(\mathrm{D}u)\,\mathrm{d}x^1 \wedge \dots \wedge \mathrm{d}x^n = \mathrm{d}u^1 \wedge \dots \wedge \mathrm{d}u^r \wedge \mathrm{d}x^{r+1} \wedge \dots \wedge \mathrm{d}x^n.$$

Note that then

$$M(\mathrm{D}u)\,\mathrm{d}x^1 \wedge \dots \wedge \mathrm{d}x^n = d(u^1 \wedge \mathrm{d}u^2 \wedge \dots \wedge \mathrm{d}u^r \wedge \mathrm{d}x^{r+1} \wedge \dots \wedge \mathrm{d}x^n). \tag{6.16}$$

In particular, by the (generalised) divergence theorem,

$$\begin{aligned}
\int_\Omega M(\mathrm{D}u)\,\mathrm{d}x^1 \wedge \dots \wedge \mathrm{d}x^n &= \int_\Omega d(u^1 \wedge \mathrm{d}u^2 \wedge \dots \wedge \mathrm{d}u^r \wedge \mathrm{d}x^{r+1} \wedge \dots \wedge \mathrm{d}x^n) \\
&= \int_{\partial\Omega} u^1 \wedge \mathrm{d}u^2 \wedge \dots \wedge \mathrm{d}u^r \wedge \mathrm{d}x^{r+1} \wedge \dots \wedge \mathrm{d}x^n \\
&= \int_{\partial\Omega} v^1 \wedge \mathrm{d}v^2 \wedge \dots \wedge \mathrm{d}v^r \wedge \mathrm{d}x^{r+1} \wedge \dots \wedge \mathrm{d}x^n \\
& \int_\Omega d(v^1 \wedge \mathrm{d}v^2 \wedge \dots \wedge \mathrm{d}v^r \wedge \mathrm{d}x^{r+1} \wedge \dots \wedge \mathrm{d}x^n) \\
&= \int_\Omega M(\mathrm{D}v)\,\mathrm{d}x^1 \wedge \dots \wedge \mathrm{d}x^n.
\end{aligned}$$

□

It is an immediate consequence that minors are quasiconvex.

Corollary 6.25. *Let $r \in \{1,\dots,\min(m,n)\}$. All $r \times r$ minors $M\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ are quasi-affine, that is M and $-M$ are quasi-convex.*

Proof. Let $F \in \mathbb{R}^{m\times n}$, $\psi \in \mathrm{W}^{1,\infty}_0(B_1(0),\mathbb{R}^m)$. Then due to Lemma 6.24,

$$M(F) = \fint_{B_1(0)} M(F)\,\mathrm{d}z = \fint_{B_1(0)} M(F + \mathrm{D}\psi(z))\,\mathrm{d}z.$$

The claim now follows directly. □

In fact, it was shown by Ball that any quasi-affine function is an affine function of minors. Further to quasi-convexity, minors enjoy a surprising weak continuity property.

Lemma 6.26. *Let $r \in \{1, \dots, \min(m, n)\}$. Suppose $M : \mathbb{R}^{m\times n} \to \mathbb{R}$ is a $r \times r$-minor and $(u_j) \subset \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$ for $p \in (r, \infty]$. If $u_j \rightharpoonup u$ in $\mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$ ($\overset{*}{\rightharpoonup}$ if $p = \infty$), then $M(\mathrm{D}u_j \rightharpoonup M(\mathrm{D}u)$ in $\mathrm{L}^{\frac{p}{r}}(\Omega)$ ($\overset{*}{\rightharpoonup}$ if $p = \infty$).*

Proof. We only consider the case $p < \infty$ and $m = n \in \{2, 3\}$. Here we utilise the explicit divergence structure of minors we employed in the proof of Lemma 6.24.

Note that 1×1 minors $M(\mathrm{D}u)$ are just entries of $\mathrm{D}u$ and hence the statement is immediate. Thus let $M^{\neg k}_{\neg l}$ be a 2×2 minor in 3-dimensions. In the 2-dimensional case, we use the same notation, even if the determinant is the only 2×2 minor. Using cyclical indices to explicitly write (6.16) in this set-up, we find

$$\int_\Omega M^{\neg k}_{\neg l}(\mathrm{D}u_j)\,\mathrm{d}x = -\int_\Omega u_j^{k+1}\partial_{l+2}u_j^{k+2}\partial_{l+1}\psi - u_j^{k+1}\partial_{l+1}u_j^{k+2}\partial_{l+2}\psi\,\mathrm{d}x$$

for all $\psi \in C^\infty_c(\Omega)$. By density, the identity holds for all $\psi \in (\mathrm{L}^{\frac{p}{2}}(\Omega))^* \equiv \mathrm{L}^{\frac{p}{p-2}}(\Omega)$. The product on the right-hand side is a product of a L^p-strongly convergent sequence, a L^p-weakly convergent sequence and a fixed function in $\mathrm{L}^{\frac{p}{p-2}}$. By Hölder's inequality, we thus deduce

$$\int_\Omega M^{\neg k}_{\neg l}(\mathrm{D}u_j)\psi\,\mathrm{d}x \to \int_\Omega M^{\neg k}_{\neg l}(\mathrm{D}u)\psi\,\mathrm{d}x$$

It remains to consider the case of the determinant in 3-dimensions. Recalling that the cofactor-matrix may be written as a sum of minors, our work above shows that $\mathrm{cof}(\mathrm{D}u_j) \rightharpoonup \mathrm{cof}(\mathrm{D}u)$ in $\mathrm{L}^{\frac{p}{2}}(\Omega)$. Using Cramer's formula and the Piola identity (which states $\operatorname{div}\mathrm{cof}(\mathrm{D}u) = 0$) shows that

$$\det(\mathrm{D}u) = \sum_{l=1}^{3} \partial_l u^1 (\mathrm{cof}(\mathrm{D}u))^1_l = \sum_{l=1}^{3} \partial_l(u^1 \mathrm{cof}(\mathrm{D}u)^1_l).$$

Arguing first for smooth functions ψ and then by density, we deduce that for all test functions $\psi \in (\mathrm{L}^{\frac{p}{3}}(\Omega))^* \equiv \mathrm{L}^{\frac{p}{p-3}}(\Omega)$,

$$\begin{aligned}\int_\Omega \det(\mathrm{D}u_j)\psi\,\mathrm{d}x &= -\sum_{l=1}^{3}\int_\Omega u_j^1 \mathrm{cof}(\mathrm{D}u_j)^1_l \partial_l\psi\,\mathrm{d}x \\ \to -\sum_{l=1}^{3}\int_\Omega u^1\mathrm{cof}(\mathrm{D}u)^1_l\partial_l\psi\,\mathrm{d}x &= \int_\Omega \det(\mathrm{D}u)\,\mathrm{d}x.\end{aligned}$$

□

6.4 A Jensen-type inequality and rigidity

The relationship between quasi-convex functions and Young measures is given by the following Jensen-type inequality.

Lemma 6.27. *Let $p \in (1, \infty]$ and suppose $\nu \in GY^p(B_1(0), \mathbb{R}^{m\times n})$ is a homogeneous Young measure. Then for all $h : \mathbb{R}^{m\times n} \to \mathbb{R}$ quasiconvex with p-growth (no growth condition if $p = \infty$),*

$$h([\nu]) \le \int h(z)\,\mathrm{d}\nu. \tag{6.17}$$

Note that if h is convex, the lemma is just Jensen's inequality and holds for all homogeneous Young measures, not just gradient Young measures.

Proof. Write $F = [\nu]$ and let $h\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ be quasiconvex with p-growth. Let $u_j \subset \mathrm{W}^{1,p}_{Fx}(B_1(0), \mathbb{R}^m)$ with $\mathrm{D}u_j \xrightarrow{Y} \nu$. Due to Lemma 6.12, we may assume that $(\mathrm{D}u_j)$ is L^p equi-integrable. Due to quasi-convexity of h,

$$h(F) \leq \fint_{B_1(0)} h(\mathrm{D}u_j)\,\mathrm{d}x \to \fint_{B_1(0)} \int h\,\mathrm{d}\nu = \int h\,\mathrm{d}\nu.$$

The convergence holds since $h(\mathrm{D}u_j)$ is equi-integrable due to the equi-integrability of $(\mathrm{D}u_j)$ in L^p and the p-growth of h. $\square$

It can be shown that the converse also holds, that is (6.17) characterises homogeneous gradient Young measures within the class of homogeneous Young measures. This is due to Kinderlehrer-Pedregal and in fact can be extended to the inhomogeneous case.

We record a consequence of Lemma 6.27 for quasi-affine functions. In light of Corollary 6.25, the following applies in particular to minors.

Corollary 6.28. *Let $p \in (1,\infty)$ and $\nu \in GY^p(\Omega, \mathbb{R}^{m\times n})$ be a homogeneous gradient Young measure. Then for all $h\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ quasi-affine with p-growth, we have*

$$h([\nu]) = \int h\,\mathrm{d}\nu.$$

Lemma 6.27 raises the question of whether every Young measure is a gradient Young measure. For inhomogeneous Young measures the answer is clearly no: Whenever ν is a gradient Young measure, $[\nu]$ is a gradient. In particular, $\mathrm{curl}([\nu]) = 0$. Thus, take V with $\mathrm{curl}(V) \neq 0$. Then $\delta[V]$ cannot be a gradient Young measure. For homogeneous Young measures ν, $[\nu]$ is constant and hence automatically a gradient.

Nevertheless, proving that a homogeneous Young measure is not a gradient Young measure is in general not easy. A possible strategy is to prove that (6.17) fails. In the following example we pursue a more direct strategy, based on a rigidity result of Ball-James. Let $A, B \in \mathbb{R}^{m\times n}$, $A \neq B$ and $\theta \in (0,1)$. Consider the Young measure

$$\nu = \theta\delta_A + (1-\theta)\delta_B \in Y^\infty(B_1(0), \mathbb{R}^{m\times n}).$$

Note that we saw in Example 6.8 that for $\mathrm{rank}(A-B) \leq 1$, ν is a gradient Young measure. We will show that for $\mathrm{rank}(A-B) \geq 2$, ν is not a gradient Young measure. The key is the following rigidity result due to Ball-James.

Theorem 6.29. *Let $\Omega \subset \mathbb{R}^n$ be open, bounded and connected. Fix $A, B \in \mathbb{R}^{m\times n}$. Then the following statements holds.*

(i) *Suppose $u \in \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$ is such that $\mathrm{D}u \in \{A, B\}$ almost everywhere in Ω.*
  - *If $\mathrm{rank}(A-B) \geq 2$, then $\mathrm{D}u = A$ almost everywhere or $\mathrm{D}u = B$ almost everywhere.*
  - *If $B - A = a \otimes b$ for $a \in \mathbb{R}^m$, $b \in S^{n-1}$ and Ω is convex, then there exists $h\colon \mathbb{R} \to \mathbb{R}$ Lipschitz with $h' \in \{0,1\}$ almost everywhere and $\nu_0 \in \mathbb{R}^m$ such that*

$$u(x) = \nu_0 + Ax + h(x \cdot b)a.$$

(ii) *If* $\operatorname{rank}(A-B) \geq 2$ *and* $(u_j) \subset \mathrm{W}^{1,p}(\Omega, \mathbb{R}^m)$ *is such that* $\operatorname{dist}(\mathrm{D}u_j, \{A, B\}) \to 0$ *in measure and* $u_j \overset{*}{\rightharpoonup} u$ *in* $\mathrm{W}^{1,\infty}(\Omega, \mathbb{R}^m)$ *for some* $u \in \mathrm{W}^{1,\infty}(\Omega, \mathbb{R}^m)$*, then either* $\mathrm{D}u_j \to \mathrm{D}u = A$ *in measure or* $\mathrm{D}u_j \to \mathrm{D}u = B$ *in measure*

Using Theorem 6.29, it is not difficult to show that ν cannot be a gradient Young measure if $\operatorname{rank}(A-B) \geq 2$. Assume for a contradiction that $\operatorname{rank}(A-B) \geq 2$ and $(u_j) \subset \mathrm{W}^{1,\infty}(B_1(0), \mathbb{R}^m)$ is such that $\mathrm{D}u_j \overset{Y}{\to} \nu$. By Lemma 6.10, $\operatorname{dist}(\mathrm{D}u_j, \{A, B\}) \to 0$ in measure. Using Theorem 6.29, we deduce that either $\mathrm{D}u_j \to A$ or $\mathrm{D}u_j \to B$ in measure. This is a contradiction by Lemma 6.10.

Proof of Theorem 6.29. To prove (i), after a translation, we may assume without loss of generality that $B = 0$. Thus, we have $\mathrm{D}u = Ag$ for some function $g \colon \Omega \to \mathbb{R}$. Mollifying u, we may assume $g \in C^\infty(\Omega)$. We will use that $\mathrm{D}u$ is curl-free, in the sense that

$$\partial_i [\mathrm{D}u]_j^k = \partial_i \partial_j u^k = \partial_j \partial_i u^k = \partial_j (\mathrm{D}u)_i^k.$$

Applying this to our situation, we deduce

$$A_j^k \partial_i g = A_i^k \partial_j g. \tag{6.18}$$

We claim that if $\operatorname{rank}(A) \geq 2$, then $\mathrm{D}g = 0$. Indeed, suppose for a contradiction that $\xi = \nabla g(x) \neq 0$ for some $x \in \Omega$. Set $a_k(x) = \frac{A_j^k}{\xi_j(x)}$ for any j such that $\xi_j(x) \neq 0$. Due to (6.18), a_k does not depend on j. Then

$$A_j^k = a_k(x)\xi_j(x) \Leftrightarrow A = a(x) \otimes \xi(x).$$

This gives the desired contradiction. Thus $\mathrm{D}g = 0$ and consequently, as Ω is connected, u is affine in Ω. Note that the property of being affine is preserved under mollification, so this completes the proof.

If $\operatorname{rank}(A) \leq 1$, $A = a \otimes b$ for some $a \in \mathbb{R}^m$, $b \in S^{n-1}$. Pick $v \in \mathbb{R}^n$ with $v \perp b$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} u(x + tv)\bigg|_{t=0} = \mathrm{D}u(x) \cdot v = [ab^T v]g(x) = 0.$$

Thus u is constant in direction v. As v was an arbitrary direction orthogonal to v and Ω is convex, $u \equiv u(x \cdot b)$. This implies the claim.

We now prove (ii). Again, after a translation, we may assume $B = 0$. As then $\operatorname{rank}(A) \geq 2$, there exists a 2×2 minor $M \colon \mathbb{R}^{m \times n} \to \mathbb{R}$ with $M(A) \neq 0$. Set

$$D_j = \left\{ x \in \Omega \colon |\mathrm{D}u_j - A| < \frac{|A|}{2} \right\}.$$

Then $\mathrm{D}u_j - A1_{D_j} \to 0$ in measure. Choosing a subsequence, we may further assume that $1_{D_j} \overset{*}{\rightharpoonup} \xi$ in L^∞ for some ξ. Now for all $w \in \mathrm{L}^1(\Omega)$ and $\varepsilon > 0$,

$$\int_\Omega (\mathrm{D}u_j - A1_{D_j}) w \, \mathrm{d}x \leq \|\mathrm{D}u_j - A1_{D_j}\|_{\mathrm{L}^\infty(\Omega)} \int_{\{|\mathrm{D}u_j - A1_{D_j} rvert > \varepsilon\}} w \, \mathrm{d}x + \varepsilon \|w\|_{\mathrm{L}^1(\Omega)}.$$

Letting $j \to \infty$, the first term on the right-hand side tends to 0. Thus as ε is arbitrary, $\mathrm{D}u_j \overset{*}{\rightharpoonup} \mathrm{D}u = A\xi$ in L^∞. By Lemma 6.26,

$$M(\mathrm{D}u_j) \overset{*}{\rightharpoonup} M(A\xi) = M(A)\xi^2.$$

Further, similarly to above,

$$M(\mathrm{D}u_j) - M(A)1_{D_j} \overset{*}{\rightharpoonup} 0 \text{ in } \mathrm{L}^\infty(\Omega).$$

In particular, we deduce $M(\mathrm{D}u_j) \overset{*}{\rightharpoonup} M(A)\xi$ and hence $\xi^2 = \xi$. In particular, $\xi = 1_D$ for some set $D \subset \Omega$ and $\mathrm{D}u = A1_D$. As $\|1_{D_j}\|_{\mathrm{L}^2(\Omega)} \to \|1_D\|_{\mathrm{L}^2(\Omega)}$, by Radon-Riesz theorem, $1_{D_j} \to 1_D$ in $\mathrm{L}^2(\Omega)$. In particular, the same convergence holds in measure. We deduce that $\mathrm{D}u_j \to A1_D = \mathrm{D}u$ in measure. Now by (i), $\mathrm{D}u = A$ or $\mathrm{D}u = 0$ almost everywhere. Since by weak*-convergence of (u_j), the limit is unique, this concludes the proof. $\square$

6.5 Lower semi-continuity

We finally turn to one of the main results of this section. We return to studying the problem

$$\min_{u \in \mathrm{W}^{1,p}_g(\Omega,\mathbb{R}^m)} \mathscr{F}[u] \qquad \mathscr{F}[u] = \int_\Omega f(x, \mathrm{D}u)\,\mathrm{d}x,$$

where $\Omega \subset \mathbb{R}^n$ is a bounded Lipschitz domain, $p \in (1,\infty)$, $g \in \mathrm{W}^{1-\frac{1}{p},p}(\partial\Omega,\mathbb{R}^m)$ and f is a Carathéodory function with p-growth, that is for some $\Lambda > 0$ and almost every $x \in \Omega$, every $z \in \mathbb{R}^{m\times n}$,

$$|f(x,z)| \leq \Lambda(1 + |z|^p).$$

The main difference with respect to Section 4 is that we want to assume only quasi-convexity of $f(x,\cdot)$, rather than convexity.

Suppose that we have a norm-bounded sequence $(\mathrm{D}u_j) \subset \mathrm{W}^{1,p}(\Omega,\mathbb{R}^{m\times n})$, such that $(f(x,\mathrm{D}u_j))$ is equi-integrable. Then up to subsequence there is $\nu \in GY^p(\Omega,\mathbb{R}^{m\times n})$ such that $\mathrm{D}u_j \xrightarrow{Y} \nu$. In particular, this gives us a limit for $f(x,\mathrm{D}u_j)$,

$$\int_\Omega f(x,\mathrm{D}u_j)\,\mathrm{d}x \to \int_\Omega \int f(x,z)\,\mathrm{d}\nu_x(z)\,\mathrm{d}x.$$

Thus, to prove sequential weak lower semi-continuity of $\mathscr{F}$ in $\mathrm{W}^{1,p}_g$, it suffices to show

$$\int f(x,z)\,\mathrm{d}\nu_x(z) \geq f(x,\mathrm{D}u)$$

for almost every $x \in \Omega$. If ν is homogeneous, this is just Lemma 6.27. If ν is non-homogeneous, the statement follows from a technique known as blow-up or localisation technique.

Proposition 6.30. *For $p \in [1,\infty)$, let $\nu = (\nu_x)_{x\in\Omega} \in GY^p(\Omega,\mathbb{R}^{m\times n})$. Then for almost every $x_0 \in \Omega$, ν_{x_0} is a homogeneous gradient Young measure, $\nu_{x_0} \in GY^p(B_1(0),\mathbb{R}^{m\times n})$.*

Proof. Let $\{\phi_k \otimes h_k\} \subset C_0(\Omega) \times C_0(\mathbb{R}^m)$ be a countable family of test functions. Then almost every $x_0 \in \Omega$ is a simultaneous Lebesgue point for all the maps $x \to \langle h_k, \nu_x\rangle$. Fix such $x_0 \in \Omega$. Let $(u_j) \subset \mathrm{W}^{1,p}(\Omega,\mathbb{R}^{m\times n})$ be a sequence generating the Young measure ν. By Lemma 6.12, we may assume $(\mathrm{D}u_j)$ is uniformly L^p-norm bounded and L^p equi-integrable.

Set for $y \in B_1(0)$ and writing $[u_j]_{B_r(x_0)} = \fint_{B_r(x_0)} u_j\,\mathrm{d}x$,

$$v_j^r(y) = \frac{u_j(x_0 + ry) - [u_j]_{B_r(x_0)}}{r}.$$

Then

$$
\begin{aligned}
\int_{B_1(0)} \phi_k(y) h_k(\nabla v_j^r)\,\mathrm{d}x &= \int_{B_1(0)} \phi_k(y) h_k(\mathrm{D}u_j(x_0+ry))\,\mathrm{d}y \\
&= \frac{1}{r^n}\int_{B_r(x_0)} \phi_k\left(\frac{x-x_0}{r}\right) h_k(\mathrm{D}u_j(x))\,\mathrm{d}x \\
&\xrightarrow{j\to\infty} \frac{1}{r^n}\int_{B_r(x_0)} \phi_k\left(\frac{x-x_0}{r}\right) \langle h_k,\nu_x\rangle\,\mathrm{d}x \\
&= \frac{1}{r^n}\int_{B_1(0)} \phi_k(y)\langle h_k,\nu_{x_0+ry}\rangle\,\mathrm{d}x \\
&\xrightarrow{r\to 0} \int_{B_1(0)} \phi_k(y)\langle h_k,\nu_{x_0}\rangle\,\mathrm{d}x.
\end{aligned}
$$

Further

$$
\int_{B_1(0)} |\mathrm{D}v_j^r|^p\,\mathrm{d}y = \int_{B_1(0)} |\mathrm{D}u_j(x_0+ry)|^p\,\mathrm{d}x = \frac{1}{r^n}\int_{B_r(x_0)} |\mathrm{D}u_j|^p\,\mathrm{d}x.
$$

In particular $(\mathrm{D}v_j)$ is uniformly bounded in $\mathrm{L}^p(B_1(0),\mathbb{R}^{m\times n})$. Note that this implies that (up to passing to a subsequence) there exists a weak* limit of $|\mathrm{D}u_j|^p\mathscr{L}^n\llcorner\Omega$, which we denote $\lambda\in\mathscr{M}^+(\overline{\Omega})$. By the Besicovitch differentiation theorem,

$$
\limsup_{r\to 0}\frac{\lambda(\overline{B(x_0,r)})}{r^n} < \infty
$$

for almost every $x_0\in\Omega$. Hence we assume this property from now on. Then, it holds that

$$
\limsup_{r\to 0}\lim_{j\to\infty}\int_{B_1(0)} |\mathrm{D}v_j^r|^p\,\mathrm{d}x < \infty.
$$

Noting that $[v_j^r]_{B_1(0)} = 0$, by Poincaré, we may extract a diagonal subsequence $w_n = v_{j(n)}^{r(n)}$ such hat (w_n) is uniformly bounded in $\mathrm{W}^{1,p}(B_1(0),\mathbb{R}^m)$ and as $n\to\infty$,

$$
\int_{B_1(0)} \phi_k(y) h_k(\mathrm{D}w_n(y))\,\mathrm{d}y \to \int_{B_1(0)} \phi_k(y)\langle h_k,\nu_{x_0}\rangle\,\mathrm{d}y.
$$

Using Lemma 6.7, this shows that $\mathrm{D}w_n \xrightarrow{Y} \nu_{x_0}$ for almost every $x_0\in\Omega$. ☐

Proposition 6.30 actually also holds if $p=\infty$, but proving this requires an additional component known as Zhang's lemma. We instead turn to using Proposition 6.30 to prove a lower semi-continuity statement, first proven by Morrey (1952) under additional technical assumptions and proven under the stated assumptions by Acerbi-Fusco (1984), using different techniques than the one we use.

Theorem 6.31. *Let $p\in(1,\infty)$. Suppose $f\colon\Omega\times\mathbb{R}^{m\times n}\to[0,\infty)$ is Carathéodory and has p-growth. Assume moreover that $f(x,\cdot)$ is quasiconvex for almost every $x\in\Omega$. Then the associated functional $\mathscr{F}$ is weakly sequentially lower semi-continuous on $\mathrm{W}^{1,p}(\Omega,\mathbb{R}^m)$.*

Proof. Let $(u_j)\subset\mathrm{W}^{1,p}(\Omega,\mathbb{R}^m)$ and $u\in\mathrm{W}^{1,p}(\Omega,\mathbb{R}^m)$ with $u_j\rightharpoonup u$ in $\mathrm{W}^{1,p}$. Then up to subsequence there exists $\nu=(\nu_x)_{x\in\Omega}\in GY^p(\Omega,\mathbb{R}^{m\times n})$ with $\mathrm{D}u_j\xrightarrow{Y}\nu$ and $[\nu]=\mathrm{D}u$. By Proposition 6.6,

$$
\liminf_{j\to\infty}\int_\Omega f(x,\mathrm{D}u_j)\,\mathrm{d}x \geq \langle\langle f,\nu\rangle\rangle.
$$

For almost every $x \in \Omega$ by Proposition 6.30,

$$\int f(x,z)\,\mathrm{d}\nu_x(z) \geq f(x,\mathrm{D}u).$$

Thus, combining estimates we see

$$\liminf_{j\to\infty} \mathscr{F}[u_j] \geq \mathscr{F}[u].$$

□

Combining Theorem 6.31 with the coercivity statement of Proposition 4.5 and applying the direct method, we immediately obtain the following existence result:

Theorem 6.32. *Suppose $f\colon \Omega \times \mathbb{R}^{m\times n} \to [0,\infty)$ is Carathéodory and satisfies the following assumptions:*

(i) *f has p-growth for some $p \in (1,\infty)$.*

(ii) *f is p-coercive, in the sense that for some $\lambda > 0$, $f(x,z) \geq \lambda|z|^p$.*

(iii) *$f(x,\cdot)$ is quasiconvex for almost every $x \in \Omega$.*

Then the functional $\mathscr{F}$ over $\mathrm{W}^{1,p}_g(\Omega,\mathbb{R}^m)$ for $g \in \mathrm{W}^{1-\frac{1}{p},p}(\partial\Omega,\mathbb{R}^m)$ admits a minimiser.

Similar to Proposition 4.8, quasiconvexity also turns out to be a necessary condition for sequential weak lower semi-continuity. We only show this for homogeneous integrands $f \equiv f(z)$. The general case can be obtained by a localisation argument.

Proposition 6.33. *Suppose $f\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ is continuous and has p-growth. If the associated functional $\mathscr{F}$ is sequentially weakly lower semi-continuous on $\mathrm{W}^{1,p}(\Omega,\mathbb{R}^m)$ (with or without boundary values), then f is quasiconvex.*

Proof. After possibly translating and rescaling Ω, we may assume that $B_1(0) \Subset \Omega$. Let $A \in \mathbb{R}^{m\times n}$, $\phi \in \mathrm{W}^{1,\infty}_0(B_1(0),\mathbb{R}^m)$. By a Vitaly covering argument, we may write

$$B_1(0) = Z^j \cup \bigcup_{k=1}^{\infty} B(a^j_k, r^j_k),$$

where $|Z^j| = 0$, $a^j_k \in B_1(0)$, $0 < r^j_k \leq \frac{1}{j}$ and $B(a^j_k, r^j_k) = B_{r^j_k}(a^j_k)$. Fix a smooth function $h\colon \Omega \setminus B_1(0) \to \mathbb{R}^m$ such that $h(x) = Ax$ for $x \in \partial B(0,1)$ and with $h\big|_{\partial\Omega}$ equal to the given boundary value, if any are given. We define

$$u_j(x) = \begin{cases} Ax + r^j_k \phi\left(\frac{x-a^j_k}{r^j_k}\right) & \text{for } x \in B(a^j_k, r^j_k) \\ h(x) & \text{for } x \in \Omega \setminus B_1(0). \end{cases}$$

Since ϕ is uniformly bounded, $u_j \to u$ in $\mathrm{W}^{1,p}(\Omega)$ where

$$u(x) = \begin{cases} Ax & \text{for } x \in B_1(0) \\ h(x) & \text{if } x \in \Omega \setminus B_1(0). \end{cases}$$

Moreover, since $\mathscr{F}$ is sequentially weakly lower semi-continuous,

$$\int_{B_1(0)} f(A)\,\mathrm{d}x \leq \liminf_{j\to\infty} \int_{B_1(0)} f(\mathrm{D}u_j)\,\mathrm{d}x$$

$$
\begin{aligned}
&= \liminf_{j\to\infty} \sum_k \int_{B(a_k^j, r_k^j)} f\left(A + \mathrm{D}\phi\left(\frac{x-a_k^j}{r_k^j}\right)\right) \mathrm{d}x \\
&= \liminf_{j\to\infty} \sum_k (r_k^j)^n \int_{B_1(0)} f(A + \mathrm{D}\phi(y))\,\mathrm{d}y \\
&= \int_{B_1(0)} f(A + \mathrm{D}\phi)\,\mathrm{d}x.
\end{aligned}
$$

Thus f is quasiconvex. □

One advantage of our proof of Theorem 6.31 is that it easily adapts to integrands depending on u. Consider $f\colon \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ with

$$|f(x,y,z)| \le \Lambda(1 + |y|^p + |z|^p) \tag{6.19}$$

for some $\Lambda > 0$, $p \in (1,\infty)$, almost every $x \in \Omega$ and every $(y,z) \in \mathbb{R}^m \times \mathbb{R}^{m\times n}$. The key idea is to identify the Young measure created by the joint sequence $(u_j, \mathrm{D}u_j)$.

Lemma 6.34. *Suppose $(u_j) \subset \mathrm{L}^p(\Omega,\mathbb{R}^m)$, $(V_j) \subset \mathrm{L}^p(\Omega,\mathbb{R}^N)$ are norm-bounded such that $u_j \to u$ pointwise almost everywhere in Ω and $V_j \xrightarrow{Y} \nu$ for some $\nu \in Y^p(\Omega,\mathbb{R}^N)$. Then*

$$(u_j, V_j) \xrightarrow{Y} \mu = (\mu_x)_{x\in\Omega} \quad \text{where} \quad \mu_x = \delta_{u(x)} \otimes \nu_x.$$

Proof. Due to a density argument as for Lemma 6.7, it suffices to show convergence for test functions

$$f(x,y,z) = \phi(x)\psi(y)h(z) \quad \text{where} \quad \phi \in C_0(\Omega), \psi \in C_0(\mathbb{R}^m), h \in C_0(\mathbb{R}^N).$$

We know that $h(V_j) \overset{*}{\rightharpoonup} (x \to \langle h, \nu_x\rangle)$ in L^∞. Moreover, $\psi(u_j) \to \psi(u)$ pointwise almost everywhere and hence strongly in L^1 as ψ is bounded. In particular, this implies that $\psi(u_j)h(V_j)$ converges weak* in the sense of measures. In other words,

$$\int_\Omega \phi(x)\psi(u_j)h(V_j)\,\mathrm{d}x \to \int_\Omega \phi(x)\psi(u)\langle h, \nu_x\rangle\,\mathrm{d}x.$$

□

Lemma 6.34 allows us to 'freeze' the u-coefficient and effectively argue as in the proof of Theorem 6.31 to obtain a lower semi-continuity statement also for integrands with u-dependence, originally due to Acerbi-Fusco (1984).

Theorem 6.35. *Let $p \in (1,\infty)$ and suppose $f\colon \Omega\times\mathbb{R}^m\times\mathbb{R}^{m\times n}$ is Carathéodory with p-growth. Assume $f(x,y,\cdot)$ is quasiconvex for almost every $x \in \Omega$ and every $y \in \mathbb{R}^m$. Then the associated functional $\mathscr{F}$ is sequentially weakly lower semi-continuous on $\mathrm{W}^{1,p}$. If f is in addition p-coercive, then $\mathscr{F}$ admits a minimiser over $\mathrm{W}^{1,p}_g(\Omega,\mathbb{R}^m)$ for $g \in \mathrm{W}^{1-\frac{1}{p},p}(\partial\Omega,\mathbb{R}^m)$.*