Universitext

Filip Rindler

Calculus of Variations

Springer

Universitext

Universitext

Universitext is a series of textbooks that presents material from a wide variety of mathematical disciplines at master's level and beyond. The books, often well class-tested by their author, may have an informal, personal even experimental approach to their subject matter. Some of the most successful and established books in the series have evolved through several editions, always following the evolution of teaching curricula, into very polished texts.

Thus as research topics trickle down into graduate-level teaching, first textbooks written for new, cutting-edge courses may make their way into *Universitext*.

More information about this series at http://www.springer.com/series/223

Filip Rindler

Calculus of Variations

Springer

Filip Rindler
Mathematics Institute
University of Warwick
Coventry
UK

ISSN 0172-5939 ISSN 2191-6675 (electronic)
Universitext
ISBN 978-3-319-77636-1 ISBN 978-3-319-77637-8 (eBook)
https://doi.org/10.1007/978-3-319-77637-8

Library of Congress Control Number: 2017958602

Mathematics Subject Classification (2010): Primary: 49–01, 49–02; Secondary: 49J45, 35J50, 28B05, 49Q20

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The calculus of variations has its roots in the first problems of optimality studied in classical antiquity by Archimedes (ca. 287–212 BC in Syracuse, *Magna Graecia*) and Zenodorus (ca. 200–140 BC). The beginning of the field as a branch of modern mathematics can be traced back to June 1696, when Johann Bernoulli published a description of the *brachistochrone problem* (see Fig. 1.1 on p. 5) and Leonhard Euler's eponymous 1766 treatise *Elementa calculi variationum*.

The field has seen a sweeping revolution since the formulation of David Hilbert's 19th, 20th, and 23rd problems in 1900, which anticipated the modern treatment of minimization problems. This is particularly true for the theory of so-called multiple integrals, that is, integral functionals defined on spaces of vector-valued maps in several variables. Minimization problems for such functionals have been systematically investigated from the 1950s onward, most notably in the works of Charles B. Morrey Jr., Ennio De Giorgi, and John M. Ball. These developments were further fueled by the adaptation of sophisticated mathematical techniques from measure theory, geometric analysis, and the theory of nonlinear PDEs.

On the application side, the discovery of powerful variational principles to investigate questions of material science, in particular in the theories of nonlinear elasticity and microstructure, was (and is) a rich source of challenging problems, which have shaped the field into its modern form. The methods of the modern calculus of variations are now among the most powerful to study highly nonlinear problems in applications from physics, technology, and economics.

The intent of this book is to give an introduction to the classical and modern calculus of variations with a focus on the theory of integral functionals defined on spaces of vector-valued maps in several variables. It leads the reader from the most fundamental results to topics at the forefront of current research. Almost all of the results presented here are not original, but I have reorganized much of the material and also improved some proofs with ideas that were not known when the original arguments were conceived.

This is not an encyclopedic work. While I do aim to show the big picture, many interesting and important results are omitted and often I only present a special case of a more general theorem. Naturally, the choice of topics that I treat in detail is biased by my own personal preferences.

The presentation of the material in this book is based on a few principles:

- Modern techniques are used whenever this leads to a clearer exposition. Most prominently, Young measures are introduced early in the book since they provide a unified and convenient framework to understand a variety of topics.
- I try to use reasonable assumptions, not the most general ones.
- When presented with a choice of how to prove a result, I have usually chosen what is in my opinion the most conceptually clear approach over more elementary ones.
- This book considers minimization problems over vector-valued maps right from the start since this situation has many applications and, in fact, much of the advanced theory was specifically developed for this case.
- Occasionally, I refer to recent theorems without giving a proof. The rationale here is that I want the reader to see the frontier of research without compromising the coherence of the text.
- I include some pointers to the literature and a few (incomplete) historical comments at the end of every chapter.
- The 120 problems are an integral part of the book and I encourage the reader to attempt as many as possible.

This book has two parts: The first seven chapters form the *Basic Course* and are intended to be read in order. They can form the basis of a 30-hour or 40-hour lecture course for an advanced undergraduate or graduate audience (with some selection on the part of the lecturer of what material to cover in detail). In fact, this part is based on lecture notes for the MA4G6 course on the calculus of variations that I lectured at the University of Warwick in 2015 and 2017 (with Richard Gratwick in 2015 and Kamil Kosiba in 2017).

Part II of the book on *Advanced Topics* contains further material that is suitable for a topics course, a reading seminar, or self-study. Here, three themes with only minimal interdependence are covered: rigidity and microstructure in Chapters 8 and 9; linear growth functionals, singularities in measures, and generalized Young measures in Chapters 10–12; and Γ-convergence for sharp-interface limits and homogenization in Chapter 13. Some results presented in these chapters have so far only been accessible in the research literature, and I hope that even seasoned professionals will find something of interest there.

The prerequisites for this book are a good knowledge of functional analysis, measure theory, and some Sobolev space theory. Most of the results that are required throughout the book are recalled in the appendix.

This book is strongly influenced by several previous works. I note in particular the lecture notes on microstructure by Müller [203], Dacorogna's treatise on the calculus of variations [76], Kirchheim's advanced lecture notes on differential inclusions [160], the monograph on Young measures by Pedregal [222], Giusti's introduction to the calculus of variations [137], Dolzmann's book on microstructure in materials [100], as well as lecture notes on several related courses by Jan Kristensen and Alexander Mielke.

I am grateful for any comments, corrections, and suggestions. They can be sent via the book's website, where a list of corrections will also be maintained:

http://www.calculusofvariations.com

comment@calculusofvariations.com

I would like to thank in particular my mathematical teachers Jan Kristensen and Alexander Mielke. Through their generosity and enthusiasm in sharing their knowledge, they have provided me with the foundation of my study and research. I am also immensely grateful to the following people for many helpful discussions and comments on preliminary versions of the manuscript: Adolfo Arroyo-Rabasa, Lisa Beck, Filippo Cagnetti, Guido De Philippis, Francesco Ghiraldin, Richard Gratwick, Martin Jesenko, Kamil Kosiba, Konstantinos Koumatos, Jan Kristensen, Rajnath Laud, Stefan Müller, Harald Rindler, Angkana Rüland, Bernd Schmidt, Sebastian Schwarzacher, Hanuš Seiner, Giles Shaw, Parth Soneji, Vladimír Šverák, Florian Theil, Jack Thomas, Günter von Häfen. I would also like to thank the production team at Springer and the anonymous referees for their very helpful comments and suggestions. I am hugely indebted to my wife Laura, my daughter Alice, my mother Karin, and my wider family for all their love and support throughout the process of writing this book. I am grateful to Kaye and Prakash for their constant encouragement. Finally, I would like to acknowledge the support from an EPSRC Research Fellowship on "Singularities in Nonlinear PDEs" (EP/L018934/1) and from the University of Warwick.

Coventry, UK
December 2017

Filip Rindler

Contents

Part I
Basic Course

Chapter 1
Introduction

In the quest to formulate useful mathematical models of aspects of the world, it turns out on surprisingly many occasions that the model becomes clearer, more compact, or more tractable if one introduces some form of *variational principle*. This means that one can find a quantity, such as energy or entropy, which obeys a minimization, maximization or saddle-point law.

How much we perceive a variational quantity as “fundamental” or “artificial” depends on the situation at hand. For example, in classical mechanics, one calls forces *conservative* if they are path-independent and hence originate from changing an energy potential. It turns out that many forces in physics are conservative, which seems to imply that the concept of energy should be considered “fundamental”. On the other hand, the entropy as a measure of *missing information* has a more “artificial” flavor.

Our approach to variational quantities here is a pragmatic one: We think of them as providing structure to a problem, which enables us to use powerful *variational methods*. For instance, in elasticity theory it is usually unrealistic to assume that a body will attain a *global* energy-minimizing shape by itself, but this does not mean that a minimum principle cannot be useful in practice. If we wait long enough, the inherent noise in a realistic physical system will move the system’s state around until it is *with high probability* close to a state that has globally minimal energy. The reader interested in the more philosophical aspects of the effectiveness of mathematics in the description of the natural world, and the calculus of variations in particular, is directed to Wigner’s very well-known essay “The Unreasonable Effectiveness of Mathematics in the Natural Sciences” [279] and the book “Mathematics and Optimal Form” by Hildebrandt & Tromba [150] as places to start.

In this book we focus on minimization problems for integral functionals defined on maps from an open and bounded set $\Omega \subset \mathbb{R}^d$ and with values in $\mathbb{R}^m$ ($d, m \in \mathbb{N}$). Thus, we aim to minimize

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8_1

$$\mathscr{F}[u] := \int_\Omega f(x, u(x), \nabla u(x))\, \mathrm{d}x, \qquad u\colon \Omega \to \mathbb{R}^m,$$

usually under conditions on the boundary values of u and possibly under further side constraints. These problems form the original core of the calculus of variations and are as relevant today as they have always been.

From the 1950s onwards, the main research focus has been on variational principles in the *vectorial case* ($d, m > 1$), which exhibit many mathematical difficulties. In particular, it turned out that new forms of (generalized) convexity had to be introduced, most notably Charles B. Morrey Jr.'s *quasiconvexity* [195] and John M. Ball's *polyconvexity* [25]. Another strong driving force of the development in the calculus of variations during the second half of the 20th century was the *Italian School*, which has produced many important discoveries, for instance in regularity theory, geometric problems, and variational convergence (most notably Ennio De Giorgi's Γ-convergence). Some further history of the calculus of variations can be found in [68, 129, 150].

We start by looking at a parade of examples, which we treat at varying levels of detail. The purpose of these examples is to place the mathematical theory in its applied context and to motivate the themes that have guided the development of the field. We will return to all of these examples once we have developed the necessary mathematical tools.

As some examples treat problems from other scientific disciplines, the reader is asked to take some statements on trust and to simply ignore the sections that are of no interest. No knowledge of the following examples is required to understand the exposition of the mathematical theory starting in the next chapter.

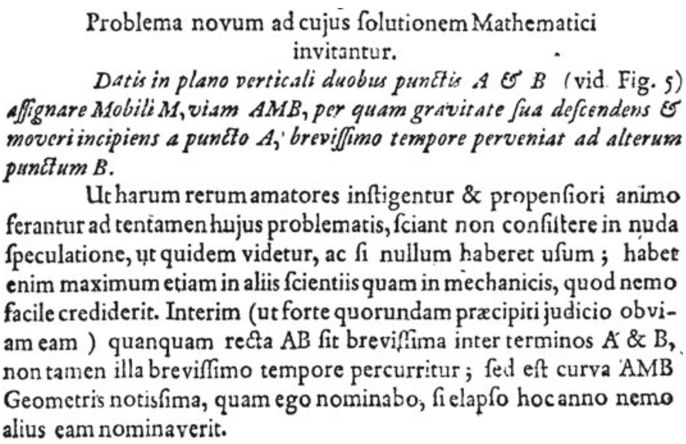
1.1 The Brachistochrone Problem

In June 1696 Johann Bernoulli published the description of a mathematical problem in the journal *Acta Eruditorum*, see Figure 1.1. Bernoulli also sent a letter containing the problem to Leibniz on 9 June 1696, who returned his solution only a few days later on 16 June, and commented that the problem tempted him "like the apple tempted Eve". Newton also published a solution (after the problem had reached him) without giving his identity, but Bernoulli identified him "ex ungue leonem" (from Latin, "by the lion's claw").

The problem that the great minds of the time found so irresistible was formulated as follows:

> *Given two points A and B in a vertical* [meaning "not horizontal"] *plane, one shall find a curve AMB for a movable point M, on which it travels from the point A to the other point B in the shortest time, only driven by its own weight.*

The resulting curve is called the *brachistochrone* (from Ancient Greek, "shortest time") curve.


Problema novum ad cujus ſolutionem Mathematici invitantur.

Datis in plano verticali duobus punctis A & B (vid. Fig. 5) *aſſignare Mobili M, viam AMB, per quam gravitate ſua deſcendens & moveri incipiens a puncto A, breviſſimo tempore perveniat ad alterum punctum B.*

Ut harum rerum amatores inſtigentur & propenſiori animo ferantur ad tentamen hujus problematis, ſciant non conſiſtere in nuda ſpeculatione, ut quidem videtur, ac ſi nullum haberet uſum; habet enim maximum etiam in aliis ſcientiis quam in mechanicis, quod nemo facile crediderit. Interim (ut forte quorundam præcipiti judicio obviam eam) quanquam recta AB ſit breviſſima inter terminos A & B, non tamen illa breviſſimo tempore percurritur; ſed eſt curva AMB Geometris notisſima, quam ego nominabo, ſi elapſo hoc anno nemo alius eam nominaverit.

Fig. 1.1 The birth certificate of the calculus of variations [40] (source: *Hathi Trust Digital Library*)

A more precise formulation of the brachistochrone problem is as follows: We look for the curve connecting the origin $(0, 0)$ to the point $(\bar{x}, \bar{y})$, where $\bar{x} > 0$, $\bar{y} < 0$, such that under the gravitational acceleration (in the negative y-direction) a point mass $m > 0$ slides from rest at $(0, 0)$ to $(\bar{x}, \bar{y})$ quickest among all such curves, see Figure 1.2. We parametrize a point (x, y) on the curve by the time $t \geq 0$ that the mass takes to reach it. The sliding point mass has kinetic and potential energies

$$\begin{aligned} E_{\text{kin}} &= \frac{m}{2}\left[\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2\right] = \frac{m}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2\left[1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2\right], \\ E_{\text{pot}} &= mgy, \end{aligned}$$

where $g \approx 9.81\,\mathrm{m/s^2}$ is the gravitational acceleration on Earth. The total energy $E_{\text{kin}} + E_{\text{pot}}$ is zero at the beginning and conserved along the path. Hence,

$$\frac{m}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2\left[1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2\right] = -mgy.$$

We can solve this for $\mathrm{d}t/\mathrm{d}x$ (where $t = t(x)$ is the inverse of the x-parameterization) to get

$$\frac{\mathrm{d}t}{\mathrm{d}x} = \sqrt{\frac{1 + (y')^2}{-2gy}} \qquad \left(\frac{\mathrm{d}t}{\mathrm{d}x} \geq 0\right),$$

where we wrote $y' = \frac{\mathrm{d}y}{\mathrm{d}x}$. Integrating over the whole x-length along the curve from 0 to $\bar{x}$, we get for the total slide duration $T[y]$ that

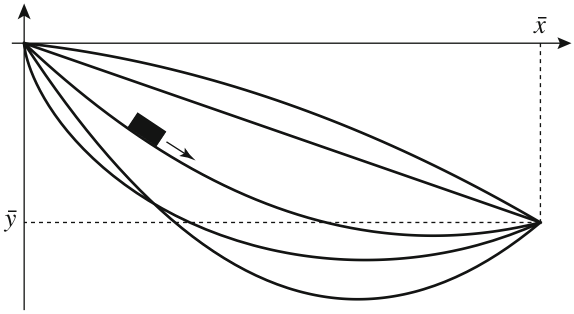


Fig. 1.2 Several slide curves from the origin to $(\bar{x}, \bar{y})$

$$T[y] = \frac{1}{\sqrt{2g}} \int_0^{\bar{x}} \sqrt{\frac{1 + (y'(x))^2}{-y(x)}} \, \mathrm{d}x.$$

We may drop the constant in front of the integral since it does not influence the minimization problem, and set $\bar{x} = 1$ by a reparameterization, to arrive at the problem

$$\begin{cases} \text{Minimize} & \mathscr{F}[y] := \displaystyle\int_0^1 \sqrt{\frac{1 + (y'(x))^2}{-y(x)}} \, \mathrm{d}x \\ \text{subject to} & y(0) = 0, \ \ y(1) = \bar{y} < 0. \end{cases}$$

Notice that the integrand is *convex* in $y'(x)$, which will be important for the solution theory. We will come back to this problem in Example 3.25.

1.2 The Isoperimetric Problem

This problem, which dates back to antiquity and is among the oldest questions in the calculus of variations, asks to enclose a given area with the shortest possible circumference. We pose it in the following version: Given $\alpha, \beta > 0$, find a map $u \colon [0, 1] \to \mathbb{R}$ such that

$$\mathscr{F}[u] := \int_0^1 \sqrt{1 + (u(s)')^2} \, \mathrm{d}s$$

is minimal among all such u with $u(0) = \alpha$, $u(1) = \beta$, and

$$\int_0^1 u(s) \, \mathrm{d}s = A,$$

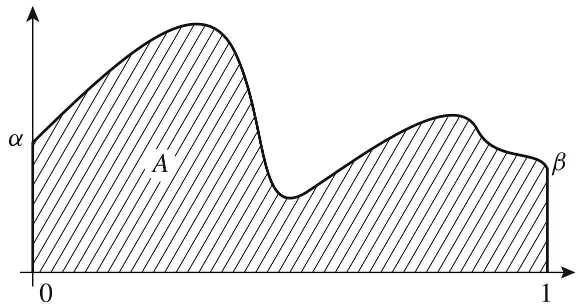


Fig. 1.3 A candidate curve for the isoperimetric problem

where $A > 0$ is the prescribed area under the curve, see Figure 1.3. Note that $\mathscr{F}[u]$ is the length of the curve $\gamma(s) := (s, u(s))^T$. We refer to [45] for more information on the history of this question and the research it inspired.

The difficulty with this problem arises as follows: The integrand $f(a) := \sqrt{1 + a^2}$ behaves like $|a|$ for large values of $|a|$, so it seems possible that solutions have vertical pieces. Thus, it is not clear what kind of candidate functions we should allow in the minimization. We will address this question in Chapter 11, see in particular Example 11.20.

1.3 Electrostatics

Consider an electric charge density $\rho\colon \mathbb{R}^3 \to \mathbb{R}$ (in units of $\mathrm{C/m^3}$) in a three-dimensional vacuum. Let $E\colon \mathbb{R}^3 \to \mathbb{R}^3$ (in V/m) and $B\colon \mathbb{R}^3 \to \mathbb{R}^3$ (in $\mathrm{T} = \mathrm{Vs/m^2}$) be the electric and magnetic fields, respectively, which we assume to be constant in time (hence electro*statics*). The *Gauss law* for electricity reads

$$\nabla \cdot E = \operatorname{div} E = \frac{\rho}{\varepsilon_0},$$

where $\varepsilon_0 \approx 8.854 \cdot 10^{-12}\,\mathrm{C/(Vm)}$ is the vacuum permittivity (electric constant). Moreover, we have the *Faraday law of induction*

$$\nabla \times E = \operatorname{curl} E = \frac{\mathrm{d}B}{\mathrm{d}t} = 0,$$

where t denotes time. Thus, since E is curl-free, there exists an *electric potential* $\phi\colon \mathbb{R}^3 \to \mathbb{R}$ (in V) such that

$$E = -\nabla\phi.$$

Combining this with the Gauss law, we arrive at the *Poisson equation*,

$$\Delta\phi = \nabla \cdot [\nabla\phi] = -\frac{\rho}{\varepsilon_0}. \tag{1.1}$$

We can also look at electrostatics in a variational way: With the norming condition $\phi(0) = 0$, the electric potential energy $U_E(x; q)$ of a point charge q (in C) at the point $x \in \mathbb{R}^3$ in the electric field E is given by the path integral

$$U_E(x; q) = -\int_0^x qE \cdot \mathrm{d}s = -\int_0^1 qE(hx) \cdot x \,\mathrm{d}h = q\phi(x),$$

which does not depend on the path chosen since E is a gradient. Hence, the total *electric energy* of our charge distribution ρ in its own electrical field is

$$U_E := \frac{1}{2}\int_{\mathbb{R}^3} \rho\phi \,\mathrm{d}x = \frac{\varepsilon_0}{2}\int_{\mathbb{R}^3} (\nabla \cdot E)\phi \,\mathrm{d}x,$$

which has units of CV $=$ J (the factor $1/2$ is necessary to count mutual reaction forces correctly). Using the identity

$$(\nabla \cdot E)\phi = \nabla \cdot (E\phi) - E \cdot (\nabla\phi),$$

the Gauss–Green theorem, and the natural assumption that ϕ vanishes at infinity, we get

$$U_E = \frac{\varepsilon_0}{2}\int_{\mathbb{R}^3} \nabla \cdot (E\phi) - E \cdot (\nabla\phi) \,\mathrm{d}x = -\frac{\varepsilon_0}{2}\int_{\mathbb{R}^3} E \cdot (\nabla\phi) \,\mathrm{d}x = \frac{\varepsilon_0}{2}\int_{\mathbb{R}^3} |\nabla\phi|^2 \,\mathrm{d}x.$$

The integral

$$\int_\Omega \frac{1}{2}|\nabla\phi(x)|^2 \,\mathrm{d}x$$

is called the *Dirichlet functional* or the *Dirichlet integral*.

In Example 3.4 we will see that the solutions ϕ of (1.1) are precisely the minimizers of the variational problem

$$\text{Minimize} \quad \phi \mapsto U_E - \int_{\mathbb{R}^3} \rho(x)\phi(x) \,\mathrm{d}x = \int_{\mathbb{R}^3} \frac{\varepsilon_0}{2}|\nabla\phi(x)|^2 - \rho(x)\phi(x) \,\mathrm{d}x.$$

The second term can be interpreted as the interaction energy between the electric field and the charge density ρ. The existence and regularity of solutions to this minimization problem will be established in Examples 2.8 and 3.15.

1.4 Stationary States in Quantum Mechanics

The non-relativistic evolution of a quantum mechanical system with N degrees of freedom in an electric field is described completely through its *wave function* $\Psi : \mathbb{R}^N \times \mathbb{R} \to \mathbb{C}$ that satisfies the *Schrödinger equation*

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t}\Psi(x,t) = \left[\frac{-\hbar^2}{2\mu}\Delta + V(x,t)\right]\Psi(x,t), \qquad (x,t) \in \mathbb{R}^N \times [0,\infty),$$

where $\hbar \approx 1.05 \cdot 10^{-34}$ Js is the reduced Planck constant, μ is the reduced mass (in kg), and $V = V(x,t) \in \mathbb{R}$ is the potential energy (in J). The operator $H := -(2\mu)^{-1}\hbar^2\Delta + V$ is called the *Hamiltonian* of the system.

The value of the wave function itself at a given point in spacetime has no obvious physical meaning, but according to the *Copenhagen interpretation* of quantum mechanics, $x \mapsto |\Psi(x,t)|^2$ is the probability density of finding a particle at the point x in a measurement at time t. In order for $|\Psi(\bullet,t)|^2$ to be a probability density, we need to impose the side constraint

$$\|\Psi(\bullet,t)\|^2_{\mathrm{L}^2(\mathbb{R}^N)} = \int_{\mathbb{R}^N} |\Psi(x,t)|^2 \,\mathrm{d}x = 1 \qquad \text{for all } t \in [0,\infty).$$

In particular, $|\Psi(x,t)|$ has to decay as $|x| \to \infty$.

Of special interest are the so-called *stationary states*, that is, solutions of the *stationary Schrödinger equation*

$$\left[\frac{-\hbar^2}{2\mu}\Delta + V(x)\right]\Psi(x) = E\Psi(x), \qquad x \in \mathbb{R}^N,$$

where $E > 0$ is an *energy level*. If we are just interested in the lowest-energy state, the so-called *ground state*, we can instead find minimizers of the energy functional

$$\mathscr{E}[\Psi] := \int_{\mathbb{R}^N} \frac{\hbar^2}{4\mu}|\nabla\Psi(x)|^2 + \frac{1}{2}V(x)|\Psi(x)|^2 \,\mathrm{d}x,$$

again under the side constraint

$$\|\Psi\|^2_{\mathrm{L}^2} = 1.$$

The two parts of the integral above correspond to kinetic and potential energy, respectively. We will continue this investigation in Example 3.22.

1.5 Optimal Saving and Consumption

Consider a capitalist worker earning a (constant) wage w per year, which the worker can either spend on consumption or save. Denote by $S(t)$ the accumulated savings at time t, where $t \in [0, T]$ is in years, with $t = 0$ denoting the start of employment and $t = T$ retirement. Let $C(t) \geq 0$ be the consumption rate (consumption per time) at time t. On the saved capital, the worker earns interest, say with gross-continuous rate $\rho > 0$, meaning that a capital amount $m > 0$ grows as $\exp(\rho t)m$. If we were given an effective APR $\rho_1 > 0$ instead of ρ, then $\rho = \ln(1 + \rho_1)$. We further assume that the salary is paid continuously, not in intervals, for simplicity. So, w is really the rate of pay, given in money per time. Then, the worker's savings evolve according to the differential equation

$$\dot{S}(t) = w + \rho S(t) - C(t). \tag{1.2}$$

We now make the (totally unreasonable) assumption that the worker's happiness only depends on his consumption rate. Suppose that our worker wants to optimize total life happiness by finding the optimal amount of consumption at any given time. So, if we denote by $U(C)$ the *marginal utility function*, that is, the marginal "happiness" due to the consumption rate C, our worker wants to find $C : [0, T] \to \mathbb{R}$ such that

$$\mathscr{H}[C] := \int_0^T U(C(t))\,\mathrm{d}t$$

is maximized. The choice of U depends on our worker's personality, but it is sensible to assume that there is a *law of diminishing returns*, i.e., for twice as much consumption, our worker is happier, but not twice as happy. So, let us assume $U' > 0$ and $U'(C) \to 0$ as $C \to \infty$. Also, we should have $U(0) = 0$ (starvation). Moreover, it is realistic for U to be *concave*, which in particular implies that there are no local maxima. One function that satisfies all of these requirements is

$$U(C) = \ln(1 + C), \qquad C > 0.$$

Let us also assume that the worker starts with no savings, $S(0) = 0$, and wants to retire with savings $S(T) = S_T \geq 0$. Rearranging (1.2) for $C(t)$ and plugging this into the formula for $\mathscr{H}$, we therefore want to solve the *optimal saving problem*

$$\begin{cases} \text{Minimize} \quad \mathscr{F}[S] := \displaystyle\int_0^T -\ln(1 + w + \rho S(t) - \dot{S}(t))\,\mathrm{d}t \\ \text{subject to} \quad S(0) = 0,\ S(T) = S_T \geq 0,\ C(t) := w + \rho S(t) - \dot{S}(t) \geq 0. \end{cases}$$

This will be solved in Example 3.7.

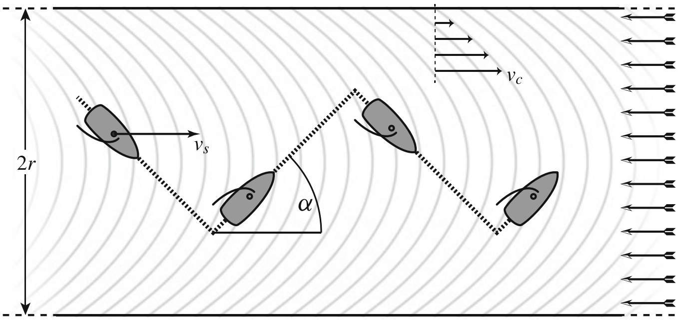


Fig. 1.4 Sailing against the wind in a channel

1.6 Sailing Against the Wind

Every sailor knows how to sail against the wind by "beating": One has to sail at an angle of approximately 45° to the wind (in real boats, the maximum might be at a lower angle, i.e., "closer to the wind"), then tack (turn the bow through the wind) and finally, after the sail has caught the wind on the other side, continue again at approximately 45° to the wind. Repeating this procedure makes the boat follow a zig-zag motion, which gives a net movement directly against the wind, see Figure 1.4. A mathematically inclined sailor might ask the question of "how often to tack". In an idealized model we can assume that the wind has the same speed and direction everywhere, tacking costs no time, and the forward sailing speed v_s of the boat depends on the angle α to the wind as follows (at least qualitatively):

$$v_s(\alpha) = v_{\max} \cdot \frac{1 - \cos(4\alpha)}{2},$$

where $v_{\max}$ is the maximum speed of the boat at the current wind speed, which we assume to be constant. Then, $v_s(\alpha)$ is non-negative and has maxima at $\alpha = \pm 45°$.

Assume furthermore that our sailor is sailing along a straight river with the current. Now, the current is fastest in the middle of the river and disappears at the banks. In fact, a good approximation for the flow speed is given by the formula of *Poiseuille (channel) flow*, which can be derived from the flow equations of fluids: At distance r from the center of the river the current's flow speed is approximately

$$v_c(r) := v_{\text{flow}}\left(1 - \frac{r^2}{R^2}\right),$$

where $R > 0$ is half the width of the river.

If we denote by $r(t)$ the distance of the boat from the middle of the channel at time $t \in [0, T]$, then the total speed (called the "velocity made good" in sailing parlance) is

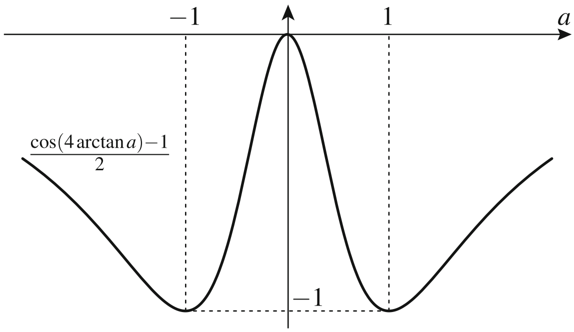


Fig. 1.5 The double-well potential in the sailing example

$$\begin{aligned} v(t) &:= v_s(\arctan r'(t)) + v_c(r(t)) \\ &= v_{\max} \cdot \frac{1-\cos(4\arctan r'(t))}{2} + v_{\text{flow}}\left(1 - \frac{r(t)^2}{R^2}\right). \end{aligned}$$

The key to understanding this problem is the observation that the function given by $a \mapsto (\cos(4\arctan a) - 1)/2$ has precisely two minima, namely at $a = \pm 1$. We say that this function is a *double-well potential*, see Figure 1.5.

The total forward distance traveled over the time interval $[0, T]$ is

$$\int_0^T v(t)\, \mathrm{d}t = \int_0^T v_{\max} \cdot \frac{1-\cos(4\arctan r'(t))}{2} + v_{\text{flow}}\left(1 - \frac{r(t)^2}{R^2}\right) \mathrm{d}t.$$

If we also require the initial and terminal conditions $r(0) = r(T) = 0$, we arrive at the *optimal beating problem*

$$\begin{cases} \text{Minimize} \quad \mathscr{F}[r] := \displaystyle\int_0^T v_{\max} \cdot \frac{\cos(4\arctan r'(t)) - 1}{2} + v_{\text{flow}}\left(\frac{r(t)^2}{R^2} - 1\right) \mathrm{d}t \\ \text{subject to} \quad r(0) = r(T) = 0, \ |r(t)| \leq R. \end{cases}$$

Our intuition tells us that in this idealized model, where tacking costs no time, we should be tacking "infinitely fast" in order to stay in the middle of the river. Later, once we have advanced tools at our disposal, we will make this idea precise, see Example 7.13.

1.7 Hyperelasticity

Elasticity theory is one of the most important theories of *continuum mechanics*, that is, the study of the mechanics of (idealized) continuous media. We will not

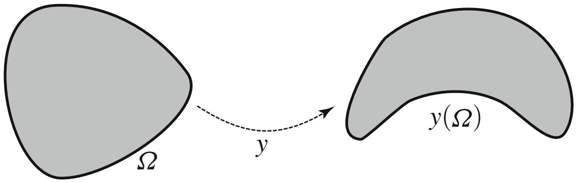


Fig. 1.6 A deformed body

go into much detail about elasticity modeling here and refer to [64] for a thorough introduction.

Consider a body of mass occupying a bounded and connected domain $\Omega \subset \mathbb{R}^3$ such that $\partial\Omega$ is a Lipschitz manifold (the union of finitely many Lipschitz graphs). We call Ω the *reference configuration*. If we deform the body, any *material point* $x \in \Omega$ is mapped to a *spatial point* $y(x) \in \mathbb{R}^3$ and we call $y(\Omega)$ the *deformed configuration*, see Figure 1.6. We also require that $y\colon \Omega \to y(\Omega)$ is a differentiable bijection and that it is *orientation-preserving*, i.e.,

$$\det \nabla y(x) > 0, \qquad x \in \Omega.$$

For convenience let us also introduce the *displacement*

$$u(x) := y(x) - x.$$

Next, we need a measure of local "stretching", called a *strain tensor*, which should serve as the argument for a local energy density. On physical grounds, *rigid body motions*, that is, deformations of the form $u(x) = Rx + u_0$ with a rotation $R \in \mathbb{R}^{3\times 3}$ ($R^T = R^{-1}$, $\det R = 1$) and $u_0 \in \mathbb{R}^3$, should not cause strain. In this sense, strain measures the deviation of the deformation from a rigid body motion. One common choice is the *Green–St. Venant strain tensor*

$$G := \frac{1}{2}\big(\nabla u + \nabla u^T + \nabla u^T \nabla u\big). \tag{1.3}$$

We first consider fully nonlinear ("finite strain") elasticity. For our purposes we simply postulate the existence of a stored-energy density $W\colon \mathbb{R}^{3\times 3} \to [0, \infty]$ and an external body force field $b\colon \Omega \to \mathbb{R}^3$ (e.g. gravity) such that

$$\mathscr{F}[y] := \int_\Omega W(\nabla y(x)) - b(x)\cdot y(x)\,\mathrm{d}x$$

represents the total elastic energy stored in the system. If the elastic energy can be written in this way as

$$\int_\Omega W(\nabla y(x))\,\mathrm{d}x,$$

we call the material *hyperelastic*. In applications, W is sometimes given as depending on the Green–St. Venant strain tensor G instead of ∇y, but for the mathematical theory the above form is more convenient. We require several properties of W:

(i) *Norming:* $W(\mathrm{Id}) = 0$ (the undeformed state costs no energy).
(ii) *Frame-indifference:* $W(QA) = W(A)$ for all $Q \in \mathrm{SO}(3)$, $A \in \mathbb{R}^{3\times 3}$.
(iii) *Infinite compression costs infinite energy:* $W(A) \to +\infty$ as $\det A \downarrow 0$.
(iv) *Infinite stretching costs infinite energy:* $W(A) \to +\infty$ as $|A| \to \infty$.

The fundamental task of nonlinear hyperelasticity is to minimize $\mathscr{F}$ as above over all $y\colon \Omega \to \mathbb{R}^3$ with given boundary values. Of course, it is not a priori clear in which space we should look for a solution. Indeed, this depends on the growth properties of W. For example, for the prototypical choice

$$W(A) := \mathrm{dist}(A, \mathrm{SO}(3))^2, \qquad \text{where} \qquad \mathrm{dist}(A, K) := \inf_{B\in K} |A - B|,$$

we would look for square-integrable functions. However, this W does not satisfy (iii) from our list of requirements. More realistic in applications are the *Mooney–Rivlin materials*, where W is of the form

$$W(A) := \begin{cases} a|A|^2 + b|\operatorname{cof} A|^2 + \Gamma(\det A) & \text{if } \det A > 0, \\ +\infty & \text{if } \det A \le 0, \end{cases}$$

with $a, b > 0$ and $\Gamma(d) = \alpha d^2 - \beta \log d$ for $\alpha, \beta > 0$. If $b = 0$, the material is called *neo-Hookean*. An even larger class is given by the *Ogden materials*, for which

$$W(A) := \sum_{i=1}^{M} a_i \operatorname{tr}\bigl[(A^T A)^{\gamma_i/2}\bigr] + \sum_{j=1}^{N} b_j \operatorname{tr} \operatorname{cof}\bigl[(A^T A)^{\delta_j/2}\bigr] + \Gamma(\det A),$$

where $M, N \in \mathbb{N}$, $a_i > 0$, $\gamma_i \ge 1$, $b_j > 0$, $\delta_j \ge 1$, and $\Gamma\colon \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is a convex function with $\Gamma(d) \to +\infty$ as $d \downarrow 0$, $\Gamma(d) = +\infty$ for $d \le 0$. These materials occur in a wide range of applications. We will consider such problems in Example 6.8.

In the setting of *linearized elasticity*, we make the "small strain" assumption that y is an orientation-preserving bijection with ∇u "small" such that the quadratic term in (1.3) can be neglected. In this case, we work with the *linearized strain tensor*

$$\mathscr{E}u := \frac{1}{2}\bigl(\nabla u + \nabla u^T\bigr).$$

Now, the displacements that do not create strain are precisely the skew-affine maps $u(x) = Wx + u_0$ with $W^T = -W$ and $u_0 \in \mathbb{R}$. This becomes more meaningful if we consider a bit more algebra: The Lie group SO(3) of rotations has as its Lie algebra $\mathrm{Lie}(\mathrm{SO}(3)) = \mathfrak{so}(3)$, the space of all skew-symmetric matrices, which then can be seen as "infinitesimal rotations".

For linearized elasticity we consider an energy of the special quadratic form

$$\mathscr{W}[u] := \int_{\Omega} \frac{1}{2} \mathscr{E}u(x) : \mathbf{C}(x)\, \mathscr{E}u(x) \, \mathrm{d}x,$$

where $\mathbf{C}(x) = \mathbf{C}_{jl}^{ik}(x)$ ($x \in \Omega$) is a symmetric, positive definite ($A : \mathbf{C}(x)A \geq c|A|^2$ for some $c > 0$) fourth-order tensor, called the *elasticity tensor*.

For homogeneous, isotropic media, $\mathbf{C}$ does not depend on x or the direction of strain, which translates into the additional condition

$$(AQ) : \mathbf{C}(AQ) = A : \mathbf{C}A \qquad \text{for all } A \in \mathbb{R}^{3\times 3},\ Q \in \mathrm{SO}(3).$$

In this case, it can be shown that $\mathscr{W}$ simplifies to

$$\mathscr{W}[u] = \int_{\Omega} \mu|\mathscr{E}u(x)|^2 + \frac{1}{2}\Big(\kappa - \frac{2}{3}\mu\Big)|\operatorname{tr} \mathscr{E}u(x)|^2 \, \mathrm{d}x$$

for $\mu > 0$ the *shear modulus* and $\kappa > 0$ the *bulk modulus*, which are material constants. For example, for cold-rolled steel $\mu \approx 75$ GPa and $\kappa \approx 160$ GPa. As in the nonlinear setting, we then consider the minimization problem for the total energy

$$\mathscr{F}[u] := \int_{\Omega} \mu|\mathscr{E}u(x)|^2 + \frac{1}{2}\Big(\kappa - \frac{2}{3}\mu\Big)|\operatorname{tr} \mathscr{E}u(x)|^2 - b(x) \cdot u(x) \, \mathrm{d}x,$$

where $b \colon \Omega \to \mathbb{R}^3$ is the external body force (now with respect to u). We will consider this functional further in Examples 2.12 and 3.16.

1.8 Microstructure in Crystals

In a single crystal of a metal like iron or an alloy like CuAlNi (Copper–Aluminium–Nickel), the atoms are arranged in a regular lattice. Assume that such a material specimen occupies an open, bounded, and connected reference domain $\Omega \subset \mathbb{R}^3$. We then want to determine the resulting deformed shape subject to external forces.

It turns out that on a microscopic scale the deformation of a single crystal (subject to given boundary conditions) often exhibits very fine locally periodic oscillations in the deformation, that is, the crystal exhibits *microstructure*, see Figure 1.7. This behavior has profound implications for the macroscopic behavior of the material.

The fundamental Cauchy–Born hypothesis postulates that for small linear displacements the crystal lattice atoms will follow this displacement (this assumption is often made, but is not always justified, see [70, 105, 108, 128]). Assuming that the microstructure does not reach down to atomic length scales, we can then model the crystal as a *continuum* and assign the energy density $W(F) \geq 0$ to the linear deformation $x \mapsto Fx$. The crucial point here is that, thanks to the Cauchy–Born

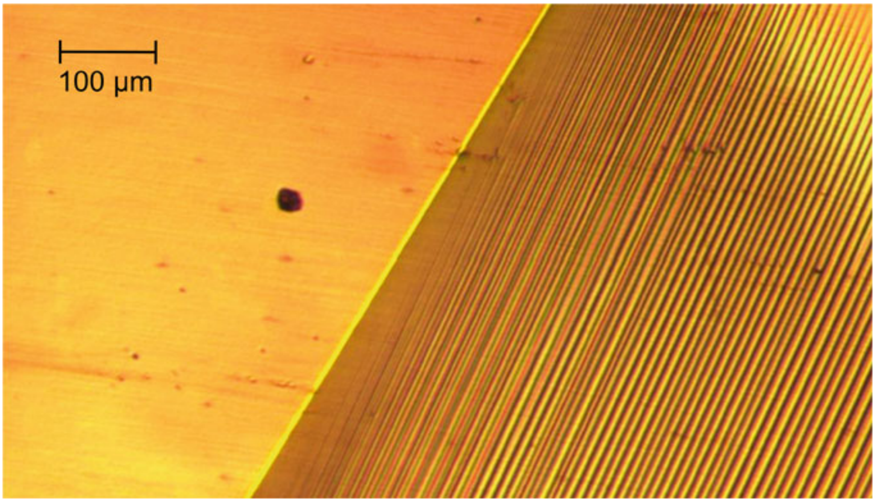


Fig. 1.7 CuAlNi microstructure undergoing a transition from cubic austenite (left) to orthorhombic martensite (right), see [241] for more details on this particular microstructure (source: *original micrograph by Hanuš Seiner, reproduced with kind permission*)

hypothesis, W depends only on F and no other "microscopic structure" of the crystal, at least for small to moderate crystal deformations. In this approach, the total energy of a deformation $y \colon \Omega \to \mathbb{R}^3$ is given as

$$\mathscr{F}[y] := \int_\Omega W(\nabla y(x))\,\mathrm{d}x, \qquad y \colon \Omega \to \mathbb{R}^3.$$

Here, on $W \colon \mathbb{R}^{3\times 3} \to [0, \infty)$ we make the following assumptions:

(i) *Norming:* $W(\mathrm{Id}) = 0$ (the undeformed state costs no energy).
(ii) *Frame-indifference:* $W(QA) = W(A)$ for all $Q \in \mathrm{SO}(3)$, $A \in \mathbb{R}^{3\times 3}$.
(iii) *Symmetry-invariance:* $W(AS) = W(S)$ for all $S \in \mathscr{S}$ and all $A \in \mathbb{R}^{3\times 3}$, where $\mathscr{S} \subset \mathrm{SO}(3)$ is the compact (symmetry) point group of the crystal.

The basic variational postulate is that the observed macroscopic deformation is a minimizer of $\mathscr{F}$ under the given boundary conditions. In fact, it is often experimentally observed that the deformation $y \colon \Omega \to \mathbb{R}^3$ is close to a *pointwise* minimizer of the integrand, at least in a very large portion of Ω. Thus, we are led to consider the *differential inclusion*

$$\nabla y(x) \in K := W^{-1}(0) = \left\{ A \in \mathbb{R}^{3\times 3} \;:\; W(A) = \min W \right\}, \qquad x \in \Omega.$$

The set K is compact in the study of crystals, but other applications also lead to differential inclusions with non-compact K.

In concrete applications, one usually has the following (idealized) situation: Above a critical temperature, K is simply SO(3), which is the simplest possible set that is compatible with the frame-indifference (ii). This is called the *austenite phase*. Below the critical temperature, however, the material undergoes a solid–solid phase transition to the *martensite phase*, where K is the union of several *wells*, that is,

$$K = \mathrm{SO}(3)U_1 \cup \cdots \cup \mathrm{SO}(3)U_N$$

for distinct matrices $U_1, \ldots, U_N \in \mathbb{R}^{3\times 3}$ with $\det U_i > 0$ ($i = 1, \ldots, N$). By the polar decomposition of matrices with positive determinants into a product of a rotation and a symmetric positive definite matrix, we can assume that all the U_i are symmetric and positive definite.

If $N \geq 2$ and other *compatibility conditions* between the matrices U_i are satisfied, microstructure can indeed be observed. It should be noted that while our model as formulated above may imply "infinitely fast" oscillations in the microstructure, in reality other (atomistic) effects limit the length scales that are observed.

As a concrete example, the NiAl (Nickel–Aluminium) alloy undergoes a *cubic-to-tetragonal* phase transition and below the critical temperature we have

$$K = \mathrm{SO}(3)U_1 \cup \mathrm{SO}(3)U_2 \cup \mathrm{SO}(3)U_3$$

with

$$U_1 = \begin{pmatrix} \beta & & \\ & \alpha & \\ & & \alpha \end{pmatrix}, \qquad U_2 = \begin{pmatrix} \alpha & & \\ & \beta & \\ & & \alpha \end{pmatrix}, \qquad U_3 = \begin{pmatrix} \alpha & & \\ & \alpha & \\ & & \beta \end{pmatrix},$$

for $\alpha \approx 0.9392$, $\beta \approx 1.1302$, see [41, 107].

As another example, the CuAlNi alloy undergoes a *cubic-to-orthorhombic* phase transition and below the critical temperature we have

$$K = \mathrm{SO}(3)U_1 \cup \cdots \cup \mathrm{SO}(3)U_6$$

with

$$U_1 = \begin{pmatrix} \xi & 0 & \eta \\ 0 & \beta & 0 \\ \eta & 0 & \xi \end{pmatrix}, \qquad U_2 = \begin{pmatrix} \xi & 0 & -\eta \\ 0 & \beta & 0 \\ -\eta & 0 & \xi \end{pmatrix}, \qquad U_3 = \begin{pmatrix} \xi & \eta & 0 \\ \eta & \xi & 0 \\ 0 & 0 & \beta \end{pmatrix},$$

$$U_4 = \begin{pmatrix} \xi & -\eta & 0 \\ -\eta & \xi & 0 \\ 0 & 0 & \beta \end{pmatrix}, \qquad U_5 = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \xi & \eta \\ 0 & \eta & \xi \end{pmatrix}, \qquad U_6 = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \xi & -\eta \\ 0 & -\eta & \xi \end{pmatrix},$$

where

$$\xi = \frac{\alpha + \gamma}{2}, \qquad \eta = \frac{\alpha - \gamma}{2}$$

for $\alpha \approx 1.0619$, $\beta \approx 0.9178$, $\gamma \approx 1.0230$, see [41, 104].

A first mathematical question that can be asked about such microstructures concerns their effective representation: What are the salient features of the oscillations in the material and how can they be captured mathematically? Moreover, which deformations with linear boundary values $x \mapsto Fx$ have almost zero energy? It turns out that by relying on very high-frequency oscillations, the set of these F can actually be much larger than K and defines a certain "hull" of K. This hull explains the observed microstructure, as we will see in Chapter 9.

In engineering applications, one striking property of NiAl and CuAlNi is the *shape-memory effect*, where a material specimen "remembers" the shape it had when it was hotter than the *critical temperature*. After cooling, the specimen can be freely deformed, but when it is again heated above the critical temperature, it "snaps back" into its original shape. This effect is directly related to the formation of microstructure (below the critical temperature), which accommodates the deformations through microstructure changes, but without changing the structure of the crystal lattice itself. Upon heating the specimen above the critical temperature, all microstructure disappears and the original shape (which is determined by the crystal lattice in the cubic phase) reappears. Note that all the matrices $U_1, U_2, \ldots$ for both NiAl and CuAlNi have determinant very close to 1, which is a common feature of shape-memory alloys, because it is necessary for the *self-accommodation* effect, where upon cooling through the critical temperature the microstructure arranges itself in a such way that the macroscopic shape does not change. See [41] for a detailed study of the shape-memory effect.

In Chapters 8, 9 we will consider the basic principles underlying this problem, see Examples 8.10, 9.17. Concrete applications are left to more specialized treatises like [41, 100].

1.9 Phase Transitions

Consider a (bounded, open, connected) container $\Omega \subset \mathbb{R}^d$ ($d \in \{2, 3\}$ are the physically interesting cases) containing two mixed fluids. We let $\rho\colon \Omega \to [0, 1]$ model the density of the first fluid and prescribe the relative amounts of the two fluids by requiring that

$$\int_\Omega \rho(x)\,\mathrm{d}x = \tilde{\gamma} \in (0, |\Omega|). \tag{1.4}$$

The Gibbs free energy of the mixture is given as

$$\mathscr{G}[\rho] := \int_\Omega W_0(\rho(x))\,\mathrm{d}x, \qquad \rho\colon \Omega \to [0, 1],$$

where $W_0\colon \mathbb{R} \to [0, \infty)$. Often, the energy density W_0 has precisely two minima $\alpha, \beta \in [0, 1]$ with $\alpha < \beta$ and thus W_0 is a *double-well potential*. In the simplest case, the fluids do not mix well (e.g. water and oil), and the two local minima of W_0 are

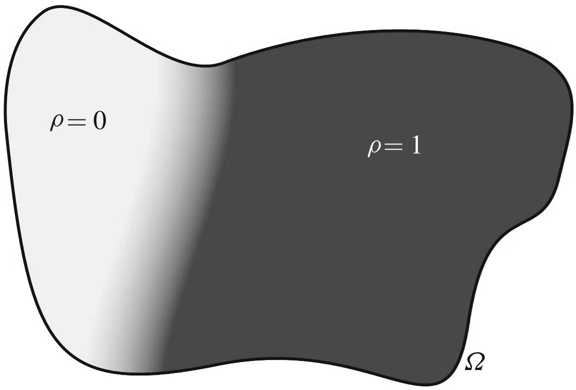


Fig. 1.8 A phase transition

located at $\alpha = 0$ (all oil) and $\beta = 1$ (all water), see Figure 1.8. It is a classical problem, first considered by Cahn–Hilliard and Gurtin (see [54, 146, 147]), to determine the equilibrium mixture, i.e., to find a minimizer of $\mathscr{G}$. In order for the problem to be interesting, we assume that

$$\tilde{\gamma} \in (\alpha|\Omega|, \beta|\Omega|).$$

However, in the above form this problem is not well-posed: We can just choose any $\rho \colon \Omega \to [0, 1]$ with $\rho(x) \in \{\alpha, \beta\}$ for all $x \in \Omega$ such that (1.4) holds. This will be a minimizer of $\mathscr{F}_0$, but this formulation is unsatisfactory: The shape of the two *phases*

$$E_\alpha := \{ x \in \Omega \;:\; \rho(x) = \alpha \}, \qquad E_\beta := \{ x \in \Omega \;:\; \rho(x) = \beta \},$$

is clearly not uniquely determined and *no regularity* can be assumed on the *phase boundary* $\partial E_\alpha \cap \Omega = \partial E_\beta \cap \Omega$. The remedy to this problem comes from physics in the form of the additional assumption that the interface between E_α and E_β should have the minimal surface area among all competitors. We can incorporate this minimum principle in two different ways. First, we can penalize changes in the function ρ by adding a (quadratic) gradient term and set

$$\widetilde{\mathscr{F}}_\varepsilon[\rho] := \int_\Omega W_0(\rho(x)) + \varepsilon^2 |\nabla \rho(x)|^2 \,\mathrm{d}x,$$

where $\varepsilon > 0$ is a (small) parameter. For purely mathematical reasons it turns out to be beneficial to transform this functional into

$$\mathscr{F}_\varepsilon[u] := \int_\Omega \frac{1}{\varepsilon} W(u(x)) + \varepsilon |\nabla u(x)|^2 \,\mathrm{d}x, \qquad u \colon \Omega \to [-1, 1],$$

where $W \colon \mathbb{R} \to [0, \infty)$ is given as (see, for instance, Figure 7.1 for an illustration of such a double-well potential)

$$W(s) := W_0\left(\alpha + \frac{s+1}{2}(\beta - \alpha)\right) - W_0(\alpha) \cdot \frac{1-s}{2} - W_0(\beta) \cdot \frac{s-1}{2},$$

so that $W(\pm 1) = 0$ are the two global minima of W. From (1.4) it can be verified easily that each ρ minimizing $\widetilde{\mathscr{F}}_\varepsilon$ corresponds to precisely one u minimizing $\mathscr{F}_\varepsilon$ via the transformation

$$\rho(x) = \alpha + \frac{u(x)+1}{2}(\beta - \alpha),$$

whereby one can compute that for such pairs (u, ρ),

$$\widetilde{\mathscr{F}}_\varepsilon[\rho] = \frac{\varepsilon(\beta-\alpha)}{2}\mathscr{F}_{\varepsilon(\beta-\alpha)/2}[u]$$

and

$$\int_\Omega u(x)\,\mathrm{d}x = \gamma := \frac{2(\tilde{\gamma} - \alpha|\Omega|)}{\beta - \alpha} - |\Omega| \in (-|\Omega|, |\Omega|).$$

We remark that the balancing of the ε-terms turns out to be necessary if we want to consider the limit as $\varepsilon \downarrow 0$. Notice also that a minimizer of $\mathscr{F}_\varepsilon$ will have a square-integrable gradient and so, besides the *pure phases*

$$E_{\pm 1} := \left\{ x \in \Omega \ : \ u(x) = \pm 1 \right\},$$

there will also be a non-empty *transition region*

$$\Delta := \left\{ x \in \Omega \ : \ u(x) \in (-1, 1) \right\}.$$

Intuitively, Δ will shrink to a phase interface surface as $\varepsilon \downarrow 0$ since the regularizing effect of the gradient term in $\mathscr{F}_\varepsilon$ gets weaker as $\varepsilon \downarrow 0$.

An alternative way to model the physical situation is to prescribe that $u\colon \Omega \to \{-1, 1\}$ splits the domain into the two phases $E_{\pm 1} = E_{\pm 1}(u)$ and the transition region is empty. In this case, we could consider those u minimizing the surface tension between the phases,

$$\mathscr{F}_0[u] := \sigma \operatorname{Per}_\Omega(E_{-1}(u)),$$

to be the physically relevant solutions. Here, the *perimeter* $\operatorname{Per}_\Omega(E_{-1}(u))$ should be understood as the surface area of $\partial E_{-1}(u) \cap \Omega$, at least if $\partial E_{-1}(u) \cap \Omega$ is a smooth manifold (with boundary). In more general situations the definition of this quantity will have to be suitably extended. The constant $\sigma > 0$ takes the role of a *surface tension*. Notice that for $\mathscr{F}_0$ as defined above to make sense, the set $\partial E_{-1}(u) \cap \Omega$ has to have some regularity, so that $\mathscr{F}_0[u] < \infty$.

An important question about the above functionals is the following: As $\varepsilon \downarrow 0$, does $\mathscr{F}_\varepsilon$ "converge" to $\mathscr{F}_0$ in a sense that entails the convergence of minimizers and minimum values (for a suitably chosen σ)? We will return to this question in Chapter 13, in particular in Example 13.10.

1.10 Composite Elastic Materials

Assume we are given a linearly elastic material specimen (like in Section 1.7) occupying the domain $\Omega \subset \mathbb{R}^3$, whose stored energy for a displacement $u\colon \Omega \to \mathbb{R}^3$ is

$$\mathscr{W}[u] := \int_\Omega \frac{1}{2}\mathscr{E}u(x) : \mathbf{C}(x)\,\mathscr{E}u(x)\;\mathrm{d}x,$$

where, as in Section 1.7, $\mathbf{C}(x) = \mathbf{C}^{ik}_{jl}(x)$ $(x \in \Omega)$ is a symmetric, positive definite fourth-order elasticity tensor. Here we suppose in addition that $\mathbf{C}$ depends on x in an ε-periodic manner for a small $\varepsilon > 0$. For instance, we could imagine our specimen to be a *composite* consisting of thin alternating material layers of two different types, see Figure 1.9. We denote the elasticity tensors of these layers by $\mathbf{C}_1$ and $\mathbf{C}_2$, respectively, and assume that the layers alternate in the first coordinate direction with thicknesses $\theta\varepsilon$ and $(1-\theta)\varepsilon$, respectively, where $\theta \in (0, 1)$. Then, the elastic energy has the form

$$\mathscr{W}_\varepsilon[u] := \int_\Omega \frac{1}{2}\mathscr{E}u(x) : \mathbf{C}\Big(\frac{x}{\varepsilon}\Big)\,\mathscr{E}u(x)\;\mathrm{d}x,$$

where

$$\mathbf{C}(x) = \mathbf{C}_1 + (\mathbf{C}_2 - \mathbf{C}_1)h(x_1), \qquad x \in \mathbb{R}^3,$$

and

$$h(t) := \begin{cases} 0 & \text{if } t - \lfloor t \rfloor \le \theta, \\ 1 & \text{if } t - \lfloor t \rfloor > \theta. \end{cases}$$

Here, $\lfloor t \rfloor$ denotes the largest integer less than or equal to $t \in \mathbb{R}$. The total energy to be minimized is

$$\begin{aligned} \mathscr{F}_\varepsilon[u] &:= \int_\Omega f_\varepsilon\Big(\frac{x}{\varepsilon}, \nabla u(x)\Big) - b(x) \cdot u(x)\;\mathrm{d}x \\ &:= \int_\Omega \frac{1}{2}\mathscr{E}u(x) : \mathbf{C}\Big(\frac{x}{\varepsilon}\Big)\,\mathscr{E}u(x) - b(x) \cdot u(x)\;\mathrm{d}x, \end{aligned}$$

where $b\colon \Omega \to \mathbb{R}^3$ is the external body force.

In many applications, one is predominantly interested in the *homogenized* behavior of the specimen, that is, its large-scale, averaged properties. Mathematically, this corresponds to a form of "variational limit" of the $\mathscr{F}_\varepsilon$, which entails the convergence of minimizers and minimum values. Ideally, we want to compute a homogenized density $f_{\mathrm{hom}}\colon \mathbb{R}^{3\times 3} \to \mathbb{R}$ (not x-dependent), such that $\mathscr{F}_\varepsilon$ "variationally converges"

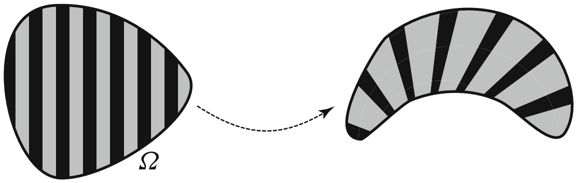


Fig. 1.9 A deformed composite material body

to a limit $\mathscr{F}_0$ of the form

$$\mathscr{F}_0[u] = \int_\Omega f_{\mathrm{hom}}(\nabla u(x)) - b(x) \cdot u(x) \, \mathrm{d}x.$$

The following questions are of importance:

- Does an $\mathscr{F}_0$ as above exist and can it be written as an *integral* functional?
- Does f_{hom} (if it exists) have the same quadratic structure as the f_ε, that is, is there a symmetric, positive definite fourth-order tensor $\mathbf{C}_{\mathrm{hom}}$ such that $f_{\mathrm{hom}}(A) = \frac{1}{2} A^{\mathrm{sym}} : \mathbf{C}_{\mathrm{hom}} A^{\mathrm{sym}}$ (here, A^{sym} is the symmetric part of A)?
- In the special case when $\mathbf{C}_1 = \alpha \, \mathbf{I}$, $\mathbf{C}_2 = \beta \, \mathbf{I}$ for $\alpha, \beta > 0$ (here, $\mathbf{I}$ denotes the tensor such that $A : \mathbf{I} B = A : B$, so $\mathbf{I}^{ik}_{jl} = \delta_{ik}\delta_{jl}$), is $\mathbf{C}_{\mathrm{hom}}$ (if it exists) also of the form $\mathbf{C}_{\mathrm{hom}} = \gamma \, \mathbf{I}$ for some $\gamma > 0$?

We will investigate these questions in Example 13.25.

Chapter 2
Convexity

In this chapter we start to develop the mathematical theory that will allow us to analyze the problems presented in the introduction, and many more. The basic minimization problem that we are considering is the following:

$$\begin{cases} \text{Minimize} & \mathscr{F}[u] := \displaystyle\int_\Omega f(x, u(x), \nabla u(x)) \,\mathrm{d}x \\ \text{over all} & u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m) \text{ with } u|_{\partial\Omega} = g. \end{cases}$$

Here, and throughout the text if not stated otherwise, we will make the standard assumption that $\Omega \subset \mathbb{R}^d$ is a **bounded Lipschitz domain**, that is, Ω is open, bounded, connected, and has a boundary that is the union of finitely many Lipschitz manifolds. The function

$$f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \to \mathbb{R}$$

is required to be measurable in the first and (jointly) continuous in the second and third arguments, which makes f a so-called **Carathéodory integrand**. Furthermore, in this chapter we (usually) let $p \in (1, \infty)$ and for the prescribed boundary values g we assume

$$g \in \mathrm{W}^{1-1/p,p}(\partial\Omega; \mathbb{R}^m).$$

In this context recall that $\mathrm{W}^{1-1/p,p}(\partial\Omega; \mathbb{R}^m)$ is the space of traces of Sobolev maps in $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$, see Appendix A.5 for some background on Sobolev spaces.

Below, we will investigate the solvability of the above minimization problem (under additional technical assumptions). We first present the main ideas of the so-called Direct Method of the calculus of variations in an abstract setting, namely for (nonlinear) functionals on Banach spaces. Then we will begin our study of integral functionals, where we will in particular take a close look at the way in which *convexity* properties of f in its gradient (third) argument determine whether $\mathscr{F}$ is

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8_2

lower semicontinuous. We also consider the question of which function space should be chosen for the candidate functions. Finally, we explain basic aspects of general convex analysis, in particular the Legendre–Fenchel duality.

2.1 The Direct Method

Fundamental to all of the existence theorems in this book is the conceptually simple, yet powerful, *Direct Method* of the calculus of variations. It is called "direct" since we prove the existence of solutions to minimization problems without the detour through a differential equation.

Let X be a complete metric space (e.g. a Banach space with the norm topology or a closed and convex subset of a reflexive Banach space with the weak topology). Let $\mathscr{F} : X \to \mathbb{R} \cup \{+\infty\}$ be our *objective functional* that we require to satisfy the following two assumptions:

(H1) **Coercivity:** For all $\Lambda \in \mathbb{R}$, the sublevel set

$$\left\{ u \in X \;:\; \mathscr{F}[u] \leq \Lambda \right\} \qquad \text{is sequentially precompact,}$$

that is, if $\mathscr{F}[u_j] \leq \Lambda$ for a sequence $(u_j) \subset X$ and some $\Lambda \in \mathbb{R}$, then (u_j) has a converging subsequence in X.

(H2) **Lower semicontinuity:** For all sequences $(u_j) \subset X$ with $u_j \to u$ in X it holds that

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

Note that here and in all of the following we use the *sequential* notions of compactness and lower semicontinuity, which are better suited to our needs than the corresponding topological concepts. For more on this point see the notes section at the end of this chapter.

The Direct Method for the abstract problem

$$\text{Minimize } \mathscr{F}[u] \text{ over all } u \in X \tag{2.1}$$

is encapsulated in the following simple result.

Theorem 2.1. *Assume that $\mathscr{F}$ is both coercive and lower semicontinuous. Then, the abstract minimization problem (2.1) has at least one solution, that is, there exists a $u_* \in X$ with $\mathscr{F}[u_*] = \min\{ \mathscr{F}[u] \;:\; u \in X \}$.*

Proof. Let us assume that there exists at least one $u \in X$ such that $\mathscr{F}[u] < +\infty$; otherwise, any $u \in X$ is a "solution" to the (degenerate) minimization problem.

To construct a minimizer we take a **minimizing sequence** $(u_j) \subset X$ such that

$$\lim_{j\to\infty} \mathscr{F}[u_j] \to \alpha := \inf \left\{ \mathscr{F}[u] \;:\; u \in X \right\} < +\infty.$$

Then, there exists a $\Lambda \in \mathbb{R}$ such that $\mathscr{F}[u_j] \le \Lambda$ for all $j \in \mathbb{N}$. Hence, by the coercivity, we may select a subsequence, which we do not make explicit in our notation, such that

$$u_j \to u_* \in X.$$

By the lower semicontinuity we immediately conclude that

$$\alpha \le \mathscr{F}[u_*] \le \liminf_{j\to\infty} \mathscr{F}[u_j] = \alpha.$$

Thus, $\mathscr{F}[u_*] = \alpha$ and u_* is the sought minimizer. □

Example 2.2. Using the Direct Method, one can easily see that the lower semicontinuous function

$$h(t) := \begin{cases} 1-t & \text{if } t < 0, \\ t & \text{if } t \ge 0, \end{cases}$$

has the minimizer $t = 0$.

Despite its nearly trivial proof, the Direct Method is very useful and flexible in applications. Indeed, it pushes the difficulty in proving the existence of a minimizer into establishing coercivity and lower semicontinuity. This, however, is a big advantage, since we have many tools at our disposal to establish these two hypotheses separately. In particular, for integral functionals, lower semicontinuity is tightly linked to *convexity* properties of the integrand, as we will see throughout this book.

At this point it is crucial to observe how coercivity and lower semicontinuity interact with the topology on X: If we choose a stronger topology, i.e., one for which there are fewer converging sequences, then it is easier for $\mathscr{F}$ to be lower semicontinuous, but harder for $\mathscr{F}$ to be coercive. The opposite holds if we choose a weaker topology. In the mathematical treatment of a problem from applications, we are most likely in a situation where $\mathscr{F}$ and the set X are given. We then need to find a suitable topology in which we can establish both coercivity and lower semicontinuity. It is remarkable that the topology that turns out to be *mathematically* convenient is often also *physically* relevant.

In this book, X will always be an infinite-dimensional Banach space (or a subset thereof) and we have a real choice between using the strong or weak convergence. Usually, it turns out that coercivity with respect to the strong convergence is *false* since strongly compact sets in infinite-dimensional spaces are very restricted, whereas coercivity with respect to the weak convergence is true under reasonable assumptions. On the other hand, while strong lower semicontinuity poses few challenges, lower semicontinuity with respect to weakly converging sequences is a more delicate matter and we will spend considerable time on this topic.

As a result of this discussion, we will almost always use the Direct Method in the following version:

Theorem 2.3. *Let X be a reflexive Banach space or a closed affine subset of a reflexive Banach space and let $\mathscr{F}\colon X \to \mathbb{R} \cup \{+\infty\}$. Assume the following:*

(WH1) ***Weak coercivity:*** *For all $\Lambda \in \mathbb{R}$ the sublevel set*

$$\left\{ u \in X \;:\; \mathscr{F}[u] \le \Lambda \right\} \qquad \text{is sequentially weakly precompact,}$$

that is, if $\mathscr{F}[u_j] \le \Lambda$ for a sequence $(u_j) \subset X$ and some $\Lambda \in \mathbb{R}$, then (u_j) has a weakly converging subsequence.

(WH2) ***Weak lower semicontinuity:*** *For all sequences $(u_j) \subset X$ with $u_j \rightharpoonup u$ in X (weak convergence) it holds that*

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

Then, the problem

$$\text{Minimize } \mathscr{F}[u] \text{ over all } u \in X$$

has at least one solution.

The proof of this theorem is analogous to the proof of Theorem 2.1, also taking into account the fact that all (strongly) closed affine subsets of a Banach space are weakly closed.

2.2 Functionals with Convex Integrands

As a first instance of the theory of integral functionals to be developed in this book, we now consider the minimization problem for

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

over all $u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$, where $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain and $p \in (1,\infty)$ will be chosen later (depending on growth properties of f). The reader is referred to Appendix A.5 for an overview of Sobolev spaces.

The following lemma shows that the integrand is measurable if f is a so-called *Carathéodory integrand*, which from now on we assume.

Lemma 2.4. *Let $f\colon \Omega \times \mathbb{R}^N \to \mathbb{R}$ be a* ***Carathéodory integrand****, that is,*

(i) $x \mapsto f(x, A)$ is Lebesgue-measurable for every fixed $A \in \mathbb{R}^N$;
(ii) $A \mapsto f(x, A)$ is continuous for (Lebesgue-)almost every fixed $x \in \Omega$.

Then, for any Borel-measurable map $V\colon \Omega \to \mathbb{R}^N$ the composition $x \mapsto f(x, V(x))$ is Lebesgue-measurable.

Proof. Assume first that V is a simple function,

$$V = \sum_{k=1}^{m} v_k \mathbb{1}_{E_k},$$

where the sets $E_k \subset \Omega$ are Borel-measurable ($k \in \{1, \dots, m\}$), $\bigcup_{k=1}^m E_k = \Omega$, and $v_k \in \mathbb{R}^N$. For $t \in \mathbb{R}$ we have

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

which is a Lebesgue-measurable set by assumption. Hence, $x \mapsto f(x, V(x))$ is Lebesgue-measurable.

Turning to the general case, every Borel-measurable function V can be approximated by simple functions V_k with

$$f(x, V_k(x)) \to f(x, V(x)) \quad \text{for all } x \in \Omega \qquad \text{as } k \to \infty,$$

see Lemma A.5. We conclude that the right-hand side is Lebesgue-measurable as the pointwise limit of Lebesgue-measurable functions. □

It is possible that the (compound) integrand in $\mathscr{F}$ is measurable, but that the integral is not well-defined. These pathological cases can, for example, be avoided if $f \geq 0$ or if one imposes the **p-growth bound**

$$|f(x, A)| \leq M(1 + |A|^p), \qquad (x, A) \in \Omega \times \mathbb{R}^{m \times d},$$

for some $M > 0$, which implies the finiteness of $\mathscr{F}[u]$ for all $u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$. In this chapter, however, this bound is not otherwise needed.

We next investigate the coercivity of $\mathscr{F}$. If $p \in (1, \infty)$, then the most basic assumption to guarantee coercivity, and the only one we consider here, is the **p-coercivity bound**

$$\mu |A|^p \leq f(x, A), \qquad (x, A) \in \Omega \times \mathbb{R}^{m \times d}, \tag{2.2}$$

for some $\mu > 0$. This coercivity also determines the exponent p for the Sobolev space where we look for solutions. Note that in the literature sometimes the coercivity bound is given as the seemingly more general $\mu |A|^p - C \leq f(x, A)$ for some $\mu, C > 0$. This, however, does not increase generality since we may pass from the integrand $f(x, A)$ to the integrand $\tilde{f}(x, A) := f(x, A) + C$, which now satisfies (2.2), without changing the minimization problem (recall that Ω is assumed bounded throughout this book).

Proposition 2.5. *If the Carathéodory integrand* $f : \Omega \times \mathbb{R}^{m\times d} \to [0,\infty)$ *satisfies the p-coercivity bound* (2.2) *with* $p \in (1,\infty)$*, then* $\mathscr{F}$ *is weakly coercive on the space*

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

where $g \in \mathrm{W}^{1-1/p,p}(\partial\Omega;\mathbb{R}^m)$.

Proof. We need to show that any sequence $(u_j) \subset \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^m)$ with

$$\sup_{j\in\mathbb{N}} \mathscr{F}[u_j] < \infty$$

is weakly precompact. From (2.2) we get

$$\mu \cdot \sup_{j\in\mathbb{N}} \int_\Omega |\nabla u_j|^p \,\mathrm{d}x \le \sup_{j\in\mathbb{N}} \mathscr{F}[u_j] < \infty,$$

whereby $\sup_j \|\nabla u_j\|_{\mathrm{L}^p} < \infty$. Fix $u_0 \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^m)$. Then, $u_j - u_0 \in \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^m)$ and $\sup_j \|\nabla(u_j - u_0)\|_{\mathrm{L}^p} < \infty$. From the Poincaré inequality, see Theorem A.26 (i), we therefore get

$$\sup_j \|u_j\|_{\mathrm{W}^{1,p}} \le \sup_j \|u_j - u_0\|_{\mathrm{W}^{1,p}} + \|u_0\|_{\mathrm{W}^{1,p}} < \infty.$$

This finishes the proof since bounded sets in separable and reflexive Banach spaces, like $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ for $p \in (1,\infty)$, are sequentially weakly precompact by Theorem A.2. □

Having settled the question of weak coercivity, we can now investigate the weak lower semicontinuity. The following pivotal result (in the one-dimensional case) goes back to the work of Leonida Tonelli in the early 20th century; the generalization to higher dimensions is due to James Serrin.

Theorem 2.6 (Tonelli 1920 & Serrin 1961 [242, 276]). *Let* $f : \Omega \times \mathbb{R}^{m\times d} \to [0,\infty)$ *be a Carathéodory integrand such that*

$$f(x, \cdot) \text{ is convex for almost every } x \in \Omega.$$

Then, $\mathscr{F}$ *is weakly lower semicontinuous on* $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ *for any* $p \in (1,\infty)$.

Proof. Step 1. We first establish that $\mathscr{F}$ is strongly lower semicontinuous, so let $u_j \to u$ in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ and $\nabla u_j \to \nabla u$ almost everywhere, which holds after selecting a subsequence (not explicitly labeled), see Appendix A.3. By assumption we have that $f(x, \nabla u_j(x)) \ge 0$. Applying Fatou's Lemma, we immediately conclude that

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

Since this holds for all subsequences, it also follows for our original sequence, see Problem 2.1.

Step 2. To prove the claimed weak lower semicontinuity take $(u_j) \subset \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ with $u_j \rightharpoonup u$ in $\mathrm{W}^{1,p}$. We need to show that

$$\mathscr{F}[u] \le \liminf_{j\to\infty} \mathscr{F}[u_j] =: \alpha. \tag{2.3}$$

Taking a subsequence (not explicitly labeled), we can in fact assume that $\mathscr{F}[u_j]$ converges to α.

By the Mazur Lemma A.4 we may find convex combinations

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

such that $v_j \to u$ in $\mathrm{W}^{1,p}$. As $f(x, \bullet)$ is convex for almost every x,

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

Since $\mathscr{F}[u_n] \to \alpha$ as $n \to \infty$ and $\sum_{n=j}^{N(j)} \theta_n^{(j)} = 1$, we arrive at

$$\liminf_{j\to\infty} \mathscr{F}[v_j] \le \alpha.$$

On the other hand, from the first step and since $v_j \to u$ strongly, we have $\mathscr{F}[u] \le \liminf_{j\to\infty} \mathscr{F}[v_j]$. Thus, (2.3) follows and the proof is finished. □

We can summarize our findings in the following existence theorem.

Theorem 2.7. *Let $f\colon \Omega \times \mathbb{R}^{m\times d} \to [0, \infty)$ be a Carathéodory integrand such that*

(i) f satisfies the p-coercivity bound (2.2) with $p \in (1, \infty)$;
(ii) $f(x, \bullet)$ is convex for almost every $x \in \Omega$.

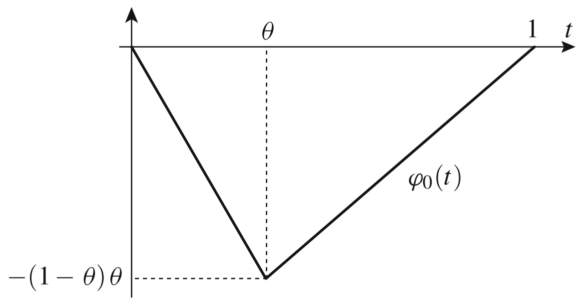
Then, the associated functional $\mathscr{F}$ has a minimizer over $\mathrm{W}_g^{1,p}(\Omega; \mathbb{R}^m)$, where $g \in \mathrm{W}^{1-1/p,p}(\partial\Omega; \mathbb{R}^m)$.

Proof. This follows immediately from the Direct Method for the weak convergence, Theorem 2.3 with $X := \mathrm{W}_g^{1,p}(\Omega; \mathbb{R}^m)$ together with Proposition 2.5 and the Tonelli–Serrin Theorem 2.6. □

Example 2.8. The **Dirichlet functional** (or **Dirichlet integral**) is

$$\mathscr{F}[u] := \int_\Omega \frac{1}{2} |\nabla u(x)|^2 \,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^m).$$

Fig. 2.1 The function φ_0



We already encountered this integral functional when considering electrostatics in Section 1.3. It is easy to see that the Dirichlet functional satisfies all requirements of Theorem 2.7 and so there exists a minimizer for any prescribed boundary values $g \in \mathrm{W}^{1/2,2}(\partial\Omega; \mathbb{R}^m)$.

We next show the following converse to the Tonelli–Serrin Theorem 2.6:

Proposition 2.9. *Let* $\mathscr{F}\colon \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m) \to \mathbb{R}$, $p \in [1, \infty)$, *be an integral functional with continuous integrand* $f\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ *(not x-dependent). If* $\mathscr{F}$ *is weakly lower semicontinuous on* $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ *and if either* $m = 1$ *or* $d = 1$ *(the* scalar case *and the* one-dimensional case*, respectively), then* f *is convex.*

Proof. We only consider the case $m = 1$ and d arbitrary; the other case is proved in a similar manner. Assume that $a, b \in \mathbb{R}^d$ with $a \neq b$ and $\theta \in (0, 1)$. Let $v := \theta a + (1-\theta)b$, $n := b - a$, and set

$$u_j(x) := v \cdot x + \frac{1}{j}\varphi_0\bigl(jx \cdot n - \lfloor jx \cdot n \rfloor\bigr), \qquad x \in \Omega,$$

where $\lfloor s \rfloor$ denotes the largest integer less than or equal to $s \in \mathbb{R}$, and

$$\varphi_0(t) := \begin{cases} -(1-\theta)t & \text{if } t \in [0, \theta), \\ \theta t - \theta & \text{if } t \in [\theta, 1), \end{cases}$$

see Figure 2.1. We have that

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

Hence, $(u_j) \subset \mathrm{W}^{1,\infty}(\Omega)$ and since the second term in the definition of u_j converges to zero uniformly, it holds that $u_j \rightharpoonup v \cdot x$ in $\mathrm{W}^{1,p}$ (here and in the following, "$v \cdot x$" is a shorthand notation for the linear function $x \mapsto v \cdot x$). By the weak lower semicontinuity, we conclude that

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

This proves the claim. □

In the **vectorial case**, i.e., $m \neq 1$ and $d \neq 1$, it turns out that convexity of the integrand (in the gradient variable) is far from being necessary for weak lower semicontinuity. In fact, there is indeed a weaker condition ensuring weak lower semicontinuity; we will explore this in Chapter 5.

Finally, we prove the following result concerning the *uniqueness* of the minimizer.

Proposition 2.10. *Let $\mathscr{F} \colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m) \to \mathbb{R}$, $p \in [1,\infty)$, be an integral functional with Carathéodory integrand $f \colon \Omega \times \mathbb{R}^{m\times d} \to \mathbb{R}$. If f is **strictly convex**, that is,*

$$f(x, \theta A + (1-\theta)B) < \theta f(x, A) + (1-\theta) f(x, B)$$

for all $x \in \Omega$, $A, B \in \mathbb{R}^{m\times d}$ with $A \neq B$, $\theta \in (0,1)$, then the minimizer $u_* \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^m)$ ($g \in \mathrm{W}^{1-1/p,p}(\partial\Omega;\mathbb{R}^m)$) of $\mathscr{F}$, if it exists, is unique.

Proof. Assume there are two different minimizers $u, v \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^m)$ of $\mathscr{F}$. Then set

$$w := \frac{1}{2}u + \frac{1}{2}v \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^m)$$

and observe that

$$\mathscr{F}[w] = \int_\Omega f\Big(x, \frac{1}{2}\nabla u(x) + \frac{1}{2}\nabla v(x)\Big) < \frac{1}{2}\mathscr{F}[u] + \frac{1}{2}\mathscr{F}[v] = \min_{\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^m)} \mathscr{F},$$

yielding an immediate contradiction. □

2.3 Integrands with *u*-Dependence

If we try to extend the results in the previous section to more general functionals

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

we discover that our proof strategy via the Mazur lemma runs into difficulties: We cannot "pull out" the convex combination inside

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

any more. Nevertheless, a lower semicontinuity result analogous to the one for the u-independent case turns out to be true:

Theorem 2.11. *Let $f \colon \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to [0,\infty)$ be a Carathéodory integrand, which here means that*

(i) $x \mapsto f(x,v,A)$ is Lebesgue-measurable for every fixed $(v,A) \in \mathbb{R}^m \times \mathbb{R}^{m\times d}$;
(ii) $(v,A) \mapsto f(x,v,A)$ is continuous for (Lebesgue-)almost every fixed $x \in \Omega$.

Assume also that

$$f(x,v,\bullet) \text{ is convex for every } (x,v) \in \Omega \times \mathbb{R}^m.$$

Then, for $p \in (1,\infty)$, the functional

$$\mathscr{F}[u] := \int_\Omega f(x,u(x),\nabla u(x))\,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m),$$

is weakly lower semicontinuous.

While it would be possible to give an elementary proof of this theorem here, we postpone the detailed study of integral functionals with u-dependent integrands until Section 5.6. There, using more advanced techniques, we will establish a much more general lower semicontinuity result, albeit under an additional p-growth assumption $|f(x,v,A)| \le M(1+|v|^p+|A|^p)$. A proof of the above theorem without this growth assumption can be found in Section 3.2.6 of [76].

Example 2.12. In the prototypical problem of linearized elasticity from Section 1.7 we are tasked to solve

$$\begin{cases} \text{Minimize} & \mathscr{F}[u] := \dfrac{1}{2}\displaystyle\int_\Omega 2\mu|\mathscr{E}u|^2 + \Big(\kappa - \frac{2}{3}\mu\Big)|\operatorname{tr}\mathscr{E}u|^2 - b\cdot u\,\mathrm{d}x \\ \text{over all} & u \in \mathrm{W}^{1,2}(\Omega;\mathbb{R}^3) \text{ with } u|_{\partial\Omega} = g, \end{cases}$$

where $\mu,\kappa > 0$, $b \in \mathrm{L}^2(\Omega;\mathbb{R}^3)$, and $g \in \mathrm{W}^{1/2,2}(\partial\Omega;\mathbb{R}^m)$. It is clear that $\mathscr{F}$ has quadratic growth. We assume that $\kappa - 2\mu/3 \ge 0$ and $g = 0$ for simplicity. Then, we first show that

$$\|\nabla u\|_{\mathrm{L}^2} \le \sqrt{2}\|\mathscr{E}u\|_{\mathrm{L}^2} \tag{2.4}$$

for all $u \in \mathrm{W}^{1,2}(\Omega;\mathbb{R}^3)$ with $u|_{\partial\Omega} = 0$. This can be seen as follows: An elementary computation shows that for $\varphi \in \mathrm{C}_c^\infty(\Omega;\mathbb{R}^3)$ it holds that

$$2(\mathscr{E}\varphi : \mathscr{E}\varphi) - \nabla\varphi : \nabla\varphi = \operatorname{div}\big[(\nabla\varphi)\varphi - (\operatorname{div}\varphi)\varphi\big] + (\operatorname{div}\varphi)^2.$$

Thus, by the divergence theorem,

$$\begin{aligned}2\|\mathscr{E}\varphi\|_{L^2}^2 - \|\nabla\varphi\|_{L^2}^2 &= \int_\Omega \operatorname{div}\left[(\nabla\varphi)\varphi - (\operatorname{div}\varphi)\varphi\right] dx + \int_\Omega (\operatorname{div}\varphi)^2 \, dx \\ &= \int_\Omega (\operatorname{div}\varphi)^2 \, dx \\ &\geq 0.\end{aligned}$$

This is (2.4) for φ. The general case follows from the density of $C_c^\infty(\Omega; \mathbb{R}^3)$ in $W_0^{1,2}(\Omega; \mathbb{R}^3)$. Then, using Young's inequality and the Poincaré inequality (see Theorem A.26 (i), we denote the L^2-Poincaré constant by $C_P > 0$), we get for any $\delta > 0$,

$$\begin{aligned}\mathscr{F}[u] &\geq \mu\|\mathscr{E}u\|_{L^2}^2 - \|b\|_{L^2}\|u\|_{L^2} \\ &\geq \mu\|\mathscr{E}u\|_{L^2}^2 - \frac{1}{2\delta}\|b\|_{L^2}^2 - \frac{\delta}{2}\|u\|_{L^2}^2 \\ &\geq \frac{\mu}{2}\|\nabla u\|_{L^2}^2 - \frac{1}{2\delta}\|b\|_{L^2}^2 - \frac{C_P^2\delta}{2}\|\nabla u\|_{L^2}^2.\end{aligned}$$

Choosing $\delta = \mu/(2C_P^2)$, we obtain the coercivity estimate

$$\mathscr{F}[u] \geq \frac{\mu}{4}\|\nabla u\|_{L^2}^2 - \frac{C_P^2}{\mu}\|b\|_{L^2}^2.$$

Hence, applying the Poincaré inequality again, $\mathscr{F}[u]$ controls $\|u\|_{W^{1,2}}$ and our functional is weakly coercive. Moreover, it is clear that the integrand is convex in the $\mathscr{E}u$-argument. Hence, Theorem 2.11 yields the existence of a solution $u_* \in W^{1,2}(\Omega; \mathbb{R}^3)$ to our minimization problem of linearized elasticity. In fact, one could also argue using the Tonelli–Serrin Theorem 2.6 and the elementary fact that the lower-order term $\int_\Omega b(x) \cdot u(x) \, dx$ is weakly continuous on $W^{1,2}$. More on the topic of linearized elasticity can be found in Sections 6.2 and 6.3 of [64].

2.4 The Lavrentiev Gap Phenomenon

We have chosen the function space in which we look for the solution of a minimization problem from the scale of Sobolev spaces according to a coercivity assumption such as (2.2). However, at first sight, classically differentiable functions may appear to be more appealing. So the question arises whether the infimum value is actually the same when considering different function spaces. Formally, given two linear or affine spaces $X \subset Y$ such that X is dense in Y, and a functional $\mathscr{F} : Y \to \mathbb{R} \cup \{+\infty\}$, we ask whether

$$\inf_X \mathscr{F} = \inf_Y \mathscr{F}.$$

Note that even if the infima agree, it is a priori unlikely that this infimum is *attained* in both spaces unless we have additional regularity of a minimizer (which we will investigate in Section 3.2).

For $X = \mathrm{C}^\infty$ and $Y = \mathrm{W}^{1,p}$ the equality of infima turns out to be true under suitable growth conditions:

Theorem 2.13. *Let* $f\colon \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}$ *be a Carathéodory integrand with* p*-growth, i.e.,*

$$|f(x,v,A)| \le M(1+|v|^p+|A|^p), \qquad (x,v,A) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times d},$$

for some $M > 0$, $p \in [1,\infty)$*. Then, the functional*

$$\mathscr{F}[u] := \int_\Omega f(x,u(x),\nabla u(x))\,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m),$$

is strongly continuous. Consequently,

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

The same equality of infima also holds with fixed boundary values.

Proof. Let $u_j \to u$ in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ and additionally assume that $u_j \to u$, $\nabla u_j \to \nabla u$ almost everywhere (which holds after selecting a subsequence). Then, from the p-growth assumption we get

$$\mathscr{F}[u_j] = \int_\Omega f(x,u_j,\nabla u_j)\,\mathrm{d}x \le \int_\Omega M(1+|u_j|^p+|\nabla u_j|^p)\,\mathrm{d}x$$

and via Pratt's Theorem A.10 we infer that

$$\mathscr{F}[u_j] \to \mathscr{F}[u].$$

Since this holds for a subsequence of any subsequence of the original sequence (u_j), we have established the continuity of $\mathscr{F}$ with respect to the strong convergence in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$.

The assertion about the equality of infima now follows readily since $\mathrm{C}^\infty(\Omega;\mathbb{R}^m)$ is dense in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$. The equality of the infima under an additional boundary value constraint follows from the continuity of the trace operator under the $\mathrm{W}^{1,p}$-convergence, see Theorem A.24, and the fact that any map in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ can be approximated with smooth functions with the same boundary values, see Theorem A.29. □

If we dispense with the p-growth assumption, however, the infimum over different spaces may indeed be different – this is called the *Lavrentiev gap phenomenon*, discovered in 1926 by Mikhail Lavrentiev. Here, we give an example between the spaces $\mathrm{W}^{1,1}$ and $\mathrm{W}^{1,\infty}$ (with boundary conditions):

Example 2.14 (*Manià 1934 [178]*). Consider the minimization problem

$$\begin{cases} \text{Minimize} & \mathscr{F}[u] := \displaystyle\int_0^1 (u(t)^3 - t)^2 \dot{u}(t)^6 \,\mathrm{d}t \\ \text{subject to} & u(0) = 0,\ u(1) = 1 \end{cases}$$

for u from either $\mathrm{W}^{1,1}(0,1)$ or $\mathrm{W}^{1,\infty}(0,1)$. We claim that

$$\inf_{\mathrm{W}^{1,1}(0,1)} \mathscr{F} < \inf_{\mathrm{W}^{1,\infty}(0,1)} \mathscr{F},$$

where here and in the following these infima are to be taken only over functions u with boundary values $u(0) = 0$, $u(1) = 1$.

Clearly, $\mathscr{F} \geq 0$, and for $u_*(t) := t^{1/3} \in (\mathrm{W}^{1,1} \setminus \mathrm{W}^{1,\infty})(0,1)$ we have $\mathscr{F}[u_*] = 0$. Thus,

$$\inf_{\mathrm{W}^{1,1}(0,1)} \mathscr{F} = 0.$$

On the other hand, every $u \in \mathrm{W}^{1,\infty}(0,1)$ is Lipschitz continuous. Thus, also using $u(0) = 0$, $u(1) = 1$, there exists a $\tau \in (0,1)$ with

$$u(t) \leq h(t) := \frac{t^{1/3}}{2} \quad \text{for all } t \in [0,\tau] \qquad \text{and} \qquad u(\tau) = h(\tau).$$

Then, $u(t)^3 - t \leq h(t)^3 - t$ for $t \in [0,\tau]$ and, since both of these terms are negative,

$$(u(t)^3 - t)^2 \geq (h(t)^3 - t)^2 = \frac{7^2}{8^2} t^2 \qquad \text{for all } t \in [0,\tau].$$

We then estimate

$$\mathscr{F}[u] \geq \int_0^\tau (u(t)^3 - t)^2 \, \dot{u}(t)^6 \,\mathrm{d}t \geq \frac{7^2}{8^2} \int_0^\tau t^2 \, \dot{u}(t)^6 \,\mathrm{d}t.$$

Further, by Hölder's inequality,

$$\begin{aligned} \int_0^\tau \dot{u}(t)\,\mathrm{d}t &= \int_0^\tau t^{-1/3} \cdot t^{1/3}\, \dot{u}(t)\,\mathrm{d}t \\ &\leq \left(\int_0^\tau t^{-2/5}\,\mathrm{d}t \right)^{5/6} \cdot \left(\int_0^\tau t^2\, \dot{u}(t)^6\,\mathrm{d}t \right)^{1/6} \\ &= \frac{5^{5/6}}{3^{5/6}} \tau^{1/2} \left(\int_0^\tau t^2\, \dot{u}(t)^6\,\mathrm{d}t \right)^{1/6}. \end{aligned}$$

Since also

$$\int_0^\tau \dot{u}(t)\,\mathrm{d}t = u(\tau) - u(0) = h(\tau) = \frac{\tau^{1/3}}{2},$$

we arrive at

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

Thus,

$$\inf_{\mathrm{W}^{1,\infty}(0,1)} \mathscr{F} > \inf_{\mathrm{W}^{1,1}(0,1)} \mathscr{F},$$

and $\mathscr{F}$ can be seen to exhibit the Lavrentiev gap phenomenon.

In a more recent example, Ball & Mizel [34] showed that the problem

$$\begin{cases} \text{Minimize} \quad \mathscr{F}[u] := \displaystyle\int_{-1}^{1} (t^4 - u(t)^6)^2 |\dot{u}(t)|^{2m} + \varepsilon \dot{u}(t)^2 \,\mathrm{d}t \\ \text{subject to} \quad u(-1) = \alpha, \ u(1) = \beta \end{cases}$$

also exhibits the Lavrentiev gap phenomenon between the spaces $\mathrm{W}^{1,2}$ and $\mathrm{W}^{1,\infty}$ if $m \in \mathbb{N}$ satisfies $m > 13$, $\varepsilon > 0$ is sufficiently small, and $-1 \leq \alpha < 0 < \beta \leq 1$. This example is significant because the Ball–Mizel functional is *coercive* on $\mathrm{W}^{1,2}(-1, 1)$ thanks to the second term of the integrand.

We note that the Lavrentiev gap phenomenon is a major obstacle for the numerical approximation of minimization problems. For instance, standard (piecewise affine) finite element approximations are in $\mathrm{W}^{1,\infty}$ and hence in the presence of the Lavrentiev gap phenomenon (between $\mathrm{W}^{1,p}$ and $\mathrm{W}^{1,\infty}$) we cannot approximate the true solution with such finite elements. Thus, one is forced to work with non-conforming elements and other advanced schemes. This issue does not only affect "academic" examples such as the ones above, but is also of great concern in applied problems, such as nonlinear elasticity theory.

2.5 Integral Side Constraints

In some minimization problems the class of candidate functions is restricted to include one or more integral side constraints. To establish the existence of a minimizer in these cases, we first need to extend the Direct Method to this scenario.

Theorem 2.15. *Let X be a Banach space or a closed affine subset of a Banach space and let $\mathscr{F}, \mathscr{H} \colon X \to \mathbb{R} \cup \{+\infty\}$. Assume the following:*

(WH1) ***Weak coercivity of $\mathscr{F}$****: For all $\Lambda \in \mathbb{R}$ the sublevel set*

$$\{ u \in X \,:\, \mathscr{F}[u] \leq \Lambda \} \qquad \textit{is sequentially weakly precompact,}$$

that is, if $\mathscr{F}[u_j] \le \Lambda$ for a sequence $(u_j) \subset X$ and some $\Lambda \in \mathbb{R}$, then (u_j) has a weakly converging subsequence.

(WH2) ***Weak lower semicontinuity of $\mathscr{F}$:*** *For all sequences $(u_j) \subset X$ with $u_j \rightharpoonup u$ in X it holds that*

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

(WH3) ***Weak continuity of $\mathscr{H}$:*** *For all sequences $(u_j) \subset X$ with $u_j \rightharpoonup u$ in X it holds that*

$$\mathscr{H}[u_j] \to \mathscr{H}[u].$$

Assume also that there exists at least one $u_0 \in X$ with $\mathscr{H}[u_0] = 0$. Then, the minimization problem

$$\text{Minimize } \mathscr{F}[u] \text{ over all } u \in X \text{ with } \mathscr{H}[u] = 0$$

has a solution.

Proof. The proof is almost exactly the same as the one for the standard Direct Method in Theorem 2.3. The only difference is that we need to select the u_j for a minimizing sequence with $\mathscr{H}[u_j] = 0$. Then, by (WH3), this property also holds for any weak limit u_* of a subsequence of the u_j's, which then is the sought minimizer. □

A large class of side constraints can be treated using the following simple result.

Lemma 2.16. *Let $h\colon \Omega \times \mathbb{R}^m \to \mathbb{R}$ be a Carathéodory integrand and let $p \in [1, \infty)$ such that there exists an $M > 0$ with*

$$|h(x, v)| \le M(1 + |v|^q), \qquad (x, v) \in \Omega \times \mathbb{R}^m, \tag{2.5}$$

for some $q \in [1, dp/(d-p))$ if $p \le d$, or no growth condition if $p > d$. Then, the functional $\mathscr{H}\colon \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m) \to \mathbb{R}$ defined through

$$\mathscr{H}[u] := \int_\Omega h(x, u(x))\,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m),$$

is weakly continuous.

Proof. We only prove the lemma in the case $p \le d$. The proof for $p > d$ is analogous, but easier.

Let $u_j \rightharpoonup u$ in $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$, whereby after selecting a subsequence and employing the Rellich–Kondrachov Theorem A.28 and Lemma A.8, $u_j \to u$ in L^q and almost everywhere. By assumption we have

$$\pm h(x, v) + M(1 + |v|^q) \ge 0.$$

Thus, applying Fatou's lemma separately to these two integrands, we get

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

Since $\|u_j\|_{\mathrm{L}^q}\to\|u\|_{\mathrm{L}^q}$, we can combine these two assertions to get $\mathscr{H}[u_j]\to\mathscr{H}[u]$. This holds for a subsequence of any subsequence of (u_j), hence it also holds for our original sequence. □

Combining this lemma with Theorems 2.7 and 2.15 and also the Rellich–Kondrachov Theorem A.28, we immediately get the following existence result.

Theorem 2.17. *Let $f\colon \Omega\times\mathbb{R}^{m\times d}\to[0,\infty)$ and $h\colon \Omega\times\mathbb{R}^m\to\mathbb{R}$ be Carathéodory integrands such that*

(i) f satisfies the p-coercivity bound (2.2), where $p\in(1,\infty)$;
(ii) $f(x,\cdot)$ is convex for all $x\in\Omega$;
(iii) h satisfies the q-growth condition (2.5) for some $q\in[1,dp/(d-p))$ if $p\le d$, or no growth condition if $p>d$.

Then, there exists a minimizer $u_*\in\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^m)$, where $g\in\mathrm{W}^{1-1/p,p}(\partial\Omega;\mathbb{R}^m)$, of the functional

$$\mathscr{F}[u]:=\int_\Omega f(x,u(x),\nabla u(x))\,\mathrm{d}x,\qquad u\in\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^m),$$

under the side constraint

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

2.6 The General Theory of Convex Functions and Duality

We finish this chapter by briefly considering the general theory of convex functions.

In all of the following let X be a (real) reflexive Banach space (finite or infinite-dimensional) with dual space X^*, see Appendix A.2. We denote by $\langle x,x^*\rangle=x^*(x)$ the duality product between $x\in X$ and $x^*\in X^*$. For a set $A\subset X$ we write $\operatorname{co}A$, $\overline{\operatorname{co}}\,A$ for its **convex hull** and **closed convex hull**, respectively. These hulls are defined to be the smallest (closed) convex set containing A, or, equivalently, the intersection of all (closed) convex sets containing A. For $A\subset X$ we furthermore define the **characteristic function** $\chi_A\colon X\to\mathbb{R}\cup\{+\infty\}$ as

$$\chi_A(x):=\frac{1}{\mathbb{1}_A(x)}-1=\begin{cases}0 & \text{if } x\in A,\\ +\infty & \text{if } x\notin A.\end{cases}$$

Let $F\colon X\to\mathbb{R}\cup\{+\infty\}$. The function F is called **proper** if it is not identically $+\infty$. We define the **effective domain** $\operatorname{dom}\,F\subset X$ and the **epigraph** $\operatorname{epi}\,F\subset X\times\mathbb{R}$

of F as follows:

$$\operatorname{dom} F := \{ x \in X : F(x) < +\infty \},$$
$$\operatorname{epi} F := \{ (x, \alpha) \in X \times \mathbb{R} : \alpha \geq F(x) \}.$$

It can be shown (see Problems 2.6, 2.7) that F is convex if and only if epi F is convex (as a set), and that f is (sequentially) lower semicontinuous if and only if epi F is (sequentially) closed; this holds with respect to both the strong and the weak convergence.

Lemma 2.18. *If* $\dim X < \infty$, *then every convex function* $F : X \to \mathbb{R} \cup \{+\infty\}$ *is locally bounded on the interior of its effective domain.*

Proof. If $x \in X$ is in the interior of the effective domain of F, then x lies in the convex hull co $\{x_1, \ldots, x_{n+1}\}$ of $n+1$ affinely independent points x_k (i.e., $\sum \alpha_k x_k = 0$ for some $\alpha_k \in \mathbb{R}$ with $\sum_k \alpha_k = 0$ implies $\alpha_1 = \alpha_2 = \cdots = \alpha_{n+1} = 0$) with $F(x_k) < +\infty$, where $n = \dim X$. Thus, there exists an open ball around x inside co $\{x_1, \ldots, x_{n+1}\}$ on which F is bounded by $\sup \{F(x_1), \ldots, F(x_{n+1})\}$. □

Lemma 2.19. *Let* $\mathscr{A}$ *be a non-empty family of continuous affine functions* $a(x) = \langle x, x^* \rangle + \alpha$ *for some* $x^* \in X^*$, $\alpha \in \mathbb{R}$. *Then,* $F : X \to \mathbb{R} \cup \{+\infty\}$ *defined through*

$$F(x) := \sup_{a \in \mathscr{A}} a(x)$$

is convex and lower semicontinuous. Conversely, every convex and lower semicontinuous function can be written in this form.

Proof. The convexity of F is clear since all the affine functions $a \in \mathscr{A}$ are in particular convex. For the lower semicontinuity we just need to realize that the pointwise supremum of continuous functions is always lower semicontinuous. Indeed, for a sequence $x_j \to x$ in X we have for all $\tilde{a} \in \mathscr{A}$ that

$$\tilde{a}(x) = \lim_{j \to \infty} \tilde{a}(x_j) \leq \liminf_{j \to \infty} \sup_{a \in \mathscr{A}} a(x_j) = \liminf_{j \to \infty} F(x_j).$$

Taking the supremum over all $\tilde{a} \in \mathscr{A}$, the lower semicontinuity follows.

For the converse, we may assume that F is proper; otherwise the result is trivial. Let $x \in X$ with $F(x) < +\infty$. The epigraph epi F of F is closed and convex by assumption. Hence, by the Hahn–Banach Separation Theorem A.1, for every $x \in X$ and every $\beta < F(x)$ we can find an affine function $a_{x,\beta} : X \to \mathbb{R}$ whose graph separates the point (x, β) from epi F. In particular, $\beta < a_{x,\beta}(x) < F(x)$ and $a_{x,\beta}$ lies everywhere below the graph of F. Letting $\beta \uparrow F(x)$, we arrive at

$$F(x) = \sup \{ a_{x,\beta}(x) : (x, \beta) \in X \times \mathbb{R} \text{ with } \beta < F(x) \}.$$

A similar argument also applies if $F(x) = +\infty$. Collecting all these $a_{x,\beta}$ for $(x, \beta) \in X \times \mathbb{R}$ with $\beta < F(x)$ into the set $\mathscr{A}$, the conclusion follows. □

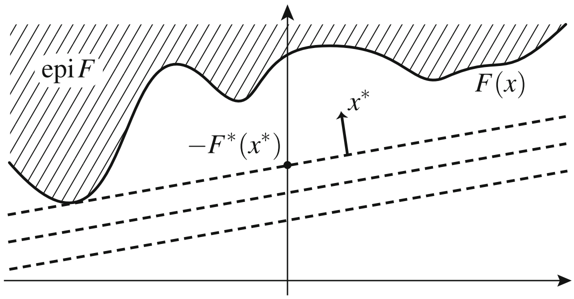


Fig. 2.2 The convex conjugate

Proposition 2.20. *Every proper convex function is continuous on the interior of its effective domain.*

We will prove this in more generality later, see Lemma 5.6 in conjunction with Lemma 2.18.

One important object in the general theory of convex functions is the **(convex) conjugate**, or **Legendre–Fenchel transform**, $F^*: X^* \to \mathbb{R} \cup \{+\infty\}$ of a proper function $F: X \to \mathbb{R} \cup \{+\infty\}$ (not necessarily convex), which is defined as follows:

$$F^*(x^*) := \sup_{x \in X} \left[\langle x, x^* \rangle - F(x)\right], \qquad x^* \in X^*.$$

Of course, we may restrict to $x \in \operatorname{dom} F$ in the supremum. The intuition here is that for a given x^* we may consider all affine hyperplanes with normal x^* (recall that all hyperplane normals are elements of X^*) that lie below epi F. Then, $-F^*(x^*)$ is the supremum of the heights at which these hyperplanes intersect the (vertical) $(\mathbb{R} \cup \{+\infty\})$-axis, see Figure 2.2. Indeed, let $\alpha \in \mathbb{R}$ be such that $F(x) \geq \langle x, x^* \rangle - \alpha$ for all $x \in X$. Then, $\alpha \geq \langle x, x^* \rangle - F(x)$ for all $x \in X$, so the highest supporting hyperplane with normal x^* is $x \mapsto \langle x, x^* \rangle - F^*(x^*)$, which intersects the vertical axis in $-F^*(x^*)$.

The following **Fenchel inequality** is immediate from the definition:

$$\langle x, x^* \rangle \leq F(x) + F^*(x^*), \qquad \text{for all } x \in X, x^* \in X^*. \tag{2.6}$$

We next collect some properties of the conjugate function:

Proposition 2.21. *Let* $F, G: X \to \mathbb{R} \cup \{+\infty\}$ *be proper and* $F^*, G^*: X^* \to \mathbb{R} \cup \{+\infty\}$ *be their conjugates.*

(i) F^* *is convex and lower semicontinuous.*
(ii) $F^*(0) = -\inf F$.
(iii) *If* $F \leq G$, *then* $G^* \leq F^*$.
(iv) *If for* $\lambda > 0$ *we denote by* F_λ *the scaled function* $F_\lambda(x) := F(\lambda x)$, *then* $F_\lambda^*(x^*) = F^*(x^*/\lambda)$.
(v) $(\lambda F)^*(x^*) = \lambda F^*(x^*/\lambda)$ *for all* $\lambda > 0$.

(vi) $(F+\gamma)^* = F^* - \gamma$ *for all* $\gamma \in \mathbb{R}$.

(vii) *If for* $a \in X$ *we denote by* F_a *the translated function* $F_a(x) := F(x-a)$*, then* $F_a^*(x^*) = F^*(x^*) + \langle a, x^* \rangle$.

Proof. The first assertion follows from Lemma 2.19, all the others are straightforward calculations, see Problem 2.8. □

We now consider a few canonical examples of convex functions.

Example 2.22 (*Support function*). Let χ_A be the characteristic function of $A \subset X$. Then, for the conjugate function we get

$$\sigma_A(x^*) := \chi_A^*(x^*) = \sup_{x \in A} \langle x, x^* \rangle, \qquad x^* \in X^*,$$

which is called the **support function** of A. It is always convex, lower semicontinuous, and **positively 1-homogeneous**, i.e., $\sigma_A(\alpha x^*) = \alpha \sigma_A(x^*)$ for all $x^* \in X^*$ and $\alpha \geq 0$, see Problem 2.9.

Example 2.23. Let $p, q \in (1, \infty)$ with $1/p + 1/q = 1$, that is, p, q are **conjugate exponents**. Then,

$$\varphi(t) := \frac{1}{p}|t|^p \qquad \text{and} \qquad \varphi^*(t) := \frac{1}{q}|t|^q, \qquad t \in \mathbb{R},$$

are conjugate. From the Fenchel inequality (2.6) we recover the **Young inequality**

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q} \qquad \text{for all } x, y \geq 0.$$

Example 2.24. For the absolute value function $\varphi(t) := |t|$ we get

$$\varphi^*(t) = \chi_{[-1,1]}(t) = \begin{cases} 0 & \text{if } |t| \leq 1, \\ +\infty & \text{if } |t| > 1, \end{cases} \qquad t \in \mathbb{R}. \tag{2.7}$$

Example 2.25. The conjugate of the exponential function is

$$\exp^*(t) = \begin{cases} +\infty & \text{if } t < 0, \\ 0 & \text{if } t = 0, \\ t \ln t - t & \text{if } t > 0, \end{cases} \qquad t \in \mathbb{R}.$$

In this case, (2.6) gives the inequality

$$xy \leq \exp(x) + y \ln y - y \qquad \text{for all } x, y > 0.$$

Example 2.26. Let $\varphi\colon \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be proper, convex, and lower semicontinuous, and let $\|\bullet\|$, $\|\bullet\|_*$ be the norms on X and on X^*, respectively. Then the functions

$$G(x) := \varphi(\|x\|) \qquad \text{and} \qquad G^*(x^*) := \varphi^*(\|x^*\|_*), \qquad x \in X,\ x^* \in X^*,$$

are conjugate. In particular, $\|\bullet\|^p/p$ and $\|\bullet\|^q/q$ for $1/p + 1/q = 1$ are conjugate. The verification of these statements is the task of Problem 2.10.

Example 2.27. Let $X = \mathbb{R}^n$ and let $S \in \mathbb{R}^{n\times n}$ be a symmetric, positive definite matrix. Then,

$$F(x) := \frac{1}{2}x^T Sx \qquad \text{and} \qquad F^*(y) := \frac{1}{2}y^T S^{-1} y, \qquad x, y \in \mathbb{R}^n,$$

are conjugate.

Iterating the construction of the conjugate, we denote by $F^{**}\colon X \to \mathbb{R} \cup \{+\infty\}$ the **biconjugate** of F, that is, the function

$$F^{**}(x) := \sup_{x^* \in X^*} \big[\langle x, x^*\rangle - F^*(x^*)\big], \qquad x \in X.$$

Proposition 2.28. *The biconjugate F^{**} is the* ***convex, lower semicontinuous envelope*** *of F, that is, the greatest convex, lower semicontinuous function below F. Moreover, $F^{***} = F^*$.*

Proof. For the moment denote the convex lower semicontinuous envelope of F by F_{clsc},

$$F_{\mathrm{clsc}}(x) := \sup\big\{\, H(x) \,:\, H \le F \text{ convex, lower semicontinuous} \,\big\}, \qquad x \in X.$$

Also define

$$G(x) := \sup\big\{\, a(x) \,:\, a \le F \text{ affine} \,\big\}, \qquad x \in X.$$

Since $G \le F$ is convex and lower semicontinuous by Lemma 2.19, $G \le F_{\mathrm{clsc}}$. On the other hand, for every convex and lower semicontinuous H from the definition of F_{clsc}, we have $H(x) = \sup_{b\in\mathscr{A}} b(x)$ for a collection of affine functions $b \le H$, again by the said lemma. However, $b \le F$ for all $b \in \mathscr{A}$ and thus b is included in the collection in the definition of G. Hence, $H \le G$, whereby $F_{\mathrm{clsc}} \le G$. In conclusion, $F_{\mathrm{clsc}} = G$.

Every affine $a \le F$ has the form $a(x) = \langle x, x^*\rangle - \alpha$ for some $x^* \in X^*$ and $\alpha \in \mathbb{R}$. We can restrict ourselves to such a with α minimal while still preserving the property $a \le F$. We see first that $a \le F$ if and only if $\alpha \ge \langle y, x^*\rangle - F(y)$ for all $y \in X$. According to the definition of the conjugate function, this condition is nothing else than

$$\alpha \ge F^*(x^*).$$

Thus, α is minimal when $\alpha = F^*(x^*)$ and we get

$$F_{\mathrm{clsc}}(x) = G(x) = \sup_{x^* \in X^*} \big[\langle x, x^* \rangle - F^*(x^*)\big] = F^{**}(x), \qquad x \in X.$$

For the second assertion it suffices to observe that F^* is convex and lower semicontinuous by Proposition 2.21 (i) and to apply the first assertion. □

As a particular consequence of the preceding result, we see that conjugation facilitates a bijection between the proper, convex, and lower semicontinuous functions on X and those on X^*, which is self-inverse in the sense above.

Corollary 2.29. $\operatorname{epi} F^{**} = \overline{\operatorname{co}} \operatorname{epi} F$.

Proof. The process of taking the convex lower semicontinuous envelope of F amounts to finding the closed convex hull of the epigraph. □

Example 2.30. For the characteristic function χ_A of $A \subset X$ we get

$$\chi_A^{**} = \sigma_A^* = \chi_{\overline{\operatorname{co}} A}.$$

In particular, A and $\overline{\operatorname{co}} A$ have the same support function.

Notes and Historical Remarks

The basic ideas concerning the Direct Method as well as lower semicontinuity and its connection to convexity are due to Leonida Tonelli and were established in a series of articles in the early 20th century [275–277]. In the 1960s James Serrin generalized the results to higher dimensions [242].

Most of the material in this chapter is very classical and can be found in a variety of books on the calculus of variations, we refer in particular to [76, 77, 137]. We note that a very general lower semicontinuity theorem for convex integrands can be found in Theorem 3.23 of [76].

All of our abstract results on the Direct Method are formulated using sequences and not using general topology tools like nets. This is justified since the weak topology on a separable, reflexive Banach space and the weak*-topology on a dual space with a separable predual are metrizable on norm-bounded sets. Thus, if the functionals under investigation satisfy suitable coerciveness assumptions, one can work with sequences. The only case where one has to be careful is when one uses the weak topology on a non-reflexive Banach space with a non-separable dual space because then the weak topology might not be metrizable. For instance, in the sequence space l^1 (with non-separable dual space l^∞), weak convergence of *sequences* is equivalent to strong convergence, but the weak and strong *topologies* still differ (see Chapter V in [74] for more details on such considerations). For us more relevant is the observation that norm-bounded sets in $\mathrm{L}^1(\Omega)$ are not weakly precompact, either sequentially or topologically (these notions turn out to be equivalent by the Eberlein–Šmulian theorem). This corresponds to functionals with linear growth, which indeed require a more involved analysis in the space of functions of bounded variation (BV). We will come back to this topic in Chapters 10–12.

For the u-dependent variational integrals the growth in the u-variable can be improved up to q-growth, where $q \in [1, p/(p-d))$ by the Sobolev embedding theorem. Moreover, we can work with the more general growth bounds $|f(x, v, A)| \leq M(h(x) + |v|^q + |A|^p)$, with $h \in \mathrm{L}^1(\Omega; [0, \infty))$ and $q \in [1, p/(p-d))$. For reasons of simplicity, we have omitted these generalizations here.

The Lavrentiev gap phenomenon was discovered in [175], our Example 2.14 is due to Manià; we follow the description in [117]. Tonelli's Regularity Theorem [118, 275] gives regularity and hence the absence of the Lavrentiev gap phenomenon, for some integral functionals with superlinear growth; also see [49, 140–143] for some recent developments in this direction.

Much of the theory of general convex functions was developed by Jean-Jacques Moreau and R. Tyrrell Rockafellar in the 1960s. The books [106, 232] and the more advanced monographs [192, 193, 233] develop these topics in great detail.

Problems

2.1. Let $\mathscr{F} \colon X \to \mathbb{R}$, where X is a complete metric space. Show that if every subsequence of the sequence $(u_j) \subset X$ with $u_j \to u$ in X has a further subsequence $(u_{j(k)})_k$ such that

$$\mathscr{F}[u] \leq \liminf_{k \to \infty} \mathscr{F}[u_{j(k)}],$$

then also

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

2.2. Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. Define

$$V := \left\{ u \in \mathrm{W}^{1,2}(\Omega) \;:\; \int_\Omega u(x) \,\mathrm{d}x = 0 \right\}.$$

Assume furthermore that $f \colon \Omega \times \mathbb{R}^d \to \mathbb{R}$ is continuously differentiable with

$$\begin{aligned} \mu|A|^2 &\leq f(x, A) && \text{for some } \mu > 0 \text{ and all } (x, A) \in \Omega \times \mathbb{R}^d, \\ |\mathrm{D}_A f(x, A)| &\leq M(1 + |A|^2) && \text{for some } M > 0 \text{ and all } (x, A) \in \Omega \times \mathbb{R}^d, \end{aligned}$$

and that $A \mapsto f(x, A)$ is convex for all $x \in \Omega$. Finally, let $g \in \mathrm{L}^2(\Omega)$. Consider the following minimization problem:

$$\begin{cases} \text{Minimize} & \mathscr{F}[u] := \int_{\Omega} f(x, \nabla u(x)) - g(x)u(x) \, \mathrm{d}x \\ \text{over all} & u \in V. \end{cases}$$

(i) Show that $\mathscr{F}$ is coercive on V, that is, there exists a $\mu > 0$ such that

$$\mathscr{F}[u] \geq \mu \|u\|^2_{\mathrm{W}^{1,2}} - \mu^{-1} \qquad \text{for all } u \in V.$$

(ii) Show that $\mathscr{F}$ is also weakly lower semicontinuous on V (weak convergence in $\mathrm{W}^{1,2}$) and hence there exists a minimizer $u_* \in V$ of $\mathscr{F}$ (minimized over V).

This problem is continued in Problem 3.9 in the next chapter.

2.3. Show that the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = xy$ is **separately convex**, that is, $x \mapsto f(x, y)$ is convex for fixed $y \in \mathbb{R}$ and $y \mapsto f(x, y)$ is convex for fixed $x \in \mathbb{R}$, but f is not convex.

2.4. Let $f : \mathbb{R}^d \to [0, \infty)$ be twice continuously differentiable and assume that there are constants $\mu, M > 0$ with

$$\mu|b|^2 \leq \mathrm{D}^2 f(a)[b, b] \leq M|b|^2 \qquad \text{for all } a, b \in \mathbb{R}^d,$$

where

$$\mathrm{D}^2 f(a)[b, b] := \frac{\mathrm{d}^2}{\mathrm{d}t^2} f(a + tb)\bigg|_{t=0} \qquad \text{for all } a, b \in \mathbb{R}^d.$$

Show that f is convex and that $|f(v)| \leq C(1 + |v|^2)$ for some $C > 0$ and all $v \in \mathbb{R}^d$.

2.5. Let $f : \mathbb{R}^d \to \mathbb{R}$ be convex and fix $x_0 \in \mathbb{R}^d$. Set

$$M := \max_{i=1,\dots,d} \big(|f(x_0 + \mathrm{e}_i) - f(x_0)|, |f(x_0 - \mathrm{e}_i) - f(x_0)| \big).$$

Prove that if $y \in \mathbb{R}^d$ satisfies $|y|_1 := |y_1| + \cdots + |y_d| \leq 1$, then $f(x_0 + y) - f(x_0) \leq M$.

2.6. Show that $F : X \to \mathbb{R} \cup \{+\infty\}$ is convex if and only if epi F is convex (as a set).

2.7. Show that $F : X \to \mathbb{R} \cup \{+\infty\}$ is (sequentially) lower semicontinuous if and only if epi F is (sequentially) closed.

2.8. Prove the statements of Proposition 2.21.

2.9. Verify the statements in Example 2.22 about the support function.

2.10. Prove the assertion in Example 2.26.

Chapter 3
Variations

In this chapter we discuss *variations* of functionals. The idea is the following: Let $\mathscr{F}: \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^m) \to \mathbb{R}$ be a functional with minimizer $u_* \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^m)$. Take a path $t \mapsto u_t \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^m)$ ($t \in \mathbb{R}$) with $u_0 = u_*$ and consider the behavior of the map

$$t \mapsto \mathscr{F}[u_t]$$

around $t = 0$. If $t \mapsto \mathscr{F}[u_t]$ is differentiable at $t = 0$, then its derivative at $t = 0$ must vanish because of the minimization property. This is analogous to the elementary fact that if $g \in \mathrm{C}^1((0,T))$ takes its minimum at a point $t_* \in (0,T)$, then $g'(t_*) = 0$.

The **first variation** $\delta\mathscr{F}[u]$ of $\mathscr{F}$ at $u \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^m)$ is the linear map

$$\delta\mathscr{F}[u]\colon \mathrm{C}^\infty_c(\Omega;\mathbb{R}^m) \to \mathbb{R}$$

defined as

$$\delta\mathscr{F}[u][\psi] := \lim_{h\downarrow 0} \frac{\mathscr{F}[u+h\psi] - \mathscr{F}[u]}{h}, \qquad \psi \in \mathrm{C}^\infty_c(\Omega;\mathbb{R}^m), \tag{3.1}$$

assuming that this limit exists. By the argument above,

$$\delta\mathscr{F}[u_*] = 0$$

at every minimizer u_*. This yields a partial differential equation, called the *Euler–Lagrange equation*, which minimizers necessarily satisfy in the weak sense. Under an additional convexity assumption on $\mathscr{F}$, the Euler–Lagrange equation turns out to be sufficient for a map to be a minimizer as well. Thus, at least for convex problems, we can find a minimizer by solving the Euler–Lagrange equation. These "calculations with variations" gave the field its name.

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8_3

A related topic, which we only touch upon in this chapter, is the *regularity theory* of minimizers. It turns out that for so-called *regular variational integrals*, minimizers are always smooth. This is the famous solution to David Hilbert's 19th problem by Ennio De Giorgi and John F. Nash, which we briefly outline (without proving the more technical aspects).

If we assume a side constraint as in Section 2.5, the paths $t \mapsto u_t$ above have to take into account this side constraint as well, which leads to the statement that the minimizer satisfies a generalization of the Euler–Lagrange equation, which involves a so-called *Lagrange multiplier*.

Finally, we discuss invariances of the integral functional, that is, nontrivial paths $t \mapsto u_t$ along which $\mathscr{F}[u_t]$ is constant. This leads to a famous theorem by Emmy Noether, which exposes "hidden" conservation laws in minimization problems.

3.1 The Euler–Lagrange Equation

Let the **directional derivative** $\mathrm{D}_A f(x, v, A) \in \mathbb{R}^{m\times d}$ of $f(x, v, \bullet)$ at A in direction B be defined via

$$\mathrm{D}_A f(x, v, A) : B := \lim_{h\downarrow 0} \frac{f(x, v, A + hB) - f(x, v, A)}{h}, \qquad A, B \in \mathbb{R}^{m\times d}.$$

We remark that when we require f to be "differentiable in A", then this entails that the derivative $\mathrm{D}_A f$ of f in A is linear in the direction B and hence the directional derivative can be represented via the Frobenius product ":" (see Appendix A.1) between $\mathrm{D}_A f(x, v, A)$ and B. A similar remark applies to the directional derivative $\mathrm{D}_v f(x, v, A)$ of $f(x, \bullet, A)$, where we now use the usual scalar product to pair a location vector with a direction vector. In fact, the matrix $\mathrm{D}_A f(x, v, A)$ and the vector $\mathrm{D}_v f(x, v, A)$ are given as

$$\mathrm{D}_A f(x, v, A) := \bigl(\partial_{A^j_k} f(x, v, A)\bigr)^j_k, \qquad \mathrm{D}_v f(x, v, A) := \bigl(\partial_{v^j} f(x, v, A)\bigr)^j.$$

The following theorem furnishes the connection between the calculus of variations and PDE theory. It is very useful if we want to actually *compute* minimizers, either by hand or numerically.

Theorem 3.1. *Let $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}$ be a Carathéodory integrand that is continuously differentiable in the second and third arguments and that satisfies the growth bounds*

$$|\mathrm{D}_v f(x, v, A)|, |\mathrm{D}_A f(x, v, A)| \le C(1 + |v|^p + |A|^p),$$

for $(x, v, A) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$, a constant $C > 0$, and $p \in [1, \infty)$. If $u_* \in \mathrm{W}^{1,p}_g(\Omega; \mathbb{R}^m)$, where $g \in \mathrm{W}^{1-1/p,p}(\partial\Omega; \mathbb{R}^m)$, minimizes the functional

$$\mathscr{F}[u] := \int_\Omega f(x, u(x), \nabla u(x)) \, \mathrm{d}x, \qquad u \in \mathrm{W}^{1,p}_g(\Omega; \mathbb{R}^m),$$

then u_* is a weak solution of the ***Euler–Lagrange equation***

$$\begin{cases} -\operatorname{div}\big[\mathrm{D}_A f(x, u, \nabla u)\big] + \mathrm{D}_v f(x, u, \nabla u) = 0 & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases} \tag{3.2}$$

Here, $u_* \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ is called a **weak solution** of (3.2) if

$$\int_\Omega \mathrm{D}_A f(x, u_*, \nabla u_*) : \nabla \psi + \mathrm{D}_v f(x, u_*, \nabla u_*) \cdot \psi \, \mathrm{d}x = 0 \tag{3.3}$$

for all $\psi \in \mathrm{C}^\infty_c(\Omega; \mathbb{R}^m)$. Note that the Euler–Lagrange "equation" is actually a *system* of PDEs (or, more precisely, a *boundary value problem* for this system). We also used the common convention to omit the x-arguments whenever this does not cause any confusion in order to curtail the proliferation of x's, for example in $f(x, u, \nabla u) = f(x, u(x), \nabla u(x))$. The boundary condition $u = g$ on $\partial\Omega$ in (3.2) is to be understood in the sense of trace, as usual.

Proof. For all $\psi \in \mathrm{C}^\infty_c(\Omega; \mathbb{R}^m)$ and all $h > 0$ we have

$$\mathscr{F}[u_*] \le \mathscr{F}[u_* + h\psi]$$

since $u_* + h\psi \in \mathrm{W}^{1,p}_g(\Omega; \mathbb{R}^m)$ is admissible in the minimization. Thus,

$$\begin{aligned} 0 &\le \int_\Omega \frac{f(x, u_* + h\psi, \nabla u_* + h\nabla\psi) - f(x, u_*, \nabla u_*)}{h} \, \mathrm{d}x \\ &= \int_\Omega \int_0^1 \frac{1}{h} \frac{\mathrm{d}}{\mathrm{d}t} \big[f(x, u_* + th\psi, \nabla u_* + th\nabla\psi) \big] \, \mathrm{d}t \, \mathrm{d}x \\ &= \int_\Omega \int_0^1 \mathrm{D}_A f(x, u_* + th\psi, \nabla u_* + th\nabla\psi) : \nabla\psi \\ &\qquad\qquad + \mathrm{D}_v f(x, u_* + th\psi, \nabla u_* + th\nabla\psi) \cdot \psi \, \mathrm{d}t \, \mathrm{d}x. \end{aligned}$$

By the growth bounds on the derivative, the integrand can be seen to have an h-uniform majorant, namely $C(1 + |u_*|^p + |\psi|^p + |\nabla u_*|^p + |\nabla\psi|^p)$ if we additionally assume $h \le 1$, and so we may apply the Lebesgue dominated convergence theorem to let $h \downarrow 0$ under the double integral. This yields

$$0 \le \int_\Omega \mathrm{D}_A f(x, u_*, \nabla u_*) : \nabla\psi + \mathrm{D}_v f(x, u_*, \nabla u_*) \cdot \psi \, \mathrm{d}x$$

and we conclude (3.3) by taking ψ and $-\psi$ in this inequality. □

Remark 3.2. If we want to allow $\psi \in \mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^m)$ in the weak formulation (3.3), then we need to assume the stronger growth conditions

$$|\mathrm{D}_v f(x,v,A)|,\, |\mathrm{D}_A f(x,v,A)| \le C(1+|v|^{p-1}+|A|^{p-1}) \tag{3.4}$$

for some $C > 0$ and $p \in [1,\infty)$ in order for (3.3) to be well-defined and finite (by Hölder's inequality). The extension to test functions $\psi \in \mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^m)$ then follows by a density argument. Indeed, (3.3) and (3.4) imply that the linear functional $T \in \mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^m)^*$ defined for $\psi \in \mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^m)$ as

$$\langle T,\psi\rangle := \int_\Omega \mathrm{D}_A f(x,u_*,\nabla u_*) : \nabla\psi + \mathrm{D}_v f(x,u_*,\nabla u_*)\cdot\psi \,\mathrm{d}x$$

satisfies

$$\langle T,\psi\rangle = 0 \qquad \text{for all } \psi \in \mathrm{C}_c^\infty(\Omega;\mathbb{R}^m)$$

and, by Hölder's inequality,

$$\begin{aligned}\big|\langle T,\psi\rangle\big| &\le C\int_\Omega \big(1+|u_*|^{p-1}+|\nabla u_*|^{p-1}\big)\big(|\psi|+|\nabla\psi|\big)\,\mathrm{d}x\\ &\le C\big(1+\|u_*\|_{\mathrm{L}^p}^{p-1}+\|\nabla u_*\|_{\mathrm{L}^p}^{p-1}\big)\|\psi\|_{\mathrm{W}^{1,p}}.\end{aligned}$$

For $\psi \in \mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^m)$ take $(\psi_j) \subset \mathrm{C}_c^\infty(\Omega;\mathbb{R}^m)$ such that $\psi_j \to \psi$ in $\mathrm{W}^{1,p}$ (such a sequence always exists by the definition of $\mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^m)$). Then, as T is continuous on $\mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^m)$ by the above estimate,

$$\langle T,\psi\rangle = \lim_{j\to\infty}\langle T,\psi_j\rangle = 0$$

and (3.3) follows.

Using the notion of first variation, see (3.1), the assertion of Theorem 3.1 can be written as

$$\delta\mathscr{F}[u_*] = 0 \quad \text{if } u_* \text{ minimizes } \mathscr{F} \text{ over } \mathrm{W}_g^{1,p}(\Omega;\mathbb{R}^m).$$

Of course, this condition is only *necessary* for u to be a minimizer. In fact, any solution of the Euler–Lagrange equation is called a **critical point** of $\mathscr{F}$, which could be a minimizer, a maximizer, or a saddle point. However, under a convexity assumption, the solution to the Euler–Lagrange equation is always a minimizer:

Proposition 3.3. *In the situation of Theorem 3.1, assume furthermore that the stronger growth conditions in* (3.4) *hold and that* $(v,A)\mapsto f(x,v,A)$ *is (jointly) convex for all* $x \in \Omega$. *If* $u_* \in \mathrm{W}_g^{1,p}(\Omega;\mathbb{R}^m)$ *solves* (3.2)*, then* u_* *is a minimizer of* $\mathscr{F}$.

Proof. Let $v \in W_g^{1,p}(\Omega;\mathbb{R}^m)$ and set $\psi := v - u_* \in W_0^{1,p}(\Omega;\mathbb{R}^m)$. Consider the function

$$g(t) := \mathscr{F}[u_* + t\psi], \qquad t \in \mathbb{R},$$

which inherits convexity from $\mathscr{F}$. By the same arguments as in the proof of Theorem 3.1, g is differentiable. Since u_* solves the Euler–Lagrange equation, we have

$$\frac{\mathrm{d}}{\mathrm{d}t} g(t)\bigg|_{t=0} = 0.$$

Then, from the convexity,

$$g(t) \geq g(0) + tg'(0) = g(0) = \mathscr{F}[u_*], \qquad t \geq 0.$$

Setting $t = 1$, we get $\mathscr{F}[v] \geq \mathscr{F}[u_*]$. As $v \in W_g^{1,p}(\Omega;\mathbb{R}^m)$ was arbitrary, u_* must be a minimizer. □

Example 3.4. Returning to the Dirichlet functional from Example 2.8, we see that the associated Euler–Lagrange equation is the **Laplace equation**

$$-\Delta u = 0 \quad \text{in } \Omega,$$

where

$$\Delta := \partial_1^2 + \cdots + \partial_d^2$$

is the **Laplace operator**. Solutions u to $-\Delta u = 0$ are called **harmonic maps** (they are always strong solutions by Example 3.15 below, so there is no distinction between weak and strong harmonic maps). Since the Dirichlet functional is convex, by Proposition 3.3 all solutions of the Laplace equation are in fact minimizers of the Dirichlet functional. Furthermore, by Proposition 2.10, it can be seen that solutions of the Laplace equation are unique for given boundary values. The same assertions also apply to the functional

$$\mathscr{F}[u] := \int_\Omega \frac{1}{2}|\nabla u(x)|^2 - h(x)\cdot u(x)\,\mathrm{d}x, \qquad u \in W^{1,2}(\Omega;\mathbb{R}^m),$$

where $h \in L^2(\Omega;\mathbb{R}^m)$. Here, the Euler–Lagrange equation is the **Poisson equation**

$$-\Delta u = h \quad \text{in } \Omega.$$

Example 3.5. In the linearized elasticity problem from Example 2.12, we may compute the Euler–Lagrange equation to be

$$\begin{cases} -\operatorname{div}\Big[2\mu\,\mathscr{E}u + \Big(\kappa - \frac{2}{3}\mu\Big)(\operatorname{tr}\mathscr{E}u)\mathrm{Id}\Big] = b & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases}$$

One crucial consequence of Theorem 3.1 is that we can use all available PDE methods to study minimizers. Immediately, one can ask about the type of PDE we are dealing with. In this respect we have the following prototypical result.

Proposition 3.6. *In the situation of Theorem 3.1, assume furthermore that f does not depend on v and is quadratic in A, i.e.,*

$$f(x, v, A) = \frac{1}{2} A : \mathbf{S}(x)A, \qquad (x, v, A) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times d},$$

for a fourth-order symmetric tensor $\mathbf{S}(x) = \mathbf{S}^{ik}_{jl}(x)$ ($x \in \Omega$). Then, the Euler–Lagrange equation is the linear PDE

$$\begin{cases} -\operatorname{div}[\mathbf{S}\nabla u] = 0 & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases}$$

Moreover,

$$B : \mathbf{S}(x)B = \mathrm{D}^2_A f(x, v, A)[B, B] := \frac{\mathrm{d}^2}{\mathrm{d}t^2} f(x, v, A + tB)\Big|_{t=0}$$

for all $x \in \Omega$, $v \in \mathbb{R}^m$, and $A, B \in \mathbb{R}^{m\times d}$. Consequently, $\mathbf{S}$ is positively semidefinite if and only if $f(x, v, \bullet)$ is convex. In this case, the above PDE is (possibly degenerate) ***elliptic****.*

The proof is immediate from Theorem 3.1 and the relevant definitions.

We now use the Euler–Lagrange equation to find concrete solutions of variational problems.

Example 3.7. Recall the optimal saving problem from Section 1.5,

$$\begin{cases} \text{Minimize} \quad \mathscr{F}[S] := \displaystyle\int_0^T -\ln(1 + w + \rho S(t) - \dot{S}(t))\,\mathrm{d}t \\ \text{subject to} \quad S(0) = 0,\ \ S(T) = S_T \ge 0,\ \ C(t) := w + \rho S(t) - \dot{S}(t) \ge 0. \end{cases}$$

Since $a \mapsto -\ln(\beta - a)$ is strictly convex for any $\beta > 0$, we know from Proposition 2.10 that if a solution exists, then it is unique. The Euler–Lagrange equation is

$$-\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{1 + w + \rho S(t) - \dot{S}(t)}\right] = \frac{\rho}{1 + w + \rho S(t) - \dot{S}(t)}. \tag{3.5}$$

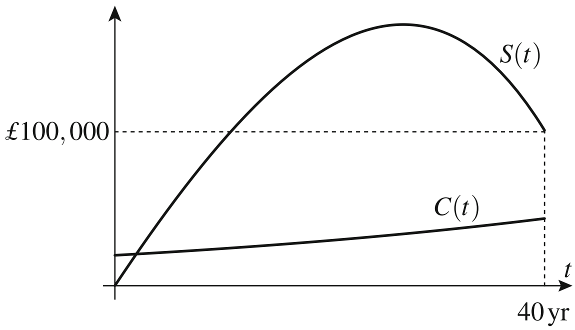


Fig. 3.1 The solution to the optimal saving problem

Technically, Theorem 3.1 is not applicable since the validity of the growth-bounds is a priori unclear. However, around the solution that will be computed below, they are in fact satisfied and so we can a posteriori justify 3.5. By Proposition 3.3 (again, a posteriori justified), we get that this solution of the Euler–Lagrange equation is also a minimizer of our functional.

In order to solve (3.5), we rearrange it into

$$\frac{\rho \dot{S}(t) - \ddot{S}(t)}{1 + w + \rho S(t) - \dot{S}(t)} = \rho.$$

With the modified consumption rate $C_*(t) := 1 + C(t) = 1 + w + \rho S(t) - \dot{S}(t)$, this is equivalent to

$$\frac{\dot{C}_*(t)}{C_*(t)} = \rho,$$

and so, if $C(0) = C_0$ (to be determined later),

$$1 + w + \rho S(t) - \dot{S}(t) = C_*(t) = \mathrm{e}^{\rho t} C_*(0) = \mathrm{e}^{\rho t}(1 + C_0).$$

This ordinary differential equation for $S(t)$ can be solved, for example via the Duhamel principle, which yields

$$\begin{aligned} S(t) &= \mathrm{e}^{\rho t} \cdot 0 + \int_0^t \mathrm{e}^{\rho(t-s)}(1 + w - \mathrm{e}^{\rho s} - \mathrm{e}^{\rho s} C_0)\, \mathrm{d}s \\ &= \frac{\mathrm{e}^{\rho t} - 1}{\rho}(1 + w) - t\mathrm{e}^{\rho t}(1 + C_0), \end{aligned}$$

and C_0 can now be chosen to satisfy the terminal condition $S(T) = S_T$, in fact, $C_0 = (1 - \mathrm{e}^{-\rho T})(1 + w)/(\rho T) - \mathrm{e}^{-\rho T} S_T / T - 1$.

In Figure 3.1 we see the optimal saving strategy for a worker earning a (constant) continuously-paid salary of $w = £30,000$ per year and having a savings goal of $S_T = £100,000$. The effective APR for savings is set at 2% per year ($\rho = 0.0198$).

The worker has to save for approximately 27 years, reaching savings of just over £168, 000, and then starts withdrawing his savings ($\dot{S}(t) < 0$) for the last 13 years. The worker's consumption $C(t) = e^{\rho t}(1 + C_0) - 1$ goes up continuously during the whole working life.

Example 3.8. For functions $u = u(t, x)\colon \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ consider the functional

$$\mathscr{F}[u] := \int_{\mathbb{R}} \int_{\mathbb{R}^d} \frac{1}{2}\bigl(-|\partial_t u|^2 + |\nabla_x u|^2\bigr) \,dx\, dt,$$

where $\nabla_x u$ is the gradient of u with respect to $x \in \mathbb{R}^d$. This functional should be interpreted as the usual Dirichlet functional with respect to the Lorentz metric. Then, the Euler–Lagrange equation is the **wave equation**

$$\partial_t^2 u - \Delta u = 0 \quad \text{in } \mathbb{R} \times \mathbb{R}^d.$$

Notice that the integrand of $\mathscr{F}$ is not convex and the wave equation is *hyperbolic*.

It is an important question whether a weak solution of the Euler–Lagrange equation (3.2) is also a **strong solution**, that is, whether $u \in W^{2,2}(\Omega; \mathbb{R}^m)$ and

$$\begin{cases} -\operatorname{div}\bigl[D_A f(x, u(x), \nabla u(x))\bigr] + D_v f(x, u(x), \nabla u(x)) = 0 & \text{for a.e. } x \in \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases} \tag{3.6}$$

If $u \in C^2(\Omega; \mathbb{R}^m) \cap C(\overline{\Omega}; \mathbb{R}^m)$ satisfies this PDE for *every* $x \in \Omega$, then we call u a **classical solution**.

Multiplying (3.6) by a test function $\psi \in C_c^\infty(\Omega; \mathbb{R}^m)$, integrating over Ω, and using the Gauss–Green theorem, it follows that any solution of (3.6) also solves (3.2) in the weak sense, i.e., (3.3) holds. The converse is true whenever u is sufficiently regular:

Proposition 3.9. *Let the integrand f be twice continuously differentiable and let $u \in W^{2,2}(\Omega; \mathbb{R}^m)$ be a weak solution of the Euler–Lagrange equation* (3.2)*. Then, u solves the Euler–Lagrange equation* (3.6) *in the strong sense.*

Proof. If $u \in W^{2,2}(\Omega; \mathbb{R}^m)$ is a weak solution, then

$$\int_\Omega D_A f(x, u, \nabla u) : \nabla\psi + D_v f(x, u, \nabla u) \cdot \psi \,dx = 0$$

for all $\psi \in C_c^\infty(\Omega; \mathbb{R}^m)$. Integration by parts (more precisely, the Gauss–Green theorem) gives

$$\int_\Omega \bigl(-\operatorname{div}\bigl[D_A f(x, u, \nabla u)\bigr] + D_v f(x, u, \nabla u)\bigr) \cdot \psi \,dx = 0$$

for all ψ as before. We conclude using the following so-called *Fundamental Lemma* of the calculus of variations. □

Lemma 3.10. *Let* $\Omega \subset \mathbb{R}^d$ *be open. If* $g \in \mathrm{L}^1(\Omega)$ *satisfies*

$$\int_\Omega g\psi \,\mathrm{d}x = 0 \qquad \textit{for all } \psi \in \mathrm{C}_c^\infty(\Omega),$$

then $g = 0$ *almost everywhere.*

Proof. We can assume that Ω is bounded by considering subdomains if necessary. Also, let g be extended by zero to all of $\mathbb{R}^d$. Fix $\varepsilon > 0$ and let $(\eta_\delta)_{\delta>0}$ be a family of mollifiers, see Appendix A.5. Then, since $\eta_\delta \star g \to g$ in L^1, there is a function $h \in \mathrm{C}_c^\infty(\mathbb{R}^d)$ with the properties

$$\|g - h\|_{\mathrm{L}^1} \leq \frac{\varepsilon}{4} \qquad \text{and} \qquad \|h\|_\infty < \infty.$$

Set $\phi(x) := h(x)/|h(x)|$ for $h(x) \neq 0$ and $\phi(x) := 0$ for $h(x) = 0$, so that $h\phi = |h|$. Then define $\psi := \eta_\delta \star \phi \in \mathrm{C}_c^\infty(\mathbb{R}^d)$ for some $\delta > 0$ such that

$$\|\phi - \psi\|_{\mathrm{L}^1} \leq \frac{\varepsilon}{2(1 + \|h\|_\infty)}.$$

Since $|\psi| \leq 1$ (this follows from the definition of the convolution),

$$\begin{aligned}
\|g\|_{\mathrm{L}^1} &\leq \|g - h\|_{\mathrm{L}^1} + \int h\phi \,\mathrm{d}x \\
&= \|g - h\|_{\mathrm{L}^1} + \int h(\phi - \psi) + (h - g)\psi + g\psi \,\mathrm{d}x \\
&\leq 2\|g - h\|_{\mathrm{L}^1} + \|h\|_\infty \cdot \|\phi - \psi\|_{\mathrm{L}^1} + 0 \\
&\leq \varepsilon.
\end{aligned}$$

We conclude by letting $\varepsilon \downarrow 0$. □

3.2 Regularity of Minimizers

We saw at the end of the last section that if a weak solution of the Euler–Lagrange equation has higher regularity (differentiability), then it is also a strong or even a classical solution. More generally, one would like to know *how much* regularity we can expect from solutions of a variational problem. Such a question was the content of Hilbert's 19th problem [149]:

> *Does every Lagrangian partial differential equation of a regular variational problem have the property of exclusively admitting analytic integrals?*[1]

[1] The German original asks "ob jede Lagrangesche partielle Differentialgleichung eines regulären Variationsproblems die Eigenschaft hat, daß sie nur analytische Integrale zuläßt."

In modern language, Hilbert asked whether "regular" variational problems (defined below) admit only analytic solutions, i.e., solutions that have a local power series representation.

In this section, we will prove some basic regularity assertions, but we will only sketch the solution of Hilbert's 19th problem as the techniques needed are quite involved. We remark at the outset that many regularity results are very sensitive to the dimensions of the domain and the target space. In particular, the behavior of the scalar case ($m = 1$) and the vector case ($m > 1$) is fundamentally different.

In the spirit of Hilbert's 19th problem, call

$$\mathscr{F}[u] := \int_\Omega f(\nabla u(x))\,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^m),$$

a **regular variational integral** if $f : \mathbb{R}^{m\times d} \to \mathbb{R}$ is twice continuously differentiable and there are constants $\mu, M > 0$ with

$$\mu|B|^2 \le \mathrm{D}^2 f(A)[B,B] \le M|B|^2, \qquad A, B \in \mathbb{R}^{m\times d}, \tag{3.7}$$

where

$$\mathrm{D}^2 f(A)[B,B] := \frac{\mathrm{d}^2}{\mathrm{d}t^2} f(A + tB)\bigg|_{t=0}.$$

Clearly, regular variational problems are convex. In fact, integrands f that satisfy the lower bound in (3.7) are called **strongly convex**. For example, the Dirichlet functional from Example 2.8 is a regular variational integral.

Since f is twice continuously differentiable,

$$\mathrm{D}^2 f(A)[B_1,B_2] = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}s} f(A + sB_1 + tB_2)\bigg|_{s,t=0}, \qquad A, B_1, B_2 \in \mathbb{R}^{m\times d},$$

is a symmetric bilinear form in B_1, B_2. One checks that for $B_1 = B_2 = B$ this agrees with $\mathrm{D}^2 f(A)[B,B]$ as defined above. It can be shown from (3.7) using basic linear algebra that

$$\left|\mathrm{D}^2 f(A)[B_1,B_2]\right| \le M|B_1||B_2|. \tag{3.8}$$

Then, by the mean value theorem, we also get that $\mathrm{D}f$ is Lipschitz continuous, that is,

$$|\mathrm{D}f(A_1) - \mathrm{D}f(A_2)| \le M|A_1 - A_2|, \qquad A_1, A_2 \in \mathbb{R}^{m\times d}, \tag{3.9}$$

and in particular (for a different $M > 0$)

$$|\mathrm{D}f(A)| \le M(1 + |A|), \qquad A \in \mathbb{R}^{m\times d}.$$

The fundamental $\mathrm{W}^{2,2}_{\mathrm{loc}}$-regularity theorem is the following.

Theorem 3.11. *Let $\mathscr{F}$ be a regular variational integral. Then, for any minimizer $u_* \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^m)$ of $\mathscr{F}$ it holds that*

$$u_* \in \mathrm{W}^{2,2}_{\mathrm{loc}}(\Omega; \mathbb{R}^m).$$

Moreover, for any ball $B(x_0, 3r) \subset \Omega$ ($x_0 \in \Omega$, $r > 0$) the ***Caccioppoli inequality***

$$\int_{B(x_0,r)} |\nabla^2 u_*(x)|^2 \,\mathrm{d}x \le \left(\frac{2M}{\mu}\right)^2 \int_{B(x_0,3r)} \frac{|\nabla u_*(x) - [\nabla u_*]_{B(x_0,3r)}|^2}{r^2} \,\mathrm{d}x \quad (3.10)$$

holds, where $[\nabla u_*]_{B(x_0,3r)} := \fint_{B(x_0,3r)} \nabla u_* \,\mathrm{d}x$. Consequently, the Euler–Lagrange equation is satisfied strongly,

$$-\operatorname{div} \mathrm{D} f(\nabla u_*) = 0 \quad \textit{a.e. in } \Omega.$$

Here, we recall that, as usual, $\fint_{B(x_0,r)} := \omega_d^{-1} r^{-d} \int_{B(x_0,r)}$, where $\omega_d := |B(0,1)|$ is the volume of the d-dimensional unit ball.

Before we come to the formal proof, let us explain the idea by establishing the Caccioppoli inequality (3.10) *assuming* that $u_* \in \mathrm{C}^\infty(\overline{\Omega}; \mathbb{R}^m)$. In this case, for any ball $B(x_0, 3r) \subset \Omega$ ($x_0 \in \Omega$, $r > 0$) take a Lipschitz cut-off function $\rho \in \mathrm{W}^{1,\infty}_0(\Omega; [0,1])$ such that

$$\mathbb{1}_{B(x_0,r)} \le \rho \le \mathbb{1}_{B(x_0,2r)} \qquad \text{and} \qquad |\nabla \rho| \le \frac{1}{r}.$$

Then test the (weak) Euler–Lagrange equation (3.3) with $\psi := \partial_k[\rho^2 \partial_k(u_* - a)]$ for some to be determined affine map $a \colon \mathbb{R}^d \to \mathbb{R}^m$ and any $k \in \{1, \ldots, d\}$. Using integration by parts,

$$\begin{aligned} 0 &= -\int_\Omega \mathrm{D} f(\nabla u_*) : \nabla\big(\partial_k[\rho^2 \partial_k(u_* - a)]\big) \,\mathrm{d}x \\ &= \int_\Omega \partial_k(\mathrm{D} f(\nabla u_*)) : \big[\rho^2 \partial_k \nabla u_* + \partial_k(u_* - a) \otimes \nabla(\rho^2)\big] \,\mathrm{d}x \\ &= \int_\Omega \rho^2 \mathrm{D}^2 f(\nabla u_*)[\partial_k \nabla u_*, \partial_k \nabla u_*] \,\mathrm{d}x \\ &\qquad + \int_\Omega \partial_k(\mathrm{D} f(\nabla u_*)) : [\partial_k(u_* - a) \otimes \nabla(\rho^2)] \,\mathrm{d}x. \end{aligned}$$

Here we remark that the tensor product "$\otimes$" is technically incorrect since we are multiplying a column vector with a row vector, but we include it here and in the following to signify that the result is a *matrix*. Then, using the bounds (3.7), (3.8) on $\mathrm{D}^2 f$ and Young's inequality,

$$
\begin{aligned}
\mu \int_{\Omega} \rho^2 |\partial_k \nabla u_*|^2 \, \mathrm{d}x &\le \int_{\Omega} \rho^2 \mathrm{D}^2 f(\nabla u_*)[\partial_k \nabla u_*, \partial_k \nabla u_*] \, \mathrm{d}x \\
&= - \int_{\Omega} \partial_k (\mathrm{D} f(\nabla u_*)) : [\partial_k (u_* - a) \otimes \nabla(\rho^2)] \, \mathrm{d}x \\
&= - \int_{\Omega} \mathrm{D}^2 f(\nabla u_*) : [\partial_k (u_* - a) \otimes \nabla(\rho^2), \partial_k \nabla u_*] \, \mathrm{d}x \\
&\le 2M \int_{\Omega} \rho |\partial_k \nabla u_*| \cdot |\partial_k (u_* - a)| \cdot |\nabla \rho| \, \mathrm{d}x \\
&\le \frac{\mu}{2} \int_{\Omega} \rho^2 |\partial_k \nabla u_*|^2 \, \mathrm{d}x + \frac{2M^2}{\mu} \int_{\Omega} |\partial_k (u_* - a)|^2 \cdot |\nabla \rho|^2 \, \mathrm{d}x.
\end{aligned}
$$

We absorb the first term on the right-hand side into the left-hand side and use the properties of ρ to infer that

$$
\frac{\mu}{2} \int_{B(x_0,r)} |\partial_k \nabla u_*|^2 \, \mathrm{d}x \le \frac{2M^2}{\mu} \int_{B(x_0,3r)} \frac{|\partial_k (u_* - a)|^2}{r^2} \, \mathrm{d}x.
$$

Multiplying by $2/\mu$, summing over k, and choosing a with $\nabla a = [\nabla u_*]_{B(x_0,3r)}$, we arrive at (3.10). This shows that for minimizers that are assumed smooth, the first-order derivatives control the second order derivatives.

For the rigorous proof we will employ the *difference quotient method*, which is fundamental in regularity theory. For $u \colon \Omega \to \mathbb{R}^m$, define the k'th **difference quotient**, $k \in \{1, \ldots, d\}$, of u at $x \in \Omega$ with height $h \in \mathbb{R} \setminus \{0\}$ to be

$$
\mathrm{D}_k^h u(x) := \frac{u(x + h\mathrm{e}_k) - u(x)}{h},
$$

where $\{\mathrm{e}_1, \ldots, \mathrm{e}_d\}$ is the standard basis of $\mathbb{R}^d$. We also set

$$
\mathrm{D}^h u := (\mathrm{D}_1^h u, \ldots, \mathrm{D}_d^h u).
$$

The key to the difference quotient method is the following characterization of Sobolev spaces.

Lemma 3.12. *Let $D \Subset \Omega \subset \mathbb{R}^d$ be open sets, $p \in (1, \infty)$, and $u \in \mathrm{L}^p(\Omega; \mathbb{R}^m)$.*

(i) If $u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$, then

$$
\|\mathrm{D}_k^h u\|_{\mathrm{L}^p(D)} \le \|\partial_k u\|_{\mathrm{L}^p(\Omega)} \qquad \text{for all } k \in 1, \ldots, d,\ |h| < \operatorname{dist}(D, \partial\Omega).
$$

(ii) If for some $0 < \delta < \operatorname{dist}(D, \partial\Omega)$ it holds that

$$
\|\mathrm{D}_k^h u\|_{\mathrm{L}^p(D)} \le C \qquad \text{for all } k \in \{1, \ldots, d\} \text{ and all } |h| < \delta,
$$

then $u \in \mathrm{W}^{1,p}(D; \mathbb{R}^m)$ and $\|\partial_k u\|_{\mathrm{L}^p(D)} \le C$ for all $k \in \{1, \ldots, d\}$.

Proof. For (i) assume first that $u \in (\mathrm{L}^p \cap \mathrm{C}^1)(\Omega; \mathbb{R}^m)$. In this case, by the fundamental theorem of calculus, at $x \in \Omega$ it holds that

$$\mathrm{D}_k^h u(x) = \frac{1}{h} \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} u(x + th\mathrm{e}_k) \,\mathrm{d}t = \int_0^1 \partial_k u(x + th\mathrm{e}_k) \,\mathrm{d}t.$$

Thus, by Jensen's inequality (see Lemma A.18),

$$\int_D |\mathrm{D}_k^h u|^p \,\mathrm{d}x \le \int_D \int_0^1 |\partial_k u(x + th\mathrm{e}_k)|^p \,\mathrm{d}t \,\mathrm{d}x \le \int_\Omega |\partial_k u|^p \,\mathrm{d}x,$$

from which the assertion is clear. The general case follows from the density of $(\mathrm{L}^p \cap \mathrm{C}^1)(\Omega; \mathbb{R}^m)$ in $\mathrm{L}^p(\Omega; \mathbb{R}^m)$.

For (ii), we observe that for fixed $k \in \{1, \dots, d\}$ by assumption $(\mathrm{D}_k^h u)_{0<h<\delta}$ is uniformly norm-bounded in $\mathrm{L}^p(D; \mathbb{R}^m)$. Thus, for an arbitrary fixed sequence of h's tending to zero, there exists a subsequence $h_j \downarrow 0$ with

$$\mathrm{D}_k^{h_j} u \rightharpoonup v_k \quad \text{in } \mathrm{L}^p$$

for some $v_k \in \mathrm{L}^p(D; \mathbb{R}^m)$. Let $\psi \in \mathrm{C}_c^\infty(D; \mathbb{R}^m)$. Using an "integration-by-parts" rule for difference quotients, which is elementary to check, we get

$$\int_D v_k \cdot \psi \,\mathrm{d}x = \lim_{j\to\infty} \int_\Omega \mathrm{D}_k^{h_j} u \cdot \psi \,\mathrm{d}x = -\lim_{j\to\infty} \int_\Omega u \cdot \mathrm{D}_k^{-h_j} \psi \,\mathrm{d}x = -\int_D u \cdot \partial_k \psi \,\mathrm{d}x.$$

Thus, $u \in \mathrm{W}^{1,p}(D; \mathbb{R}^m)$ and $v_k = \partial_k u$. The norm estimate follows from the lower semicontinuity of the norm under weak convergence (which is well-known in functional analysis, but can also be proved using the Tonelli–Serrin Theorem 2.6). □

Proof of Theorem 3.11. The idea is to emulate the a priori non-existent second derivatives using difference quotients and to derive estimates which allow one to conclude that these difference quotients are uniformly (in the height) bounded in L^2. Then we can conclude by the preceding lemma.

Let $u_* \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^m)$ be a minimizer of $\mathscr{F}$. By Theorem (3.1) and Remark (3.2),

$$0 = \int_\Omega \mathrm{D}f(\nabla u_*) : \nabla\psi \,\mathrm{d}x \qquad \text{for all } \psi \in \mathrm{W}_0^{1,2}(\Omega; \mathbb{R}^m). \tag{3.11}$$

Fix a ball $B(x_0, 3r) \subset \Omega$ and take a Lipschitz cut-off function $\rho \in \mathrm{W}_0^{1,\infty}(\Omega; [0, 1])$ such that

$$\mathbb{1}_{B(x_0,r)} \le \rho \le \mathbb{1}_{B(x_0,2r)} \qquad \text{and} \qquad |\nabla\rho| \le \frac{1}{r}.$$

Then, for any $k = 1, \dots, d$ and $|h| < r$ we let

$$\psi := \mathrm{D}_k^{-h}\big[\rho^2 \mathrm{D}_k^h (u_* - a)\big] \in \mathrm{W}_0^{1,2}(\Omega; \mathbb{R}^m),$$

where $a\colon \mathbb{R}^d \to \mathbb{R}^m$ is an affine function to be chosen later. We may plug this ψ into (3.11) to get

$$0 = \int_\Omega \mathrm{D}_k^h(\mathrm{D} f(\nabla u_*)) : \big[\rho^2 \mathrm{D}_k^h \nabla u_* + \mathrm{D}_k^h (u_* - a) \otimes \nabla(\rho^2)\big] \,\mathrm{d}x. \tag{3.12}$$

Here, we again used the "integration-by-parts" formula for difference quotients from the proof of Lemma (3.12).

Next, we estimate, using the assumptions on f,

$$\begin{aligned}
\mu |\mathrm{D}_k^h \nabla u_*|^2 &\le \int_0^1 \mathrm{D}^2 f(\nabla u_* + th\mathrm{D}_k^h \nabla u_*)[\mathrm{D}_k^h \nabla u_*, \mathrm{D}_k^h \nabla u_*] \,\mathrm{d}t \\
&= \frac{1}{h} \mathrm{D} f(\nabla u_* + th\mathrm{D}_k^h \nabla u_*) : \mathrm{D}_k^h \nabla u_* \Big|_{t=0}^1 \\
&= \mathrm{D}_k^h(\mathrm{D} f(\nabla u_*)) : \mathrm{D}_k^h \nabla u_*,
\end{aligned}$$

where for the last line we note that

$$\begin{aligned}
\mathrm{D}_k^h(\mathrm{D} f(\nabla u_*))(x) &= \frac{\mathrm{D} f(\nabla u_*(x + h\mathrm{e}_k)) - \mathrm{D} f(\nabla u_*(x))}{h} \\
&= \frac{\mathrm{D} f(\nabla u_*(x) + h\mathrm{D}_k^h \nabla u_*(x)) - \mathrm{D} f(\nabla u_*(x))}{h}.
\end{aligned}$$

On the other hand, using the Cauchy–Schwarz and Young inequalities,

$$\begin{aligned}
&\big|\mathrm{D}_k^h(\mathrm{D} f(\nabla u_*)) : \big[\mathrm{D}_k^h (u_* - a) \otimes \nabla(\rho^2)\big]\big| \\
&\quad \le 2\rho |\mathrm{D}_k^h(\mathrm{D} f(\nabla u_*))| \cdot |\mathrm{D}_k^h (u_* - a)| \cdot |\nabla \rho| \\
&\quad \le \frac{\mu}{2M^2} \rho^2 |\mathrm{D}_k^h(\mathrm{D} f(\nabla u_*))|^2 + \frac{2M^2}{\mu} |\mathrm{D}_k^h (u_* - a)|^2 \cdot |\nabla \rho|^2.
\end{aligned}$$

From the last two estimates and (3.12) we get

$$\begin{aligned}
&\mu \int_\Omega |\mathrm{D}_k^h \nabla u_*|^2 \rho^2 \,\mathrm{d}x \\
&\quad \le \int_\Omega \mathrm{D}_k^h(\mathrm{D} f(\nabla u_*)) : [\rho^2 \mathrm{D}_k^h \nabla u_*] \,\mathrm{d}x \\
&\quad = -\int_\Omega \mathrm{D}_k^h(\mathrm{D} f(\nabla u_*)) : \big[\mathrm{D}_k^h (u_* - a) \otimes \nabla(\rho^2)\big] \,\mathrm{d}x \\
&\quad \le \int_\Omega \frac{\mu}{2M^2} \rho^2 |\mathrm{D}_k^h(\mathrm{D} f(\nabla u_*))|^2 + \frac{2M^2}{\mu} |\mathrm{D}_k^h (u_* - a)|^2 \cdot |\nabla \rho|^2 \,\mathrm{d}x
\end{aligned}$$

$$\leq \int_\Omega \frac{\mu}{2}\rho^2 |\mathrm{D}_k^h \nabla u_*|^2 + \frac{2M^2}{\mu} |\mathrm{D}_k^h (u_* - a)|^2 \cdot |\nabla \rho|^2 \,\mathrm{d}x, \tag{3.13}$$

where in the last line we used the Lipschitz continuity (3.9) of $\mathrm{D}f$ to estimate

$$\begin{aligned}
\left|\mathrm{D}_k^h(\mathrm{D}f(\nabla u_*))(x)\right| &= \left|\frac{\mathrm{D}f(\nabla u_*(x+h\mathrm{e}_k)) - \mathrm{D}f(\nabla u_*(x))}{h}\right| \\
&= \left|\frac{\mathrm{D}f(\nabla u_*(x) + h\mathrm{D}_k^h \nabla u_*(x)) - \mathrm{D}f(\nabla u_*(x))}{h}\right| \\
&\leq M\frac{|\nabla u_*(x+h\mathrm{e}_k) - \nabla u_*(x)|}{h} \\
&= M|\mathrm{D}_k^h \nabla u_*(x)|.
\end{aligned}$$

Absorbing the first term on the right-hand side of (3.13) into the left-hand side and using the properties of ρ, we arrive at

$$\int_{B(x_0,r)} |\mathrm{D}_k^h \nabla u_*|^2 \,\mathrm{d}x \leq \left(\frac{2M}{\mu}\right)^2 \int_{B(x_0,2r)} \frac{|\mathrm{D}_k^h(u_* - a)|^2}{r^2} \,\mathrm{d}x.$$

Now invoke the difference-quotient lemma, part (i), to deduce that

$$\int_{B(x_0,r)} |\mathrm{D}_k^h \nabla u_*|^2 \,\mathrm{d}x \leq \left(\frac{2M}{\mu}\right)^2 \int_{B(x_0,3r)} \frac{|\partial_k(u_* - a)|^2}{r^2} \,\mathrm{d}x.$$

Applying the difference-quotient lemma again, this time part (ii), we get $u_* \in \mathrm{W}^{2,2}(B(x_0,r);\mathbb{R}^m)$. The Caccioppoli inequality (3.10) follows once we take a with $\nabla a := [\nabla u_*]_{B(x_0,3r)}$. □

The $\mathrm{W}^{2,2}_{\mathrm{loc}}$-regularity of solutions can be extended up to and including the boundary if the boundary is smooth enough, yielding a full $\mathrm{W}^{2,2}$-regularity theorem, see Section 6.3.2 in [111] and also [137].

In the special case when f is quadratic, say $\mathrm{D}^2 f(x,v,A) = \mathbf{S}$ for a symmetric fourth-order tensor $\mathbf{S}$, we can iterate, or *bootstrap*, the regularity arguments to conclude the smoothness of a minimizer. Indeed, with the $\mathrm{W}^{2,2}_{\mathrm{loc}}$-regularity result at hand, we may use $\psi = \partial_k \widetilde{\psi}$ for $\widetilde{\psi} \in \mathrm{C}_c^\infty(\Omega;\mathbb{R}^m)$, $k = 1,\ldots,d$, as test function in the weak formulation of the Euler–Lagrange equation $-\operatorname{div}[\mathrm{D}f(\nabla u_*)] = 0$ to conclude that

$$-\operatorname{div}\big[\mathbf{S}\nabla(\partial_k u_*)\big] = 0 \quad \text{in } \Omega \tag{3.14}$$

holds in the weak sense, i.e.,

$$0 = -\int_\Omega \mathrm{D}f(\nabla u_*) : \nabla(\partial_k \widetilde{\psi}) \,\mathrm{d}x = \int_\Omega \big[\mathbf{S}\nabla(\partial_k u_*)\big] : \nabla \widetilde{\psi} \,\mathrm{d}x$$

for all $\widetilde{\psi} \in \mathrm{C}_c^\infty(\Omega; \mathbb{R}^m)$. However, this PDE is itself the Euler–Lagrange equation for the integral functional

$$\mathscr{F}^{(k)}[u] := \int_\Omega \frac{1}{2} \nabla(\partial_k u(x)) : \mathbf{S}\nabla(\partial_k u(x)) \,\mathrm{d}x, \qquad u \in \mathrm{W}^{2,2}(\Omega; \mathbb{R}^m).$$

Applying Theorem 3.11, we get that $\partial_k u \in \mathrm{W}^{2,2}_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$. Iterating this procedure, we obtain the following higher-regularity result.

Corollary 3.13. *Let $\mathscr{F}$ be a quadratic regular variational integral. Then, for any minimizer $u_* \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^m)$ of $\mathscr{F}$ it holds that $u_* \in \mathrm{W}^{k,2}_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$ for all $k \in \mathbb{N}$, hence also $u_* \in \mathrm{C}^\infty(\Omega; \mathbb{R}^m)$.*

We also state a regularity result for integral functionals with a lower-order term:

Theorem 3.14. *Let $f : \mathbb{R}^{m\times d} \to \mathbb{R}$ satisfy the same assumptions as in Theorem 3.11 and let $h \in \mathrm{L}^2(\Omega; \mathbb{R}^m)$. Then, minimizers $u_* \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^m)$ of the functional*

$$\mathscr{F}[u] := \int_\Omega f(\nabla u(x)) - h(x) \cdot u(x) \,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^m),$$

lie in $\mathrm{W}^{2,2}_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$ and satisfy the strong Euler–Lagrange equation

$$-\operatorname{div} \mathrm{D} f(\nabla u_*) = h \quad \textit{a.e. in } \Omega.$$

If f is quadratic and $h \in \mathrm{C}^\infty(\Omega; \mathbb{R}^m)$, then $u_* \in \mathrm{C}^\infty(\Omega; \mathbb{R}^m)$.

The proof is the task of Problem 3.4 and an extension is in Problem 3.5.

Example 3.15. For a minimizer $u_* \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^m)$ of the Dirichlet functional as in Example 2.8, the theory presented so far immediately gives $u_* \in \mathrm{C}^\infty(\Omega; \mathbb{R}^m)$. By Theorem 3.14, the same applies to the functional

$$\mathscr{F}[u] := \int_\Omega \frac{1}{2} |\nabla u(x)|^2 - h(x) \cdot u(x) \,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^m),$$

where we assume $h \in \mathrm{C}^\infty(\Omega; \mathbb{R}^m)$. Thus, solutions $u \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^m)$ of the Laplace equation

$$-\Delta u = 0 \quad \text{in } \Omega$$

or the Poisson equation

$$-\Delta u = h \quad \text{in } \Omega$$

with $h \in \mathrm{C}^\infty(\Omega; \mathbb{R}^m)$, are smooth.

Example 3.16. Similarly, for minimizers $u_* \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^3)$ of the problem of linearized elasticity from Section 1.7 and Example 2.12, we also get $u_* \in \mathrm{C}^\infty(\Omega; \mathbb{R}^3)$.

The reasoning has to be slightly adjusted, however, to take care of the fact that we are dealing with the symmetric gradient $\mathscr{E}u$ instead of ∇u.

In the general case of regular variational integrals the Euler–Lagrange equation is *nonlinear* and the bootstrapping procedure fails. Thus, for Hilbert's 19th problem, different methods had to be developed. Full solutions were given, in the scalar case $m = 1$, by Ennio De Giorgi in 1957 [87] and, using different methods, by John F. Nash in 1958 [213]. After the proof was improved by Jürgen Moser in 1960/1961 [197, 198], the results are now collectively referred to as *De Giorgi–Nash–Moser theory*.

The standard solution (following De Giorgi's approach) is based on the *De Giorgi regularity theorem* and the classical *Schauder estimates*, which establish Hölder regularity of weak solutions of a PDE.

Theorem 3.17 **(De Giorgi 1957 [87]).** *Let $S \colon \Omega \to \mathbb{R}^{d\times d}$ be measurable, symmetric, ($S(x) = S(x)^T$ for $x \in \Omega$), and satisfy the ellipticity and boundedness estimates*

$$\mu|v|^2 \le v^T S(x)v \le M|v|^2, \qquad (x, v) \in \Omega \times \mathbb{R}^d, \tag{3.15}$$

for constants $\mu, M > 0$. If $u \in \mathrm{W}^{1,2}(\Omega)$ is a weak solution of

$$-\operatorname{div}[S\nabla u] = 0, \tag{3.16}$$

then $u \in \mathrm{C}^{0,\alpha_0}_{\mathrm{loc}}(\Omega)$, that is, u is α_0-Hölder continuous, for some $\alpha_0 = \alpha_0(d, M/\mu) \in (0, 1)$.

Theorem 3.18 **(Schauder 1934/1937 [237, 238]).** *Let $S \colon \Omega \to \mathbb{R}^{d\times d}$ be as above but in addition assumed to be of Hölder class $\mathrm{C}^{n-1,\alpha}$ for some $n \in \mathbb{N}$ and $\alpha \in (0, 1)$. If $u \in \mathrm{W}^{1,2}(\Omega)$ is a weak solution of*

$$-\operatorname{div}[S\nabla u] = 0,$$

then $u \in \mathrm{C}^{n,\alpha}_{\mathrm{loc}}(\Omega)$.

We remark that the Schauder estimates also hold for *systems* of PDEs ($m > 1$), but the De Giorgi regularity theorem does not. The proofs of these results are a bit involved (see the notes section at the end of this chapter for some pointers to the literature), but we at least establish one of their most important consequences, namely the solution of Hilbert's 19th problem in the scalar case (with smoothness instead of analyticity, however):

Theorem 3.19 **(De Giorgi 1957 & Nash 1958 & Moser 1960 [87, 197, 213]).** *Let $\mathscr{F}$ be a regular variational integral with an integrand $f \colon \mathbb{R}^{d\times d} \to \mathbb{R}$ that is n times continuously differentiable, where $n \in \{2, 3, \ldots\}$. If $u_* \in \mathrm{W}^{1,2}(\Omega)$ minimizes $\mathscr{F}$, then $u_* \in \mathrm{C}^{n-1,\alpha}_{\mathrm{loc}}(\Omega)$ for some $\alpha \in (0, 1)$. In particular, if f is smooth, then $u_* \in \mathrm{C}^\infty(\Omega)$.*

Proof We saw in (3.14) that the partial derivatives $\partial_k u_*, k = 1, \dots, d$, of a minimizer $u_* \in \mathrm{W}^{1,2}(\Omega)$ of $\mathscr{F}$ satisfy (3.16) for

$$S(x) := \mathrm{D}^2 f(\nabla u_*(x)), \qquad x \in \Omega.$$

From general properties of the Hessian we conclude that $S(x)$ is symmetric and the upper and lower estimates (3.15) on S follow from the respective properties of $\mathrm{D}^2 f$. However, we cannot conclude (yet) any regularity of S beyond measurability. Nevertheless, we may apply the De Giorgi Regularity Theorem 3.17, whereby $\partial_k u_* \in \mathrm{C}^{0,\alpha_0}_{\mathrm{loc}}(\Omega)$ for some $\alpha_0 \in (0, 1)$ and all $k = 1, \dots, d$. Hence $u_* \in \mathrm{C}^{1,\alpha_0}_{\mathrm{loc}}(\Omega)$. This is the assertion for $n = 2$.

If $n = 3$, our arguments so far in conjunction with the regularity assumptions on f imply $S(x) = \mathrm{D}^2 f(\nabla u_*(x)) \in \mathrm{C}^{0,\alpha_0}_{\mathrm{loc}}(\Omega)$. Indeed, $\mathrm{D}^2 f$ is locally Lipschitz and ∇u_* is locally α_0-Hölder continuous, hence the composite function is also locally α_0-Hölder continuous. Consequently, the Schauder estimates from Theorem 3.18 apply and yield $\partial_k u_* \in \mathrm{C}^{1,\alpha_0}_{\mathrm{loc}}(\Omega)$, whereby $u_* \in \mathrm{C}^{2,\alpha_0}_{\mathrm{loc}}(\Omega) = \mathrm{C}^{n-1,\alpha_0}_{\mathrm{loc}}(\Omega)$.

For higher n, this procedure can be iterated until we run out of f-derivatives and $u_* \in \mathrm{C}^{n-1,\alpha_0}_{\mathrm{loc}}(\Omega)$. □

A slight refinement of the above argument also yields analyticity of u_* if f is analytic. Another refinement shows that in the situation of the De Giorgi–Nash–Moser Theorem, we even have $u_* \in \mathrm{C}^{n-1,\alpha}_{\mathrm{loc}}(\Omega)$ for *all* $\alpha \in (0, 1)$. Here one needs the additional Schauder–type result that weak solutions u to $-\operatorname{div}[S\nabla u] = 0$ for $S = S(x)$ *continuous* have $\mathrm{C}^{0,\alpha}_{\mathrm{loc}}$-regularity for *all* $\alpha \in (0, 1)$. In the above proof we can apply this result at stage $n = 2$ since $S(x) = \mathrm{D}^2 f(\nabla u_*(x))$ is continuous by the De Giorgi regularity theorem. Thus, $u_* \in \mathrm{C}^{1,\alpha}_{\mathrm{loc}}(\Omega)$ for *any* $\alpha \in (0, 1)$. The other parts of the proof are adapted accordingly. See [244] for details and other refinements.

We close this section by considering the vectorial case $m > 1$. It was shown again by De Giorgi that if $d = m > 2$, then his regularity theorem does not hold:

Example 3.20 (*De Giorgi 1968 [88]*). Let $d = m > 2$ and define

$$u_*(x) := \frac{x}{|x|^\gamma}, \qquad \gamma := \frac{d}{2}\left(1 - \frac{1}{\sqrt{(2d-2)^2 + 1}}\right), \qquad x \in B(0, 1).$$

Note that $1 < \gamma < d/2$ and so $u_* \in \mathrm{W}^{1,2}(B(0, 1); \mathbb{R}^d)$ but $u_* \notin \mathrm{L}^\infty_{\mathrm{loc}}(B(0, 1); \mathbb{R}^d)$. It can be checked, though, that u_* solves (3.16) for

$$A^T \mathbf{S}(x) A := |A|^2 + \left[\left((d-2)\mathrm{Id} + d \cdot \frac{x \otimes x}{|x|^2}\right) : A\right]^2, \qquad A \in \mathbb{R}^{d\times d},$$

which satisfies all assumptions of (a vector-analogue of) the De Giorgi Regularity Theorem 3.17. Furthermore, u_* is a weak solution to the system of PDEs

$$-\operatorname{div}[\mathbf{S}\nabla u_*] = 0.$$

In fact, u_* is a minimizer of the quadratic variational integral

$$\mathscr{F}_{\mathrm{DG}}[u] := \int_{B(0,1)} \nabla u(x)^T \mathbf{S}(x) \nabla u(x) \, \mathrm{d}x, \qquad u \in \mathrm{W}^{1,2}(B(0,1); \mathbb{R}^2).$$

However, this is not a regular variational integral because the integrand depends (non-smoothly) on x.

In principle, this still leaves the possibility that the vectorial analogue of Hilbert's 19th problem has a positive solution, just that its proof would have to proceed along a different route than via the De Giorgi theorem. However, Nečas in 1975 [214] gave a (complicated) example of a regular variational integral for $d \geq 5$ that has a minimizer that is non-C^1 (but still Lipschitz).

For $d = 2$, minimizers to regular variational problems are always as regular as the data allows (as in Theorem 3.19), this is the Morrey regularity theorem from 1938, see [194, 196]. If we confine ourselves to Sobolev-regularity, then using the difference quotient technique, one can prove $\mathrm{W}^{2,2+\delta}_{\mathrm{loc}}$-regularity for minimizers of regular variational problems for some *dimension-dependent* $\delta > 0$. This result is originally due to Campanato [55]. By the Sobolev embedding theorem this yields $\mathrm{C}^{0,\alpha}$-regularity for some $\alpha \in (0, 1)$ when $d \leq 4$. In 2008 Kristensen and Melcher [166] established $\mathrm{W}^{2,2+\delta}_{\mathrm{loc}}$-regularity for a *dimension-independent* $\delta > 0$ (in fact, $\delta = \mu/(50M)$). We will discuss further regularity results for the vector case in Section 5.7.

On the negative side, the following results are known: Šverák and Yan proved in 2000–2002 [255, 256] that there exist regular variational integrals (with smooth integrands) with the following properties:

- $d \geq 3$, $m \geq 5$ or $d \geq 4$, $m \geq 3$: The minimizer is non-Lipschitz.
- $d \geq 5$, $m \geq 14$: The minimizer is unbounded.

In 2016 Mooney and Savin [191] were finally able to give a striking example that there exists a regular variational integral in dimensions $d \geq 3$, $m \geq 2$ that has a non-Lipschitz minimizer.

3.3 Lagrange Multipliers

We now continue the study of integral side constraints from Section 2.5.

Theorem 3.21 *Let $f \colon \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \to \mathbb{R}$, $h \colon \Omega \times \mathbb{R}^m \to \mathbb{R}$ be Carathéodory integrands that are continuously differentiable in v, A and that satisfy the growth bounds*

$$
\begin{aligned}
|f(x, v, A)| &\le M(1 + |v|^p + |A|^p), \\
|h(x, v)| &\le M(1 + |v|^p), \\
|D_v h(x, v)| &\le M(1 + |v|^{p-1}), \qquad (x, v, A) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times d},
\end{aligned}
$$

for some $M > 0$, $p \in [1, \infty)$. *Suppose that the map* $u_* \in W^{1,p}_g(\Omega; \mathbb{R}^m)$, *where* $g \in W^{1-1/p,p}(\partial\Omega; \mathbb{R}^m)$, *minimizes the functional*

$$
\mathscr{F}[u] := \int_\Omega f(x, u(x), \nabla u(x)) \, dx, \qquad u \in W^{1,p}_g(\Omega; \mathbb{R}^m),
$$

under the side constraint

$$
\mathscr{H}[u] := \int_\Omega h(x, u(x)) \, dx = 0.
$$

Assume furthermore that the consistency condition

$$
\delta\mathscr{H}[u_*][w] := \int_\Omega D_v h(x, u_*(x)) \cdot w(x) \, dx \neq 0 \tag{3.17}
$$

holds for at least one $w \in W^{1,p}_0(\Omega; \mathbb{R}^m)$. *Then, there exists a* ***Lagrange multiplier*** $\lambda \in \mathbb{R}$ *such that* u_* *is a weak solution of the system of PDEs*

$$
\begin{cases}
-\operatorname{div}\big[D_A f(x, u, \nabla u)\big] + D_v f(x, u, \nabla u) = \lambda D_v h(x, u) & \text{in } \Omega, \\
u = g & \text{on } \partial\Omega.
\end{cases} \tag{3.18}
$$

Proof Let $u_* \in W^{1,p}_g(\Omega; \mathbb{R}^m)$ be a minimizer of $\mathscr{F}$ under the side constraint $\mathscr{H}[u_*] = 0$. From the consistency condition (3.17) we infer that there exists a $w \in W^{1,p}_0(\Omega; \mathbb{R}^m)$ such that

$$
\delta\mathscr{H}[u_*][w] = \int_\Omega D_v h(x, u_*) \cdot w \, dx = 1.
$$

Now fix any $v \in W^{1,p}_0(\Omega; \mathbb{R}^m)$ and define for $s, t \in \mathbb{R}$,

$$
H(s, t) := \mathscr{H}[u_* + sv + tw].
$$

It is not difficult to see that H is continuously differentiable in both s and t (this uses the strong continuity of $\mathscr{H}$, see Theorem 2.13) and

$$
\begin{aligned}
\partial_s H(s, t) &= \delta\mathscr{H}[u_* + sv + tw][v], \\
\partial_t H(s, t) &= \delta\mathscr{H}[u_* + sv + tw][w].
\end{aligned}
$$

Thus, from the definition of w we infer that

$$H(0,0) = 0 \qquad \text{and} \qquad \partial_t H(0,0) = 1.$$

By the implicit function theorem there exists a continuously differentiable function $\tau : \mathbb{R} \to \mathbb{R}$ such that $\tau(0) = 0$ and

$$H(s, \tau(s)) = 0 \qquad \text{for small } |s|.$$

The chain rule yields for such s,

$$0 = \partial_s[H(s,\tau(s))] = \partial_s H(s,\tau(s)) + \partial_t H(s,\tau(s))\tau'(s),$$

whereby

$$\tau'(0) = -\partial_s H(0,0) = -\int_\Omega \mathrm{D}_v h(x,u_*) \cdot v \,\mathrm{d}x. \tag{3.19}$$

Now define for small $|s|$ as above,

$$J(s) := \mathscr{F}[u_* + sv + \tau(s)w].$$

We have

$$\mathscr{H}[u_* + sv + \tau(s)w] = H(s,\tau(s)) = 0,$$

and the continuously differentiable function J has a minimum at $s = 0$ by the minimization property of u_*. Thus, with the shorthand notations $\mathrm{D}_A f := \mathrm{D}_A f(x, u_*, \nabla u_*)$ and $\mathrm{D}_v f := \mathrm{D}_v f(x, u_*, \nabla u_*)$,

$$0 = J'(0) = \int_\Omega \mathrm{D}_A f : (\nabla v + \tau'(0)\nabla w) + \mathrm{D}_v f \cdot (v + \tau'(0)w) \,\mathrm{d}x.$$

Rearranging and using (3.19), we get

$$\begin{aligned}\int_\Omega \mathrm{D}_A f : \nabla v + \mathrm{D}_v f \cdot v \,\mathrm{d}x &= -\tau'(0) \int_\Omega \mathrm{D}_A f : \nabla w + \mathrm{D}_v f \cdot w \,\mathrm{d}x \\ &= \lambda \int_\Omega \mathrm{D}_v h(x,u_*) \cdot v \,\mathrm{d}x,\end{aligned} \tag{3.20}$$

where we have defined

$$\lambda := \int_\Omega \mathrm{D}_A f : \nabla w + \mathrm{D}_v f \cdot w \,\mathrm{d}x.$$

Since (3.20) shows that u_* is a weak solution of (3.18), the proof is finished. □

Example 3.22 (*Stationary Schrödinger equation*). When looking for ground states in quantum mechanics as in Section 1.4, we have to minimize

$$\mathscr{E}[\Psi] := \int_{\mathbb{R}^N} \frac{\hbar^2}{4\mu}|\nabla\Psi|^2 + \frac{1}{2}V(x)|\Psi|^2 \, dx$$

over all $\Psi \in W^{1,2}(\mathbb{R}^N; \mathbb{C})$ under the side constraint

$$\|\Psi\|_{L^2}^2 = \int_{\mathbb{R}^N} |\Psi|^2 \, dx = 1.$$

From Theorem 2.15 in conjunction with Lemma 2.16 (extended to also apply in the whole space $\Omega = \mathbb{R}^N$) we see that this problem always has at least one solution $\Psi_* \in W^{1,2}(\mathbb{R}^N; \mathbb{C})$. Theorem 3.21 (likewise extended to the whole space and for side constraints $\mathscr{H}[u] = \alpha \in \mathbb{R}$) yields that this Ψ_* satisfies (in the weak sense)

$$\left[\frac{-\hbar^2}{2\mu}\Delta + V(x)\right]\Psi_*(x) = E\Psi_*(x), \qquad x \in \Omega,$$

for some $E \in \mathbb{R}$ (the Lagrange multiplier is $E/2$), which is precisely the *stationary Schrödinger equation*. One can also show that $E > 0$ is the smallest eigenvalue of the operator $\Psi \mapsto [\frac{-\hbar^2}{2\mu}\Delta + V(x)]\Psi$; the proof of this fact is the task of Problem 3.10.

3.4 Invariances and Noether's Theorem

In physics and other applications of the calculus of variations, we are often interested in the *symmetries* of minimizers or, more generally, critical points. These symmetries manifest themselves in other differential or pointwise relations that are automatically satisfied for any minimizer (critical point). They can be used to identify concrete solutions or are interesting in their own right. In this section we only consider $W^{2,2}_{loc}$-minimizers (critical points), which is the natural level of regularity by Theorem 3.11.

As a first concrete example of a symmetry, consider the Dirichlet functional

$$\mathscr{F}[u] := \int_\Omega \frac{1}{2}|\nabla u(x)|^2 \, dx, \qquad u \in W^{1,2}(\Omega),$$

which was introduced in Example 2.8. First, we notice that $\mathscr{F}$ is invariant under translations in space: Let $u \in (W^{1,2} \cap W^{2,2}_{loc})(\Omega)$, $\tau \in \mathbb{R}$, $k \in \{1, \ldots, d\}$, and set

$$x_\tau := x + \tau e_k, \qquad u_\tau(x) := u(x + \tau e_k).$$

Then, for any open set $D \subset \mathbb{R}^d$ with $D_\tau := D + \tau e_k \subset \Omega$, we have the *invariance relation*

$$\int_D \frac{1}{2}|\nabla u_\tau(x)|^2 \, dx = \int_{D_\tau} \frac{1}{2}|\nabla u(x_\tau)|^2 \, dx_\tau. \tag{3.21}$$

The Dirichlet functional also exhibits an invariance with respect to *scaling*: For $\lambda > 0$ set

$$x_\lambda := \lambda x, \qquad u_\lambda(x) := \lambda^{(d-2)/2} u(\lambda x), \qquad D_\lambda := \lambda D. \tag{3.22}$$

Then it is not hard to see that (3.21) again holds if we set $\lambda = e^\tau$ (to allow $\tau \in \mathbb{R}$ as before).

The main result of this section, Noether's theorem, roughly says that "differentiable invariances of a functional give rise to conservation laws". More concretely, the two invariances of the Dirichlet functional presented above will yield two *additional* PDEs that any minimizer of the Dirichlet functional must satisfy.

To make this statement precise in the general case, we need a bit of notation: Let $u_* \in (\mathrm{W}^{1,2} \cap \mathrm{W}^{2,2}_{\mathrm{loc}})(\Omega; \mathbb{R}^m)$ be a minimizer (or, more generally, a critical point) of the functional

$$\mathscr{F}[u] := \int_\Omega f(x, u(x), \nabla u(x)) \, dx,$$

where $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}$ is assumed to be continuously differentiable in the second and third arguments. Then, u_* satisfies the strong Euler–Lagrange equation

$$-\operatorname{div}\big[D_A f(x, u_*, \nabla u_*)\big] + D_v f(x, u_*, \nabla u_*) = 0 \quad \text{a.e. in } \Omega,$$

see Proposition 3.9. We consider u_* to be extended to all of $\mathbb{R}^d$ (it will not matter below how we extend u_*).

The invariance is specified through maps $g : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ and $H : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^m$, which will depend on u_* above, with

$$g(x, 0) = x \qquad \text{and} \qquad H(x, 0) = u_*(x), \qquad x \in \mathbb{R}^d.$$

We also require that g, H are *continuously differentiable* in their second argument for almost every $x \in \Omega$. Then set for $x \in \mathbb{R}^d$, $\tau \in \mathbb{R}$ and any open set $D \Subset \mathbb{R}^d$,

$$x_\tau := g(x, \tau), \qquad u_\tau(x) := H(x, \tau), \qquad D_\tau := g(D, \tau).$$

One can think of the transformation (g, H) as a form of *homotopy*. We call $\mathscr{F}$ **invariant** under the transformation defined by (g, H) if

$$\int_D f(x, u_\tau(x), \nabla u_\tau(x)) \, dx = \int_{D_\tau} f(x', u_*(x'), \nabla u_*(x')) \, dx' \tag{3.23}$$

for all Lipschitz subdomains $D \subset \mathbb{R}^d$ and for all $\tau \in \mathbb{R}$ sufficiently small such that $D_\tau \Subset \Omega$.

The following result goes back to Emmy Noether and is considered to be one of the most important mathematical theorems ever proved. Its pivotal idea of systematically generating conservation laws from invariances has shown itself to be immensely influential in modern physics.

Theorem 3.23 **(Noether 1918 [217]).** *Let* $f\colon \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}$ *be twice continuously differentiable in* v, A *and satisfy the growth bounds*

$$|\mathrm{D}_v f(x,v,A)|, |\mathrm{D}_A f(x,v,A)| \le C(1+|v|^p+|A|^p), \qquad (x,v,A)\in\Omega\times\mathbb{R}^m\times\mathbb{R}^{m\times d},$$

for some $C>0$, $p\in[1,\infty)$. *Further, let the associated functional* $\mathscr{F}$ *be invariant under the transformation defined by* (g,H) *as above, and assume that there exists a majorant* $h\in\mathrm{L}^p(\Omega)$ *such that*

$$|\partial_\tau H(x,\tau)|, |\partial_\tau g(x,\tau)| \le h(x) \quad \text{for a.e. } x\in\Omega \text{ and all } \tau\in\mathbb{R}. \tag{3.24}$$

Then, for any minimizer or critical point $u_*\in(\mathrm{W}^{1,2}\cap\mathrm{W}^{2,2}_{\mathrm{loc}})(\Omega;\mathbb{R}^m)$ *of the functional* $\mathscr{F}$*, the conservation law*

$$\operatorname{div}\bigl[\mu^T\mathrm{D}_A f(x,u_*,\nabla u_*) - \nu f(x,u_*,\nabla u_*)\bigr] = 0 \quad \text{a.e. in } \Omega \tag{3.25}$$

holds, where

$$\mu(x) := \partial_\tau H(x,0)\in\mathbb{R}^m \qquad \text{and} \qquad \nu(x) := \partial_\tau g(x,0)\in\mathbb{R}^d, \qquad x\in\Omega,$$

are the ***Noether multipliers****.*

Corollary 3.24. *If* $f = f(v,A)\colon \mathbb{R}\times\mathbb{R}^d\to\mathbb{R}$ *does not depend on* x*, then every minimizer (or critical point)* $u_*\in(\mathrm{W}^{1,2}\cap\mathrm{W}^{2,2}_{\mathrm{loc}})(\Omega)$ *of the corresponding functional* $\mathscr{F}$ *satisfies*

$$\sum_{i=1}^d \partial_i\bigl[(\partial_k u)\cdot\partial_{A_k} f(u_*,\nabla u_*) - \delta_{ik} f(u_*,\nabla u_*)\bigr] = 0 \quad \text{a.e. in } \Omega \tag{3.26}$$

for all $k=1,\ldots,d$.

Here,

$$\delta_{ik} := \begin{cases} 1 & \text{if } i=k,\\ 0 & \text{otherwise,}\end{cases}$$

is the **Kronecker delta**.

Proof of Theorem 3.23 and Corollary 3.24. We will differentiate (3.23) at the minimizer (critical point) u_* with respect to τ. To be able to differentiate the left-hand side under the integral, we use the growth assumptions on the derivatives

$\mathrm{D}_v f(x, v, A)$, $\mathrm{D}_A f(x, v, A)$ and (3.24) to get a uniform (in τ) majorant, which allows us to move the differentiation under the integral sign. For the right-hand side, we need to employ the formula for the differentiation of an integral with respect to a moving domain (this is a special case of the *Reynolds transport theorem*), namely

$$\frac{\mathrm{d}}{\mathrm{d}\tau} \int_{D_\tau} f(x, u_*, \nabla u_*) \,\mathrm{d}x = -\int_{\partial D_\tau} f(x, u_*, \nabla u_*) \partial_\tau g(x, \tau) \cdot n \,\mathrm{d}\mathscr{H}^{d-1},$$

where n is the unit inner normal on ∂D_τ and $\mathscr{H}^{d-1}$ is the $(d-1)$-dimensional surface (Hausdorff) measure on ∂D_τ; this formula can be checked in an elementary way using the transformation formula for integrals under coordinate changes.

Abbreviating for readability

$$\begin{aligned} F &:= f(x, u_*(x), \nabla u_*(x)), \\ \mathrm{D}_A F &:= \mathrm{D}_A f(x, u_*(x), \nabla u_*(x)), \\ \mathrm{D}_v F &:= \mathrm{D}_v f(x, u_*(x), \nabla u_*(x)), \end{aligned}$$

we get as the result of the differentiation of (3.23) with respect to τ and then setting $\tau = 0$ that

$$\int_D \mathrm{D}_A F : \nabla \mu + \mathrm{D}_v F \cdot \mu \,\mathrm{d}x = -\int_{\partial D} F \nu \cdot n \,\mathrm{d}\mathscr{H}^{d-1}.$$

Next, we apply the Gauss–Green theorem to obtain

$$\begin{aligned} \int_D \big[-\operatorname{div} \mathrm{D}_A F + \mathrm{D}_v F\big] \cdot \mu \,\mathrm{d}x &= \int_{\partial D} \big[\mu^T \mathrm{D}_A F - \nu F\big] \cdot n \,\mathrm{d}\mathscr{H}^{d-1} \\ &= -\int_D \operatorname{div}\big[\mu^T \mathrm{D}_A F - \nu F\big] \,\mathrm{d}x. \end{aligned}$$

The Euler–Lagrange equation $-\operatorname{div} \mathrm{D}_A F + \mathrm{D}_v F = 0$, which holds strongly for minimizers and critical points of regularity $\mathrm{W}^{1,2} \cap \mathrm{W}^{2,2}_{\mathrm{loc}}$ (see Proposition 3.9), then yields

$$\int_D \operatorname{div}\big[\mu^T \mathrm{D}_A F - \nu F\big] \,\mathrm{d}x = 0,$$

and varying D we conclude that (3.25) holds.

The corollary follows by considering the invariance

$$x_\tau := x + \tau \mathrm{e}_k, \qquad u_\tau(x) := u_*(x + \tau \mathrm{e}_k)$$

for $k = 1, \ldots, d$, and a computation. □

Example 3.25. In the brachistochrone problem presented in Section 1.1, we were tasked to minimize the functional

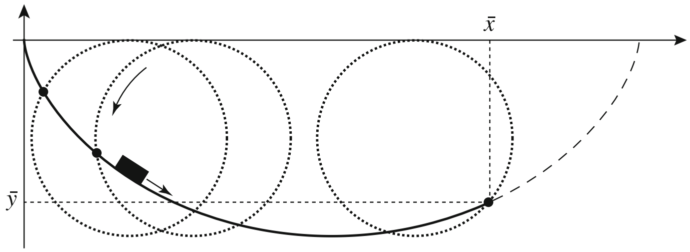


Fig. 3.2 The brachistochrone curve (here, $\bar{x} = 1$)

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

over all curves $y\colon [0,1] \to \mathbb{R}$ with $y(0) = 0$, $y(1) = \bar{y} < 0$. The integrand $f(v,a) = \sqrt{-(1+a^2)/v}$ is independent of x, hence from (3.26) we get

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

for some $r > 0$ (positive constants lead to inadmissible y). So,

$$\frac{(y')^2}{\sqrt{1+(y')^2}\cdot\sqrt{-y}} - \frac{\sqrt{1+(y')^2}}{\sqrt{-y}} = -\frac{1}{\sqrt{2r}},$$

which we transform into

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

The solution of this differential equation is called an **inverted cycloid** with radius $r > 0$. It is the curve traced out by a fixed point on a circle of radius r that touches the y-axis at the beginning and then rolls to the right on the bottom of the x-axis, see Figure 3.2. It can be written in parametric form as

$$\begin{aligned} x(t) &= r(t - \sin t), \\ y(t) &= -r(1 - \cos t), \end{aligned} \qquad t \in \mathbb{R}.$$

The radius $r > 0$ has to be chosen to satisfy the two boundary conditions on one cycloid segment.

Of course, we have not properly shown that this is the (unique) solution of the brachistochrone problem for technical reasons (e.g. growth conditions), but this can indeed be proved rigorously.

Example 3.26. We return to our canonical example, the Dirichlet functional, see Example 2.8. We know from Example 3.15 that minimizers u_* are smooth. At the beginning of this section we remarked that the Dirichlet functional is invariant under the scaling transformation (3.22) with $\lambda = e^\tau$. By Noether's theorem, this yields the (non-obvious) conservation law

$$\operatorname{div}\Big[(2\nabla u_*(x) \cdot x + (d-2)u)\nabla u_*(x) - |\nabla u_*|^2 x\Big] = 0.$$

Integrating this over $B(x_0, r) \subset \Omega$, $r > 0$, and using the Gauss–Green theorem, we get

$$(d-2)\int_{B(x_0,r)} |\nabla u_*(x)|^2 \,\mathrm{d}x = r\int_{\partial B(x_0,r)} |\nabla u_*(x)|^2 - 2\bigg(\nabla u_*(x) \cdot \frac{x}{|x|}\bigg)^2 \,\mathrm{d}\mathscr{H}^{d-1}$$

and then, after some computations,

$$\frac{\mathrm{d}}{\mathrm{d}r}\bigg(\frac{1}{r^{d-2}}\int_{B(x_0,r)} |\nabla u_*(x)|^2 \,\mathrm{d}x\bigg) = \frac{2}{r^{d-2}}\int_{\partial B(x_0,r)} \bigg(\nabla u_*(x) \cdot \frac{x}{|x|}\bigg)^2 \,\mathrm{d}\mathscr{H}^{d-1} \geq 0.$$

This *monotonicity formula* implies that

$$\frac{1}{r^{d-2}}\int_{B(x_0,r)} |\nabla u_*(x)|^2 \,\mathrm{d}x \quad \text{is increasing in } r > 0.$$

Any harmonic map ($-\Delta u = 0$) that is defined on all of $\mathbb{R}^d$ satisfies this monotonicity formula. For example, this allows us to draw the conclusion that if $d \geq 3$, then u_* cannot be compactly supported. The formula also shows that the growth around a singularity has to behave "more smoothly" than $|x|^{-2}$ (in fact, we already know that solutions are smooth). While these are not particularly strong remarks (in fact, it can be shown that $r \mapsto r^{-d} \int_{B(x_0,r)} |\nabla u_*(x)|^2 \,\mathrm{d}x$ is also increasing), they serve to illustrate how Noether's theorem restricts the candidates for solutions. Problem 3.8 exploits another invariance of the Dirichlet integral.

The last examples exhibited conservation laws that were not obvious from the Euler–Lagrange equation. While in principle they could have been derived directly, Noether's theorem gave us a *systematic* way to find these conservation laws from invariances.

3.5 Subdifferentials

Common to all the results presented in this chapter so far was that they needed some form of differentiability assumption on the functional. For *convex* functionals one can relax these differentiability assumptions by replacing differentials by affine functions that support the functional's graph.

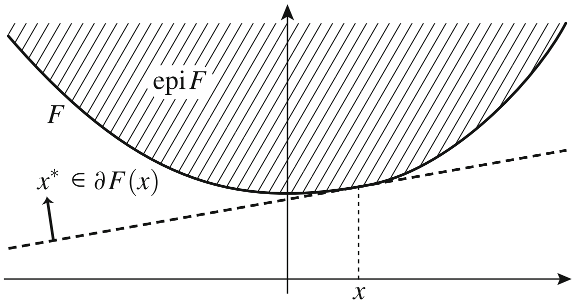


Fig. 3.3 The subdifferential

The fundamental definition is the following: Let X be a reflexive Banach space. The **subdifferential** of a proper function $F\colon X \to \mathbb{R} \cup \{+\infty\}$ at $x \in X$ is the set-valued map $\partial F\colon X \rightrightarrows X^*$, i.e., $\partial F(x) \subset X^*$ for all $x \in X$, defined by

$$\partial F(x) := \Big\{ x^* \in X^* \,:\, F(x) + \langle y - x, x^* \rangle \le F(y) \text{ for all } y \in X \Big\}, \qquad x \in X.$$

Any element x^* of $\partial F(x)$ is called a **subgradient** of F at x. The geometric intuition is that for all $x^* \in \partial F(x)$ the graph of the affine function $y \mapsto F(x) + \langle y - x, x^* \rangle$ lies everywhere below the interior of epi F and touches the graph of F at $(x, F(x))$, see Figure 3.3.

Example 3.27. Let $F(t) := |t|$, $t \in \mathbb{R}$, which is not differentiable at $t = 0$. Then,

$$\partial F(t) = \mathrm{Sgn}(t) = \begin{cases} \{-1\} & \text{if } t < 0, \\ [-1, 1] & \text{if } t = 0, \\ \{+1\} & \text{if } t > 0, \end{cases} \qquad t \in \mathbb{R}.$$

The function Sgn is called the **(multi-valued) signum function**.

Example 3.28. For the characteristic function χ_K of a compact convex set $K \subset \mathbb{R}^d$ we get

$$\partial \chi_K(x) = N_K(x) = \begin{cases} \{0\} & \text{if } x \in K \setminus \partial K, \\ \Big\{ x^* \in \mathbb{R}^d \,:\, \langle y - x, x^* \rangle \le 0 \text{ for } y \in K \Big\} & \text{if } x \in \partial K, \\ \emptyset & \text{if } x \notin K. \end{cases}$$

The set-valued function N_K is called the **normal cone** to K at x. One can verify its geometric meaning from the definition.

We first collect several properties of the subdifferential.

Proposition 3.29. *Let $F : X \to \mathbb{R} \cup \{+\infty\}$ be proper and convex and let $x \in X$.*

(i) $\partial F(x)$ *is convex and closed.*
(ii) $\partial F(x) \neq \emptyset$ *if F is finite and continuous at x.*
*(iii) If F is **Gâteaux-differentiable** at x, that is, there exists an $F'(x) \in X^*$ such that*

$$\lim_{h \downarrow 0} \frac{F(x + hv) - F(x)}{h} = \langle v, F'(x) \rangle \qquad \text{for all } v \in X,$$

then $\partial F(x) = \{F'(x)\}$.
(iv) If $\partial F(x) \neq \emptyset$, then $F(x) = F^{**}(x)$.
(v) If $F(x) = F^{**}(x)$, then $\partial F(x) = \partial F^{**}(x)$.

Proof. Ad (i): We will show in the proof of Theorem 3.32 below that

$$\partial F(x) = \left\{ x^* \in X^* \ : \ F^*(x^*) - \langle x, x^* \rangle \leq -F(x) \right\}.$$

The assertion then follows from the fact that the left-hand side of the inequality is convex and lower semicontinuous in x^*, see Proposition 2.21 (i).

Ad (ii): If F is finite and continuous at any point, then the interior of epi F is non-empty. By the Hahn–Banach Separation Theorem A.1, we can therefore find a supporting hyperplane to the interior of epi F at x, which is the graph of an affine function

$$a(y) = \langle y - x, x^* \rangle + F(x)$$

for some $x^* \in X^*$. In particular, $a \leq F$ and thus $x^* \in \partial F(x)$.

Ad (iii): We know from (ii) that there exists an $x^* \in X^*$ such that

$$F(x) + \langle y - x, x^* \rangle \leq F(y) \qquad \text{for all } y \in X.$$

Using $y = x + hv$ with $v \in X$, $h > 0$ yields

$$\frac{F(x + hv) - F(x)}{h} \geq \langle v, x^* \rangle.$$

Letting $h \downarrow 0$ and using the Gâteaux-differentiability, $\langle v, F'(x) \rangle \geq \langle v, x^* \rangle$. Applying this argument with $\pm v$, we therefore arrive at $x^* = F'(x)$.

Ad (iv), (v): These are just computations. □

The subdifferential behaves analogously to the classical differential in minimization problems:

Theorem 3.30. *Let $F : X \to \mathbb{R} \cup \{+\infty\}$ be proper and convex. Then, $x \in X$ is a minimizer for F if and only if $0 \in \partial F(x)$.*

Proof. By the definition of the subdifferential, $0 \in \partial F(x)$ is equivalent to $F(y) \geq F(x)$ for all $y \in X$. □

Example 3.31. The preceding proposition allows us to find a replacement for the Euler–Lagrange equation in the non-differentiable case. Consider the functional

$$\mathscr{F}[u] := \int_{B(0,1)} \frac{1}{4}\big(|\nabla u(x)|^2 - 1\big)^2 + |u(x)| \,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,4}(B(0,1)).$$

Then, we get that $u_* \in \mathrm{W}^{1,4}(B(0,1))$ minimizes $\mathscr{F}$ (under some boundary conditions) if and only if

$$-\operatorname{div}\big[(|\nabla u_*|^2 - 1)\nabla u_*\big] + \operatorname{Sgn}(u_*) \ni 0$$

holds in a suitable weak sense, namely as the (weak) **variational inequality**

$$\int_{B(0,1)} |u_*| \,\mathrm{d}x \le \int_{B(0,1)} |u_* + \psi(x)| + \big[(|\nabla u_*|^2 - 1)\nabla u_*\big] \cdot \nabla\psi \,\mathrm{d}x$$

for all $\psi \in \mathrm{W}_0^{1,4}(B(0,1))$. To see this, one can proceed as in the proof of Theorem 3.1 and additionally use the convexity estimate

$$|u_*(x) + \psi(x)| - |u_*(x)| \le \frac{|u_*(x) + h\psi(x)| - |u_*(x)|}{h} \qquad \text{for all } h \in (0,1].$$

A key property of the subdifferential is that it interacts well with the Legendre–Fenchel conjugate. In particular, equality in the Fenchel inequality (2.6) characterizes subgradients.

Theorem 3.32. *Let $F\colon X \to \mathbb{R} \cup \{+\infty\}$ be proper, lower semicontinuous, and convex, and let $F^*\colon X^* \to \mathbb{R} \cup \{+\infty\}$ be its conjugate. Then, the following are equivalent for $x \in X$, $x^* \in X^*$:*

(i) $x^* \in \partial F(x)$.
(ii) $x \in \partial F^*(x^*)$.
(iii) $\langle x, x^*\rangle = F(x) + F^*(x^*)$.

Proof. (i) ⇔ (iii): We already know "≤" in (iii), this is the Fenchel inequality (2.6). Thus, we only need to show the equivalence of (i) with the direction "≥" of (iii). By definition, $x^* \in \partial F(x)$ is equivalent to

$$\langle x, x^*\rangle - F(x) \ge \langle y, x^*\rangle - F(y) \qquad \text{for all } y \in X.$$

Taking the supremum over all $y \in X$ on the right-hand side, we get that this is further equivalent to

$$\langle x, x^*\rangle - F(x) \ge F^*(x^*),$$

but this is the inequality "≥" in (iii).

(ii) ⇔ (iii): This follows in the same way once we recognize that $F = F^{**}$ by Proposition 2.28 as F is convex and lower semicontinuous. □

Example 3.33. Using this theorem, the differential inclusion (here understood pointwise)

$$\operatorname{div}\bigl[(|\nabla u_*(x)|^2 - 1)\nabla u_*(x)\bigr] \in \operatorname{Sgn}(u_*(x)) \quad \text{a.e. in } B(0,1)$$

from Example 3.31 can now equivalently be written in the *dual form*

$$u_*(x) \in \partial\chi_{[-1,1]}\Bigl(\operatorname{div}\bigl[(|\nabla u_*(x)|^2 - 1)\nabla u_*(x)\bigr]\Bigr) \quad \text{a.e. in } B(0,1),$$

where $|\bullet|^* = \chi_{[-1,1]}$ is as in (2.7) from Example 2.23.

Notes and Historical Remarks

Difference quotients were already considered by Newton. Their application to regularity theory is due to work by Nirenberg in the 1940s and 1950s. Many of the fundamental results of regularity theory (albeit in a non-variational context) can be found in [136]. The books by Giusti [137] and Giaquinta–Martinazzi [133] contain theory relevant for variational questions, whereas [177] treats many questions of "fine regularity" (e.g. pointwise properties of solutions). A recent, very accessible introduction to regularity theory for PDEs is [36], also see the survey [188], which focuses on the calculus of variations. A nice framework for Schauder estimates is [244].

The Fundamental Lemma 3.10 of the calculus of variations is due to Paul Du Bois-Reymond and is sometimes named after him.

Noether's theorem has many ramifications and can be put into a very general form in Hamiltonian systems and Lie group theory. The idea is to study groups of symmetries and their actions. For an introduction to this diverse field, see [218] and also [278]. Example 3.26 about the monotonicity formula is from [111]. For more on Lagrange multipliers and Noether's theorem, see [131, 132].

Subdifferentials were introduced in the general work of Jean-Jacques Moreau and R. Tyrrell Rockafellar on convex analysis. There are also extended subdifferentials for non-convex functions; see the monographs [233] and [192, 193] for more on this.

Problems

3.1. Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain.

(i) Compute the Euler–Lagrange equation, in weak form, for a minimizer $u \in \mathrm{W}^{1,2}(\Omega)$ of the functional

$$\mathscr{F}[u] := \int_\Omega \frac{1}{2}\nabla u(x) S(x) \nabla u(x)^T - g(x)u(x)\,\mathrm{d}x,$$

where $S\colon \Omega \to \mathbb{R}^{d\times d}$, $g\colon \Omega \to \mathbb{R}$ are continuous and $S(x) = S(x)^T$ for all $x \in \Omega$.

(ii) Assume now that additionally $S(x)$ is continuously differentiable in x and that $u \in \mathrm{W}^{2,2}(\Omega)$ is a minimizer of $\mathscr{F}$ as above. State the strong Euler–Lagrange equation for u and prove that it follows from the weak version.

3.2. Let $\Omega \subset \mathbb{R}^2$ and let $u \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^2)$. In the three cases

(i) $f(A) = A^i_j$ for some $i, j \in \{1, 2\}$,
(ii) $f(A) = \det A$,
(iii) $f(A) = (\operatorname{cof} A)^i_j$ for some $i, j \in \{1, 2\}$

show through a direct calculation that

$$-\operatorname{div}[\mathrm{D}_A f(\nabla u)] = 0 \tag{3.27}$$

for *all* $u \in \mathrm{C}^2(\Omega; \mathbb{R}^2)$. This shows that these f are *null-Lagrangians*, i.e., the Euler–Lagrange equation holds for all u.

3.3. Show that also for $d \geq 3$ the above Euler–Lagrange equation (3.27) continues to hold for all $(r \times r)$-minors $f(A) = M(A)$, that is, $M(A)$ is the determinant of a selection of r rows and r columns of A. *Hint:* You can assume that you select the *first* r rows and columns, thus considering only the *principal* minors.

3.4. Prove Theorem 3.14, namely that for functionals of the form

$$\mathscr{F}[u] := \int_\Omega f(\nabla u(x)) - h(x) \cdot u(x) \,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^m),$$

where f satisfies the same assumptions as in Theorem 3.11 and $h \in \mathrm{L}^2(\Omega; \mathbb{R}^m)$, the minimizer u_* has $\mathrm{W}^{2,2}_{\mathrm{loc}}$-regularity. Show furthermore that if f is quadratic and h is smooth, then $u_* \in \mathrm{C}^\infty(\Omega; \mathbb{R}^m)$.

3.5. Prove an analogue of Theorem 3.11 for functionals of the form

$$\mathscr{F}[u] := \int_\Omega f(\nabla u(x)) - H(u(x)) \,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,2}(\Omega),$$

where f satisfies the same assumptions as in Theorem 3.11 and $H\colon \mathbb{R}^m \to \mathbb{R}$ is continuously differentiable with $|\mathrm{D}H(v)| \leq C(1 + |v|)$ for some $C > 0$ and all $v \in \mathbb{R}^m$. Is it possible to also allow the weaker growth bound $|\mathrm{D}H(v)| \leq C(1+|v|^r)$ for some $r > 1$?

3.6. Consider the minimization problem for the problem of linearized elasticity on $\Omega \subset \mathbb{R}^3$,

$$\begin{cases} \text{Minimize} & \mathscr{F}[u] := \displaystyle\int_\Omega \mu|\mathscr{E}u(x)|^2 + \frac{1}{2}\Big(\kappa - \frac{2}{3}\mu\Big)|\operatorname{tr}\mathscr{E}u(x)|^2 - b(x)\cdot u(x)\,\mathrm{d}x \\ \text{over all} & u \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^3) \text{ with } u|_{\partial\Omega} = g, \end{cases}$$

where $\mu, \kappa > 0$ are such that $\kappa - \frac{2}{3}\mu \geq 0$, $f \in \mathrm{L}^\infty(\Omega; \mathbb{R}^3)$, $g \in \mathrm{W}^{1/2,2}(\partial\Omega; \mathbb{R}^3)$, and $\mathscr{E}u(x) := (\nabla u(x) + \nabla u(x)^T)/2$. Prove that the Euler–Lagrange equation (satisfied by a minimizer in a weak sense) is

$$\begin{cases} -\operatorname{div}\Big[2\mu\,\mathscr{E}u + \Big(\kappa - \dfrac{2}{3}\mu\Big)(\operatorname{tr}\mathscr{E}u)I\Big] = b & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases}$$

3.7. In the situation of the previous problem, assume $\kappa - \frac{2}{3}\mu = 0$, $b = 0$, and that $u \in (\mathrm{W}^{1,2} \cap \mathrm{W}^{2,2}_{\mathrm{loc}})(\Omega; \mathbb{R}^3)$ is a minimizer of $\mathscr{F}$ as above. Show, using a suitable Noether symmetry, that for all skew-symmetric $W \in \mathbb{R}^{3\times 3}$ ($W^T = -W$) it holds that

$$\operatorname{div}[x^T W^T \mathscr{E}u(x)] = 0 \qquad \text{for a.e. } x \in \Omega.$$

3.8. Set for a skew-symmetric $W \in \mathbb{R}^{d\times d}$

$$x_\tau = g(x,\tau) := \exp(\tau W)x, \qquad u_\tau = H(x,\tau) := u(\exp(\tau W)x), \qquad \tau \in \mathbb{R}.$$

Show that the Dirichlet functional is invariant under the rotational transformation defined by (g, H) and conclude that any minimizer $u_* \in (\mathrm{W}^{1,2} \cap \mathrm{W}^{2,2}_{\mathrm{loc}})(\Omega; \mathbb{R}^m)$ of the Dirichlet functional (for given boundary values) satisfies the conservation law

$$\operatorname{div}\big[x^T W |\nabla u_*(x)|^2\big] = 0.$$

3.9. In the situation of Exercise 2.2, derive the weak Euler–Lagrange equation. *Hint:* Think about the class of "test variations" ψ that you need to allow.

3.10. In the situation of Example 3.22, show that $E > 0$ is the smallest eigenvalue of the operator $\Psi \mapsto [\frac{-\hbar^2}{2\mu}\Delta + V(x)]\Psi$.

Chapter 4
Young Measures

Before we continue our study of integral functionals, we first introduce an abstract, yet very versatile, tool, the *Young measure*, named after its inventor Laurence C. Young. Young measures pervade much of the modern theory of the calculus of variations and will be used throughout the remainder of the book. In the next chapter, we will see their first use in the proof of the lower semicontinuity theorem for integral functionals with *quasiconvex* integrands.

Let us motivate this device through the following fundamental question: Assume that we are given a weakly converging sequence $v_j \rightharpoonup v$ in $\mathrm{L}^2(\Omega)$, where $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain, and an integral functional

$$\mathscr{F}[w] := \int_\Omega f(x, w(x)) \,\mathrm{d}x, \qquad w \in \mathrm{L}^2(\Omega),$$

with $f : \Omega \times \mathbb{R} \to \mathbb{R}$ continuous and bounded (for simplicity). Then, $(\mathscr{F}[v_j])_j$ is a bounded sequence and up to a (non-renumbered) subsequence we may assume that $\mathscr{F}[v_j]$ converges to some limit as $j \to \infty$. The question then arises: how can we compute this limit for *every* integrand f as above?

Equivalently, we could ask for the weak* limit in $\mathrm{L}^\infty(\Omega)$ of the sequence of compound functions $F_j(x) := f(x, v_j(x))$. It is easy to see that this weak* limit in general is *not* equal to $f(x, v(x))$. For example, in $\Omega = (0, 1)$, consider the oscillating sequence

$$v_j(x) := \begin{cases} a & \text{if } jx - \lfloor jx \rfloor \in [0, \theta), \\ b & \text{if } jx - \lfloor jx \rfloor \in [\theta, 1), \end{cases} \qquad x \in (0, 1),$$

where $a, b \in \mathbb{R}$ with $a \neq b, \theta \in (0, 1)$. Then, if $f(x, a) = \alpha \in \mathbb{R}$, $f(x, b) = \beta \in \mathbb{R}$, and f is smooth and bounded, we see immediately that $v_j \overset{*}{\rightharpoonup} \theta a + (1 - \theta)b$ and

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8_4

$$f(x, v_j) \overset{*}{\rightharpoonup} \theta\alpha + (1-\theta)\beta =: F(x).$$

Of course, the right-hand side is in general not equal to $f(x, \theta a + (1-\theta)b)$. However, we can write

$$F(x) = \langle f(x, \cdot), \nu_x \rangle = \int f(x, v) \,\mathrm{d}\nu_x(v)$$

with an x-parametrized family of probability measures $\nu_x \in \mathscr{M}^1(\mathbb{R})$ (see Appendices A.3, A.4 for some basic facts of measure theory), namely

$$\nu_x = \theta\delta_a + (1-\theta)\delta_b, \qquad x \in (0, 1).$$

This family $(\nu_x)_{x\in\Omega}$ reflects the asymptotic distribution of values in the sequence (v_j) and will be called the *Young measure* generated by the sequence (v_j).

After laying the groundwork for the theory of Young measures, in this chapter we will in particular focus on the properties of *gradient* Young measures, that is, those Young measures where the generating sequence (the (v_j) above) consists entirely of gradients. This class of Young measures is the most relevant for the calculus of variations.

4.1 The Fundamental Theorem

We start with a result that nowadays is widely known as the *Fundamental Theorem* of Young measure theory:

Theorem 4.1 (Young 1937–1942 [280–282]). *Let $(V_j) \subset \mathrm{L}^p(\Omega; \mathbb{R}^N)$ be a norm-bounded sequence, where $p \in [1, \infty]$. Then, there exists a subsequence (not explicitly labeled) 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 the (sub)sequence** (V_j), such that the following assertions are true:*

*(i) The family $(\nu_x)_x$ is **weakly* measurable**, that is, for all Carathéodory integrands $f: \Omega \times \mathbb{R}^N \to \mathbb{R}$, the compound function*

$$x \mapsto \langle f(x, \cdot), \nu_x \rangle := \int f(x, A) \,\mathrm{d}\nu_x(A), \qquad x \in \Omega,$$

is Lebesgue-measurable.

(ii) If $p \in [1, \infty)$, it holds that

$$\int_\Omega \int |A|^p \,\mathrm{d}\nu_x(A) \,\mathrm{d}x < \infty,$$

or, if $p = \infty$, there exists a compact set $K \subset \mathbb{R}^N$ such that

$$\operatorname{supp} \nu_x \subset K \qquad \textit{for a.e. } x \in \Omega.$$

(iii) For all Carathéodory integrands $f\colon \Omega \times \mathbb{R}^N \to \mathbb{R}$ with the property that the family $(f(x, V_j))_j$ is uniformly L^1*-bounded and equiintegrable, it holds that*

$$f(x, V_j) \rightharpoonup \left(x \mapsto \int f(x, A) \, \mathrm{d}\nu_x(A) \right) \quad \textit{in } \mathrm{L}^1. \tag{4.1}$$

For parametrized measures $\nu = (\nu_x)_{x \in \Omega}$ that satisfy (i) and (ii) above, we write $\nu = (\nu_x)_x \in \mathbf{Y}^p(\Omega; \mathbb{R}^N)$. If for the target dimension we have $N = 1$, then we simply write $\mathbf{Y}^p(\Omega)$ instead of $\mathbf{Y}^p(\Omega; \mathbb{R})$. See Problem 4.3 for the reason why in the definition of $\mathbf{Y}^p(\Omega; \mathbb{R}^N)$ we do not need to include (iii).

The generation of a Young measure ν by a sequence (V_j), i.e., the validity of (iii) for this sequence (V_j), will be written symbolically as

$$V_j \overset{\mathbf{Y}}{\to} \nu.$$

We refer to the explanation after Vitali's Convergence Theorem A.11 for several equivalent ways to express equiintegrability. Additionally, recall from the Dunford–Pettis Theorem A.12 that $(f(x, V_j))_j$ is equiintegrable if and only if it is weakly precompact in $\mathrm{L}^1(\Omega)$. Absorbing a test function for weak convergence into f, the convergence (4.1) can equivalently be expressed as

$$\int_\Omega f(x, V_j(x)) \, \mathrm{d}x \to \int_\Omega \int f(x, A) \, \mathrm{d}\nu_x(A) \, \mathrm{d}x = \int_\Omega \bigl\langle f(x, \cdot), \nu_x \bigr\rangle \, \mathrm{d}x =: \langle\!\langle f, \nu \rangle\!\rangle$$

for all Carathéodory integrands $f\colon \Omega \times \mathbb{R}^N \to \mathbb{R}$ such that the family $(f(x, V_j))_j$ is uniformly L^1-bounded and equiintegrable. We call $\langle\!\langle f, \nu \rangle\!\rangle$ the **duality pairing** between f and ν.

For $\nu = (\nu_x)_x \in \mathbf{Y}^p(\Omega; \mathbb{R}^N)$ the **barycenter** $[\nu] \in \mathrm{L}^p(\Omega; \mathbb{R}^N)$ of ν is defined via

$$[\nu](x) := [\nu_x] := \bigl\langle \mathrm{id}, \nu_x \bigr\rangle = \int A \, \mathrm{d}\nu_x(A), \qquad x \in \Omega.$$

Remark 4.2. The boundedness assumption on the generating sequence (V_j) in the Fundamental Theorem can be weakened: For the existence of a Young measure we only need to require the **tightness condition**

$$\lim_{h \uparrow \infty} \sup_{j \in \mathbb{N}} |\{|V_j| \geq h\}| = 0,$$

which, for example, follows from $\sup_j \int_\Omega |V_j|^r \, \mathrm{d}x < \infty$ for some $r > 0$. Of course, in this case, statement (ii) needs to be suitably adapted. The proof of the statements in this remark is the task of Problem 4.2.

For the proof of the Fundamental Theorem, we first associate with each V_j an **elementary Young measure** $\delta[V_j] = (\delta[V_j]_x)_{x\in\Omega} \in \mathbf{Y}^p(\Omega;\mathbb{R}^N)$ via

$$\delta[V_j]_x := \delta_{V_j(x)}, \qquad x \in \Omega, \tag{4.2}$$

where δ_v denotes the **Dirac mass** at $v \in \mathbb{R}^N$, that is $\delta_v(B) = 1$ if and only if $v \in B$ for any Borel set $B \subset \Omega$. Clearly, $\delta[V_j]_x$ is only defined up to a $\mathscr{L}^d$-negligible set of x's. In fact, we will (implicitly) consider all Young measures to be defined only up to $\mathscr{L}^d$-negligible sets.

On our road to proving the Fundamental Theorem, we will first show the following *Young measure compactness principle*.

Lemma 4.3. *Let $p \in [1,\infty]$ and let $(\nu^{(j)})_j \subset \mathbf{Y}^p(\Omega;\mathbb{R}^N)$ be a sequence of* L^p-*Young measures. If $p \in [1,\infty)$, assume*

$$\sup_{j\in\mathbb{N}} \left\langle\!\left\langle |\cdot|^p, \nu^{(j)} \right\rangle\!\right\rangle = \sup_{j\in\mathbb{N}} \int_\Omega \int |A|^p \,\mathrm{d}\nu_x^{(j)}(A)\,\mathrm{d}x < \infty \tag{4.3}$$

or, if $p = \infty$, assume that there exists a compact set $K \subset \mathbb{R}^{m\times d}$ such that

$$\operatorname{supp} \nu_x^{(j)} \subset K \qquad \text{for a.e. } x \in \Omega \text{ and all } j \in \mathbb{N}. \tag{4.4}$$

Then, there exists a subsequence of (ν_j) (not explicitly labeled) and $\nu \in \mathbf{Y}^p(\Omega;\mathbb{R}^N)$ such that

$$\left\langle\!\left\langle f, \nu^{(j)} \right\rangle\!\right\rangle \to \left\langle\!\left\langle f, \nu \right\rangle\!\right\rangle \qquad \text{as } j \to \infty \tag{4.5}$$

for all Carathéodory integrands $f\colon \Omega\times\mathbb{R}^N \to \mathbb{R}$ for which the sequence of functions $x \mapsto \langle f(x,\bullet), \nu_x^{(j)}\rangle$ is uniformly L^1-*bounded and the equiintegrability condition*

$$\sup_{j\in\mathbb{N}} \left\langle\!\left\langle |f(x,A)| \mathbb{1}_{\{|f(x,A)|\geq h\}}, \nu^{(j)} \right\rangle\!\right\rangle \to 0 \qquad \text{as } h \to \infty \tag{4.6}$$

holds. Moreover, if $p < \infty$,

$$\left\langle\!\left\langle |\cdot|^p, \nu \right\rangle\!\right\rangle \leq \liminf_{j\to\infty} \left\langle\!\left\langle |\cdot|^p, \nu^{(j)} \right\rangle\!\right\rangle \tag{4.7}$$

or, if $p = \infty$,

$$\operatorname{supp} \nu_x \subset K \qquad \text{for a.e. } x \in \Omega. \tag{4.8}$$

We say that $\nu^{(j)}$ **converges weakly*** to ν, in symbols "$\nu^{(j)} \overset{*}{\rightharpoonup} \nu$", if

$$\left\langle\!\left\langle f, \nu^{(j)} \right\rangle\!\right\rangle \to \left\langle\!\left\langle f, \nu \right\rangle\!\right\rangle \qquad \text{as } j \to \infty \tag{4.9}$$

for all $f \in \mathrm{C}_0(\Omega \times \mathbb{R}^N)$. This is in particular implied by (4.5).

Proof. Define the measures

$$\mu^{(j)} := \mathscr{L}^d_x \llcorner \Omega \otimes \nu^{(j)}_x,$$

which is just a shorthand notation for the Radon measures $\mu^{(j)} \in \mathscr{M}(\Omega \times \mathbb{R}^N) \cong \mathrm{C}_0(\Omega \times \mathbb{R}^N)^*$ defined through their action

$$\langle f, \mu^{(j)} \rangle = \int_\Omega \int f(x, A)\, \mathrm{d}\nu^{(j)}_x(A)\, \mathrm{d}x \qquad \text{for all } f \in \mathrm{C}_0(\Omega \times \mathbb{R}^N).$$

For instance, if $\nu^{(j)} = \delta[V_j]$ with the elementary Young measure $\delta[V_j]$ defined in (4.2), we recover

$$\langle f, \mu^{(j)} \rangle = \int_\Omega f(x, V_j(x))\, \mathrm{d}x \qquad \text{for all } f \in \mathrm{C}_0(\Omega \times \mathbb{R}^N).$$

Clearly, every so-defined $\mu^{(j)}$ is a *positive* measure and

$$\left|\langle f, \mu^{(j)} \rangle\right| \le |\Omega| \cdot \|f\|_\infty,$$

whereby the $(\mu^{(j)})$ constitute a uniformly bounded sequence in the dual space $\mathrm{C}_0(\Omega \times \mathbb{R}^N)^*$. Thus, by the Sequential Banach–Alaoglu Theorem A.3, we can select a (not explicitly labeled) subsequence such that there exists a $\mu \in \mathrm{C}_0(\Omega \times \mathbb{R}^N)^*$ with

$$\langle f, \mu^{(j)} \rangle \to \langle f, \mu \rangle \qquad \text{for all } f \in \mathrm{C}_0(\Omega \times \mathbb{R}^N). \tag{4.10}$$

Next, we will show that μ can again be written in the form $\mu = \mathscr{L}^d_x \llcorner \Omega \otimes \nu_x$ for a weakly* measurable parametrized family $\nu = (\nu_x)_{x\in\Omega} \subset \mathscr{M}^1(\mathbb{R}^N)$ of probability measures. For this we will use the following measure-theoretic *disintegration theorem*, which is proved below.

Theorem 4.4. *Let $\Omega \subset \mathbb{R}^d$ be open and let $\mu \in \mathscr{M}^+(\Omega \times \mathbb{R}^N)$ 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 the measure $\kappa \in \mathscr{M}^+(\Omega)$ defined via*

$$\kappa(B) := \mu(B \times \mathbb{R}^N) \qquad \textit{for any Borel set } B \subset \Omega,$$

it holds that

$$\mu = \kappa(\mathrm{d}x) \otimes \nu_x,$$

that is,

$$\int f\, \mathrm{d}\mu = \int_\Omega \int f(x, A)\, \mathrm{d}\nu_x(A)\, \mathrm{d}\kappa(x) \qquad \textit{for all } f \in \mathrm{C}_0(\Omega \times \mathbb{R}^N). \tag{4.11}$$

Furthermore, the family $(\nu_x)_{x\in\Omega} \subset \mathscr{M}^1(\mathbb{R}^N)$ *is* κ*-essentially unique, that is, if* $(\nu'_x)_{x\in\Omega} \subset \mathscr{M}^1(\mathbb{R}^N)$ *has the properties above, then* $\nu_x = \nu'_x$ *for* κ*-almost every* $x \in \Omega$.

With this theorem at hand, we first observe that (4.10) together with Lemma A.19 implies that for every open set $U \subset \Omega$ it holds that

$$\mu(U \times \mathbb{R}^N) \le \liminf_{j\to\infty} \mu^{(j)}(U \times \mathbb{R}^N) = |U|. \tag{4.12}$$

On the other hand, employing the same lemma again, we get for every compact set $K \subset \Omega$ and any $R > 0$ that

$$\begin{aligned} \mu(K \times \overline{B(0,R)}) &\ge \limsup_{j\to\infty} \mu^{(j)}(K \times \overline{B(0,R)}) \\ &= \limsup_{j\to\infty} \int_K \int_{\{|A|\le R\}} 1 \, \mathrm{d}\nu_x^{(j)}(A) \, \mathrm{d}x \\ &\ge |K| - \frac{1}{R^p} \sup_{j\in\mathbb{N}} \langle\!\langle |\bullet|^p, \nu^{(j)} \rangle\!\rangle. \end{aligned}$$

Letting $R \to \infty$ and employing (4.3) as well as the inner regularity of Radon measures, see Appendix A.3, we arrive at

$$\mu(K \times \mathbb{R}^N) \ge |K|.$$

Together with (4.12), we thus get from the disintegration theorem that

$$\mu = \mathscr{L}_x^d \llcorner \Omega \otimes \nu_x,$$

where $(\nu_x)_x$ is a weakly*-measurable family of probability measures. Thus, for $f \in \mathrm{C}_0(\Omega \times \mathbb{R}^N)$,

$$\lim_{j\to\infty} \langle\!\langle f, \nu^{(j)} \rangle\!\rangle = \lim_{j\to\infty} \langle f, \mu^{(j)} \rangle = \langle f, \mu \rangle = \langle\!\langle f, \nu \rangle\!\rangle,$$

which is (4.5) for such integrands f.

Next, we show (4.5) in the case when f is Carathéodory and bounded and such that there exists a compact set $K \subset \mathbb{R}^N$ with $\operatorname{supp} f \subset \Omega \times K$, so that for almost every fixed $x \in \Omega$ the function $f(x, \bullet)$ is uniformly continuous. We will need the following theorem, which is proved below.

Theorem 4.5 (Scorza Dragoni 1948 [240]). *Suppose that* $f\colon \Omega \times \mathbb{R}^N \to \mathbb{R}$ *is a Carathéodory integrand such that for almost every fixed* $x \in \Omega$ *the function* $f(x, \bullet)$ *is uniformly continuous. Then, there exists an increasing sequence of compact sets* $S_k \subset \Omega$ $(k \in \mathbb{N})$ *with* $|\Omega \setminus S_k| \downarrow 0$ *such that* $f|_{S_k \times \mathbb{R}^N}$ *is continuous.*

So, let the sets $S_k \Subset \Omega$ ($k \in \mathbb{N}$) with $|\Omega \setminus S_k| \downarrow 0$ be such that $f|_{S_k \times \mathbb{R}^N}$ is continuous for our Carathéodory integrand f. Let furthermore $f_k \in \mathrm{C}_0(\Omega \times \mathbb{R}^N)$ be an extension of $f|_{S_k \times \mathbb{R}^N}$ to all of $\Omega \times \mathbb{R}^N$ and assume that the f_k are uniformly (in k) bounded; this uses the Tietze Extension Theorem A.33, a cut-off construction, and a truncation (we omit the details as they are straightforward).

Since $f_k \in \mathrm{C}_0(\Omega \times \mathbb{R}^N)$, the sequence $(\langle f_k(x, \bullet), \nu_x^{(j)} \rangle)_j$ is weakly precompact in $\mathrm{L}^1(\Omega)$. Thus, (4.10) implies

$$\big\langle f_k(x, \bullet), \nu_x^{(j)} \big\rangle \rightharpoonup \big\langle f_k(x, \bullet), \nu_x \big\rangle \quad \text{in } \mathrm{L}^1(\Omega) \qquad \text{as } j \to \infty.$$

In particular,

$$\big\langle f(x, \bullet), \nu_x^{(j)} \big\rangle \rightharpoonup \big\langle f(x, \bullet), \nu_x \big\rangle \quad \text{in } \mathrm{L}^1(S_k) \qquad \text{as } j \to \infty.$$

On the other hand,

$$\int_\Omega \Big| \big\langle f(x, \bullet), \nu_x^{(j)} \big\rangle - \mathbb{1}_{S_k} \big\langle f(x, \bullet), \nu_x^{(j)} \big\rangle \Big| \,\mathrm{d}x \le \int_{\Omega \setminus S_k} \Big| \big\langle f(x, \bullet), \nu_x^{(j)} \big\rangle \Big| \,\mathrm{d}x$$

and this converges to zero as $k \to \infty$, uniformly in j, by the boundedness of f. The same estimate holds with ν in place of $\nu^{(j)}$. Therefore, we may conclude that

$$\big\langle f(x, \bullet), \nu_x^{(j)} \big\rangle \rightharpoonup \big\langle f(x, \bullet), \nu_x \big\rangle \quad \text{in } \mathrm{L}^1(\Omega) \qquad \text{as } j \to \infty,$$

which directly yields (4.5) for bounded Carathéodory integrands f with $\operatorname{supp} f \subset \Omega \times K$.

Finally, to remove the restriction of boundedness and compact support in A, we remark that it suffices to show (4.5) under the additional constraint $f \ge 0$ by considering the positive and negative parts separately (the resulting functions are still Carathéodory integrands). Choose for any $h \in \mathbb{N}$ a cut-off function $\rho_h \in \mathrm{C}_c^\infty(\mathbb{R}; [0, 1])$ with $\rho_h = 1$ on $B(0, h)$ and $\operatorname{supp} \rho_h \subset B(0, 2h)$. Set

$$f^h(x, A) := \rho_h(|A|^{p/2}) \rho_h(f(x, A)) f(x, A).$$

Then,

$$\begin{aligned} E_{j,h} &:= \int_\Omega \int |f(x, \bullet) - f^h(x, \bullet)| \,\mathrm{d}\nu_x^{(j)} \,\mathrm{d}x \\ &\le \int_\Omega \int \big[1 - \rho_h(|A|^{p/2}) \rho_h(f(x, A))\big] |f(x, A)| \,\mathrm{d}\nu_x^{(j)}(A) \,\mathrm{d}x \\ &\le \iint_{\{(x,A) \in \Omega \times \mathbb{R}^N \,:\, |A|^{p/2} \ge h \text{ or } |f(x,A)| \ge h\}} |f(x, A)| \,\mathrm{d}\nu_x^{(j)}(A) \,\mathrm{d}x \end{aligned}$$

$$
\begin{aligned}
&\le \int_\Omega \int_{\{A\in\mathbb{R}^N \,:\, |A|^{p/2}\ge h \text{ and } |f(x,A)|<h\}} h \, \mathrm{d}\nu_x^{(j)}(A)\, \mathrm{d}x \\
&\quad + \iint_{\{(x,A)\in\Omega\times\mathbb{R}^N \,:\, |f(x,A)|\ge h\}} |f(x,A)| \, \mathrm{d}\nu_x^{(j)}(A)\, \mathrm{d}x \\
&= \frac{1}{h}\int_\Omega \int_{\{A\in\mathbb{R}^N \,:\, |A|^{p/2}\ge h\}} h^2 \, \mathrm{d}\nu_x^{(j)}(A)\, \mathrm{d}x \\
&\quad + \iint_{\{(x,A)\in\Omega\times\mathbb{R}^N \,:\, |f(x,A)|\ge h\}} |f(x,A)| \, \mathrm{d}\nu_x^{(j)}(A)\, \mathrm{d}x \\
&\le \frac{1}{h}\sup_{j\in\mathbb{N}} \left\langle\!\left\langle |A|^p, \nu^{(j)} \right\rangle\!\right\rangle + \sup_{j\in\mathbb{N}} \left\langle\!\left\langle |f(x,A)|\mathbb{1}_{\{|f(x,A)|\ge h\}}, \nu^{(j)} \right\rangle\!\right\rangle.
\end{aligned}
$$

Both of these terms converge to zero as $h \to \infty$ by the assumptions from the compactness principle, in particular (4.3) and the equiintegrability condition (4.6). Consequently, for fixed $h \in \mathbb{N}$,

$$
\begin{aligned}
&\lim_{j\to\infty} \left| \langle\!\langle f, \nu^{(j)} \rangle\!\rangle - \langle\!\langle f, \nu \rangle\!\rangle \right| \\
&\quad \le \limsup_{j\to\infty} \left(\left|\langle\!\langle f - f^h, \nu^{(j)} \rangle\!\rangle\right| + \left|\langle\!\langle f^h, \nu^{(j)} \rangle\!\rangle - \langle\!\langle f^h, \nu \rangle\!\rangle\right| + \left|\langle\!\langle f^h - f, \nu \rangle\!\rangle\right| \right) \\
&\quad \le \sup_{j\in\mathbb{N}} E_{j,h} + \limsup_{j\to\infty} \left|\langle\!\langle f^h - f, \nu \rangle\!\rangle\right|,
\end{aligned}
$$

where we used the previous step, namely that the Young measure convergence holds for f^h, to see that the middle upper limit vanishes. Now, the first term vanishes as $h \to \infty$ by the preceding estimate and the second term tends to zero by the pointwise bounded convergence of f^h to f and $f \ge 0$. Thus, (4.5) follows also in this case.

The last assertion to be shown in the compactness principle is (4.7) if $p \in [1, \infty)$ and (4.8) if $p = \infty$. For (4.7) let $h \in \mathbb{N}$ and define $|A|_h := \min\{|A|, h\}$. Then,

$$
\liminf_{j\to\infty} \langle\!\langle |\cdot|^p, \nu^{(j)} \rangle\!\rangle \ge \lim_{j\to\infty} \langle\!\langle |\cdot|_h^p, \nu^{(j)} \rangle\!\rangle = \langle\!\langle |\cdot|_h^p, \nu \rangle\!\rangle \qquad \text{for all } h \in \mathbb{N}.
$$

We conclude by letting $h \to \infty$ and using the monotone convergence theorem. For (4.8) take any $\varphi \in \mathrm{C}_0(\Omega)$, $\psi \in \mathrm{C}_0(\mathbb{R}^N)$ with $\operatorname{supp}\psi \cap K = \emptyset$. Then,

$$
\langle\!\langle \varphi \otimes \psi, \nu \rangle\!\rangle = \lim_{j\to\infty} \langle\!\langle \varphi \otimes \psi, \nu^{(j)} \rangle\!\rangle = 0.
$$

Varying φ and ψ, the assertion $\operatorname{supp}\nu_x \subset K$ for almost every $x \in \Omega$ follows. □

Proof of Theorem 4.4. For $\psi \in \mathrm{C}_0(\mathbb{R}^N)$ we define the (signed) measure $\mu_\psi \in \mathscr{M}(\Omega)$ via

$$
\mu_\psi(B) := \int_{B\times\mathbb{R}^N} \psi(A)\, \mathrm{d}\mu(x, A)
$$

for any Borel set $B \subset \Omega$. We have

$$\mu_\psi(B) \le \|\psi\|_\infty \mu(B \times \mathbb{R}^N) = \|\psi\|_\infty \kappa(B).$$

Thus, by the Besicovitch Differentiation Theorem A.23 there exists a κ-measurable function $h_\psi : \Omega \to \mathbb{R}$ with $|h_\psi| \le \|\psi\|_\infty$ such that $\mu_\psi = h_\psi \kappa$. This construction is linear in ψ, that is,

$$\mu_{\alpha\psi_1+\beta\psi_2} = \alpha\mu_{\psi_1} + \beta\mu_{\psi_2} = \alpha h_{\psi_1}\kappa + \beta h_{\psi_2}\kappa = (h_{\alpha\psi_1+\beta\psi_2})\kappa$$

for all $\psi_1, \psi_2 \in \mathrm{C}_0(\mathbb{R}^N)$ and $\alpha, \beta \in \mathbb{R}$.

Fix a countable dense family of functions $\mathscr{D} \subset \mathrm{C}_0(\mathbb{R}^N)$. Then, we can find a κ-negligible set $N \subset \Omega$ such that

$$h_{\psi_1}(x) + h_{\psi_2}(x) = h_{\psi_1+\psi_2}(x) \qquad \text{for all } x \in \Omega \setminus N \text{ and all } \psi_1, \psi_2 \in \mathscr{D}.$$

Setting $T_x[\psi] := h_\psi(x)$ for $x \in \Omega \setminus N$ and $\psi \in \mathscr{D}$, we see that $|T_x[\psi]| \le \|\psi\|_\infty$ and thus T_x can be extended to a linear bounded operator on $\mathrm{C}_0(\mathbb{R}^N)$. Hence, the Riesz Representation Theorem A.21 shows that for every $x \in \Omega \setminus N$ there exists a measure $\nu_x \in \mathscr{M}(\mathbb{R}^N)$ with $|\nu_x|(\mathbb{R}^N) \le 1$ such that

$$T_x[\psi] = \int_{\mathbb{R}^N} \psi(A) \,\mathrm{d}\nu_x(A), \qquad \psi \in \mathrm{C}_0(\mathbb{R}^N).$$

Further setting $\nu_x := \delta_0$ at points $x \in N$, we note that for all $\psi \in \mathscr{D}$ the function $x \mapsto \langle \psi, \nu_x \rangle = T_x[\psi] = h_\psi(x)$ is κ-measurable by definition. Thus, the κ-measurability also holds for $\psi \in \mathrm{C}_0(\mathbb{R}^N)$ by approximation. By a further approximation, the details of which we omit (see Problem 4.1), this implies the weak* measurability of the family $(\nu_x)_{x\in\Omega}$.

For a Borel set $B \subset \Omega$ and $\psi \in \mathscr{D}$, we have

$$\begin{aligned}
\int_{\Omega\times\mathbb{R}^N} \mathbb{1}_B(x)\psi(A) \,\mathrm{d}\mu(x, A) &= \mu_\psi(B) \\
&= \int_B h_\psi(x) \,\mathrm{d}\kappa(x) \\
&= \int_B \int_{\mathbb{R}^N} \psi(A) \,\mathrm{d}\nu_x(A) \,\mathrm{d}\kappa(x) \\
&= \int_\Omega \int_{\mathbb{R}^N} \mathbb{1}_B(x)\psi(A) \,\mathrm{d}\nu_x(A) \,\mathrm{d}\kappa(x).
\end{aligned}$$

This is (4.11) for $f := \mathbb{1}_B \otimes \psi$. By a (multi-stage) approximation, (4.11) then also holds for all $f \in \mathrm{C}_0(\Omega \times \mathbb{R}^N)$ and also for all $f := \mathbb{1}_{B\times\mathbb{R}^N}$ for any Borel set $B \subset \Omega$.

To see that the ν_x are indeed probability measures, it suffices to observe

$$\mu(B \times \mathbb{R}^N) = \int_B \nu_x(\mathbb{R}^N) \,\mathrm{d}\kappa(x) \le \int_B 1 \,\mathrm{d}\kappa(x) = \mu(B \times \mathbb{R}^N)$$

for all Borel set $B \subset \Omega$. Hence, $\nu_x(\mathbb{R}^N) = 1$ for κ-almost every $x \in \Omega$, which together with $|\nu_x|(\mathbb{R}^N) \le 1$ implies that ν_x is a probability measure.

The κ-essential uniqueness follows directly by applying (4.11) to all $f := \varphi \otimes \psi$ with $\varphi \in \mathrm{C}_0(\Omega)$, $\psi \in \mathrm{C}_0(\mathbb{R}^N)$. □

Proof of Theorem 4.5. For $j \in \mathbb{N}$ consider the functions

$$g_j(x) := \sup\Big\{ |f(x,A) - f(x,B)| \;\mathbf{:}\; A, B \in \mathbb{R}^N,\ |A-B| \le 1/j \Big\}, \qquad x \in \Omega.$$

Since f is Carathéodory and for almost every $x \in \Omega$ the function $f(x, \bullet)$ is uniformly continuous, we have $g_j \to 0$ pointwise almost everywhere as $j \to \infty$.

Let $n \in \mathbb{N}$. By Egorov's Theorem A.13, there exists a compact set $K_0 \subset \Omega$ with $|\Omega \setminus K_0| \le 1/(2n)$ and $g_j \to 0$ uniformly on K_0. Let $\{A_i\}_{i\in\mathbb{N}} \subset \mathbb{R}^N$ be dense in $\mathbb{R}^N$. By Lusin's Theorem A.16 there are compact sets K_i ($i \in \mathbb{N}$) such that $|\Omega \setminus K_i| \le 1/(2^{i+1}n)$ and $f(\bullet, A_i)$ is continuous in K_i. For

$$S_n := K_0 \cap \bigcap_{i=1}^{\infty} K_i$$

we estimate

$$|\Omega \setminus S_n| \le \frac{1}{2n}\biggl(1 + \sum_{i=1}^{\infty} \frac{1}{2^i}\biggr) = \frac{1}{n}.$$

Consequently, $|\Omega \setminus S_n| \to 0$ as $n \to \infty$.

Let $\varepsilon > 0$ and choose $\delta > 0$ such that $|A - B| \le 2\delta$ for $A, B \in \mathbb{R}^N$ implies $|f(x,A) - f(x,B)| \le \varepsilon$ for all $x \in S_n \subset K_0$ (the existence of such a δ follows from the uniform convergence $g_j \to 0$ on S_n).

For any $(\bar{x}, \bar{A}) \in S_n \times \mathbb{R}^N$ pick A_i from the dense collection $\{A_i\}$ such that $|\bar{A} - A_i| \le \delta$. For this A_i there exists an $\eta > 0$ such that for all $y \in S_n \subset K_i$ with $|\bar{x} - y| \le \eta$ it holds that $|f(\bar{x}, A_i) - f(y, A_i)| \le \varepsilon$. So, for $(x, A) \in S_n \times \mathbb{R}^N$ with

$$|\bar{x} - x| \le \eta \quad \text{and} \quad |\bar{A} - A| \le \delta,$$

we have $|A_i - A| \le |A_i - \bar{A}| + |\bar{A} - A| \le 2\delta$, and so,

$$\begin{aligned} |f(\bar{x},\bar{A}) - f(x,A)| &\le |f(\bar{x},\bar{A}) - f(\bar{x},A_i)| + |f(\bar{x},A_i) - f(x,A_i)| \\ &\quad + |f(x,A_i) - f(x,A)| \\ &\le 3\varepsilon. \end{aligned}$$

Hence, $f|_{S_n \times \mathbb{R}^N}$ is continuous at any $(\bar{x}, \bar{A}) \in S_n \times \mathbb{R}^N$. □

Proof of the Fundamental Theorem 4.1. We apply the compactness principle to the sequence $(\delta[V_j])_j$ of elementary Young measures defined in (4.2). The boundedness conditions (4.3), (4.4) are directly implied by the L^p-boundedness assumption on (V_j). For $(\delta[V_j])_j$ the assumption (4.6) expresses precisely the equiintegrability of $(f(x, V_j))_j$. Thus, (i)–(iii) follow from the compactness principle. □

The Fundamental Theorem 4.1 can also be proved in a more functional analytic way as follows: Let $L^\infty_{w*}(\Omega; \mathscr{M}(\mathbb{R}^N))$ be the set of essentially bounded weakly* measurable functions defined on Ω with values in the Radon measures $\mathscr{M}(\mathbb{R}^N)$. It turns out (see, for example, [27] or [96, 97]) that $L^\infty_{w*}(\Omega; \mathscr{M}(\mathbb{R}^N))$ is the dual space to $L^1(\Omega; C_0(\mathbb{R}^N))$. One can show that the maps $\nu^{(j)} = (x \mapsto \nu_x^{(j)})_j$ form a uniformly bounded set in $L^\infty_{w*}(\Omega; \mathscr{M}(\mathbb{R}^N))$ and by the sequential Banach–Alaoglu Theorem A.3 we can again conclude the existence of a weak* limit point ν of the $\nu^{(j)}$'s, which also inherits the property of being a collection of probability measures. The extended representation of limits for Carathéodory integrands f follows as before.

Finally, we show a "lower semicontinuity result" for the duality pairing between a Young measure and a positive integrand for which we do not have equiintegrability of the compound functions.

Proposition 4.6. *Let $(V_j) \subset L^p(\Omega; \mathbb{R}^N)$, $p \in [1, \infty]$, be a norm-bounded sequence generating the Young measure $\nu \in \mathbf{Y}^p(\Omega; \mathbb{R}^N)$ and let $f : \Omega \times \mathbb{R}^N \to [0, \infty)$ be a Carathéodory integrand (not necessarily satisfying the equiintegrability property in (iii) of the Fundamental Theorem). Then,*

$$\liminf_{j\to\infty} \int_\Omega f(x, V_j(x))\, dx = \liminf_{j\to\infty} \langle\!\langle f, \delta[V_j] \rangle\!\rangle \geq \langle\!\langle f, \nu \rangle\!\rangle.$$

Proof. For $h \in \mathbb{N}$ define $f_h(x, A) := \min\{f(x, A), h\}$. Then, (iii) from the Fundamental Theorem is applicable for the integrand f_h and we get

$$\int_\Omega f_h(x, V_j(x))\, dx \to \langle\!\langle f_h, \nu \rangle\!\rangle = \int_\Omega \int f_h(x, A)\, d\nu_x(A)\, dx.$$

Since $f \geq f_h$, we have

$$\liminf_{j\to\infty} \int_\Omega f(x, V_j(x))\, dx \geq \langle\!\langle f_h, \nu \rangle\!\rangle.$$

We conclude by letting $h \to \infty$ and using the monotone convergence theorem. □

4.2 Examples

We will now consider a few examples of Young measures. Most of these $\nu = (\nu_x)_x \in \mathbf{Y}^p(\Omega; \mathbb{R}^N)$ will in fact be **homogeneous**, that is, ν_x is almost everywhere constant in $x \in \Omega$; we then simply write ν in place of ν_x.

In order to identify the Young measure generated by some sequence, the following simple result often turns out to be useful.

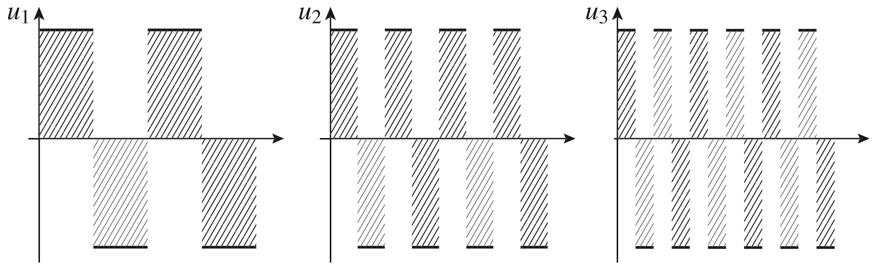


Fig. 4.1 An oscillating sequence

Lemma 4.7. *There exists a countable family* $\{\varphi_k \otimes h_k\}_{k\in\mathbb{N}} \subset \mathrm{C}_0(\Omega) \times \mathrm{C}_0(\mathbb{R}^N)$ *with the following property: If* $(V_j) \subset \mathrm{L}^p(\Omega;\mathbb{R}^N)$ *is uniformly norm-bounded and* $\nu \in \mathbf{Y}^p(\Omega;\mathbb{R}^N)$ *is such that*

$$\lim_{j\to\infty} \int_\Omega \varphi_k(x) h_k(V_j(x))\,\mathrm{d}x = \int_\Omega \varphi_k(x)\big\langle h_k, \nu_x\big\rangle\,\mathrm{d}x \qquad \textit{for all } k \in \mathbb{N},$$

then $V_j \xrightarrow{\mathbf{Y}} \nu$.

Proof. Let $\{\varphi_k\}_k$ and $\{h_l\}_l$ be countable dense subsets of $\mathrm{C}_0(\Omega)$ and $\mathrm{C}_0(\mathbb{R}^N)$, respectively. The assertion of the lemma is immediate (with a numbering of $\mathbb{N}\times\mathbb{N}$) once we recall the basic fact from functional analysis that the set of linear combinations of the functions $f_{k,l} := \varphi_k \otimes h_l$ (that is, $f_{k,l}(x, A) := \varphi_k(x)h_l(A)$) is dense in $\mathrm{C}_0(\Omega \times \mathbb{R}^N)$ and testing with such functions determines Young measure convergence, as we have seen in the proof of the Fundamental Theorem 4.1. □

Example 4.8. In $\Omega := (0, 1)$ define $u := \mathbb{1}_{(0,1/2)} - \mathbb{1}_{(1/2,1)}$ and extend this function periodically to all of $\mathbb{R}$. Then, the functions $u_j(x) := u(jx)$ for $j \in \mathbb{N}$ (see Figure 4.1) generate the homogeneous Young measure $\nu \in \mathbf{Y}^\infty((0, 1))$ with

$$\nu = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1}.$$

Indeed, for $\varphi \in \mathrm{C}_0((0, 1))$, $h \in \mathrm{C}_0(\mathbb{R})$ we have that φ is uniformly continuous, say $|\varphi(x)-\varphi(y)| \le \omega(|x-y|)$ with a **modulus of continuity** $\omega\colon [0,\infty) \to [0,\infty)$, that is, ω is continuous, increasing, and $\omega(0) = 0$. Then, since h is uniformly bounded,

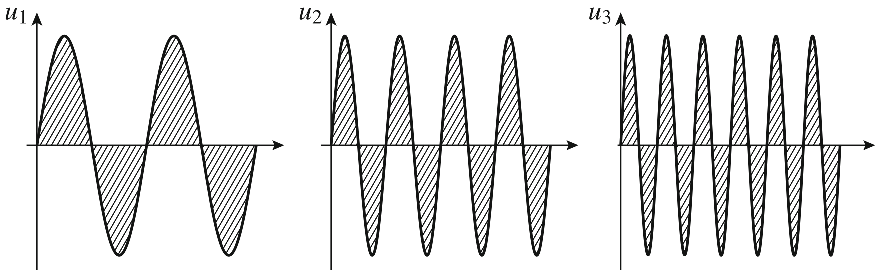


Fig. 4.2 Another oscillating sequence

$$\begin{aligned}
&\lim_{j\to\infty}\int_0^1 \varphi(x)h(u_j(x))\,\mathrm{d}x \\
&\quad= \lim_{j\to\infty}\sum_{k=0}^{j-1}\biggl[\int_{k/j}^{(k+1)/j}\varphi\Bigl(\frac{k}{j}\Bigr)h(u_j(x))\,\mathrm{d}x + \frac{1}{j}\mathrm{O}\bigl(\omega(1/j)\bigr)\biggr] \\
&\quad= \lim_{j\to\infty}\sum_{k=0}^{j-1}\frac{1}{j}\,\varphi\Bigl(\frac{k}{j}\Bigr)\int_0^1 h(u(y))\,\mathrm{d}y \\
&\quad= \int_0^1 \varphi(x)\,\mathrm{d}x\cdot\Bigl(\frac{1}{2}h(-1)+\frac{1}{2}h(+1)\Bigr).
\end{aligned}$$

For the last equality we used that the Riemann sums converge to the integral of φ. By Lemma 4.7, this identifies ν as claimed.

Example 4.9. Take $\Omega := (0,1)$ again and let $u_j(x) := \sin(2\pi jx)$ for $j \in \mathbb{N}$ (see Figure 4.2). The sequence (u_j) generates the homogeneous Young measure $\nu \in \mathbf{Y}^\infty((0,1))$ with

$$\nu = \frac{1}{\pi\sqrt{1-y^2}}\,\mathscr{L}^1_y \llcorner (-1,1).$$

This should be plausible from looking at the oscillating sequence; a formal proof is the task of Problem 4.4.

Example 4.10. Take a bounded Lipschitz domain $\Omega \subset \mathbb{R}^2$ and assume that $A, B \in \mathbb{R}^{2\times 2}$ are **rank-one connected**, that is, $B - A = a \otimes n$ for some $a, n \in \mathbb{R}^2$ (this is equivalent to $\operatorname{rank}(A-B) \le 1$). For $\theta \in (0,1)$ define

$$u(x) := Ax + \biggl(\int_0^{x\cdot n}\chi(t)\,\mathrm{d}t\biggr)a, \qquad x \in \mathbb{R}^2,$$

where

$$\chi := \mathbb{1}_{\bigcup_{z\in\mathbb{Z}}[z,z+1-\theta)}.$$

If we let $u_j(x) := u(jx)/j$, $x \in \Omega$, then the sequence (∇u_j) (restricted to $(0,1)^2$) generates the homogeneous Young measure $\nu \in \mathbf{Y}^\infty((0,1)^2;\mathbb{R}^{2\times 2})$ with

$$\nu = \theta\delta_A + (1-\theta)\delta_B.$$

4.3 Young Measures and Notions of Convergence

Next, we investigate how Young measure generation interacts with various notions of convergence.

Lemma 4.11. *Let $(V_j) \subset \mathrm{L}^p(\Omega;\mathbb{R}^N)$, $p \in (1,\infty]$, be a sequence generating the Young measure $\nu \in \mathbf{Y}^p(\Omega;\mathbb{R}^N)$. Then, setting $V(x) := [\nu](x) = [\nu_x]$ (the barycenter of ν), it holds that*

$$V_j \rightharpoonup V \quad \text{in } \mathrm{L}^p \text{ if } p \in (1,\infty) \qquad \text{or} \qquad V_j \overset{*}{\rightharpoonup} V \quad \text{in } \mathrm{L}^\infty \text{ if } p = \infty.$$

Proof. Bounded sequences in $\mathrm{L}^p(\Omega;\mathbb{R}^N)$ with $p > 1$ are weakly precompact and thus it suffices to identify the limit of any weakly(*) converging subsequence. From the Dunford–Pettis Theorem A.12 it follows that any such (sub)sequence is in fact (L^1-)equiintegrable. Now simply apply assertion (iii) of the Fundamental Theorem for the integrand $f(x,A) := A$ (or, more pedantically, $f_i(A) := A_i$ for $i = 1,\ldots,N$). □

The preceding lemma does not hold for $p = 1$. A counterexample is given by $V_j := j\mathbb{1}_{(0,1/j)}$, which *concentrates*.

Another important feature of Young measures is that they allow one to read off whether or not the generating sequence converges in measure:

Lemma 4.12. *Let $\nu \in \mathbf{Y}^p(\Omega;\mathbb{R}^N)$, $p \in [1,\infty]$, be the Young measure generated by a norm-bounded sequence $(V_j) \subset \mathrm{L}^p(\Omega;\mathbb{R}^N)$ and let $K \subset \mathbb{R}^N$ be compact. Then,*

$$\operatorname{dist}(V_j, K) \to 0 \text{ in measure} \qquad \Longleftrightarrow \qquad \operatorname{supp} \nu_x \subset K \text{ for a.e. } x \in \Omega.$$

Moreover, for $V \in \mathrm{L}^p(\Omega;\mathbb{R}^N)$,

$$V_j \to V \text{ in measure} \qquad \Longleftrightarrow \qquad \nu_x = \delta_{V(x)} \text{ for a.e. } x \in \Omega.$$

Proof. For any bounded, positive Carathéodory integrand $f\colon \Omega\times\mathbb{R}^N \to [0,1]$ and all $\delta \in (0,1)$ the Markov inequality and the Fundamental Theorem 4.1 imply

$$\begin{aligned}\limsup_{j\to\infty} \bigl|\bigl\{\, x \in \Omega \;:\; f(x,V_j(x)) \geq \delta \,\bigr\}\bigr| &\leq \lim_{j\to\infty} \frac{1}{\delta}\int_\Omega f(x,V_j(x))\,\mathrm{d}x \\ &= \frac{1}{\delta}\int_\Omega \int f(x,\,\boldsymbol{\cdot}\,)\,\mathrm{d}\nu_x\,\mathrm{d}x.\end{aligned}$$

On the other hand,

$$\int_{\Omega} \int f(x, \bullet) \, \mathrm{d}\nu_x \, \mathrm{d}x = \lim_{j\to\infty} \int_{\Omega} f(x, V_j(x)) \, \mathrm{d}x$$
$$\le \delta|\Omega| + \limsup_{j\to\infty} \bigl|\bigl\{ x \in \Omega \;:\; f(x, V_j(x)) \ge \delta \bigr\}\bigr|.$$

The preceding estimates show that $f(x, V_j(x))$ converges to zero in measure if and only if $\langle f(x, \bullet), \nu_x \rangle = 0$ for $\mathscr{L}^d$-almost every $x \in \Omega$.

For the first assertion set

$$f(x, A) := \frac{\operatorname{dist}(A, K)}{1 + \operatorname{dist}(A, K)}, \qquad (x, A) \in \Omega \times \mathbb{R}^N.$$

In this case, $f(x, V_j(x))$ converges to zero in measure if and only if $\operatorname{dist}(V_j, K) \to 0$ in measure, and $\langle f(x, \bullet), \nu_x \rangle = 0$ if and only if $\operatorname{supp} \nu_x \subset K$.

For the second assertion we choose

$$f(x, A) := \frac{|A - V(x)|}{1 + |A - V(x)|}, \qquad (x, A) \in \Omega \times \mathbb{R}^N.$$

Then, $f(x, V_j(x))$ converges to zero in measure if and only if V_j converges to V in measure, and $\langle f(x, \bullet), \nu_x \rangle = 0$ if and only if $\nu_x = \delta_{V(x)}$. □

4.4 Gradient Young Measures

The most important Young measures for our purposes are those that can be generated by a sequence of *gradients*. Let $\nu \in \mathbf{Y}^p(\Omega; \mathbb{R}^{m\times d})$ (use $\mathbb{R}^N = \mathbb{R}^{m\times d} \cong \mathbb{R}^{md}$ in the theory of the last section). We say that ν is a $\mathrm{W}^{1,p}$**-gradient Young measure**, where $p \in [1, \infty]$, in symbols $\nu \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m\times d})$, if there exists a norm-bounded sequence $(u_j) \subset \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ such that $\nabla u_j \xrightarrow{\mathbf{Y}} \nu$, i.e., the sequence (∇u_j) generates ν. Note that it is not required, and in fact never true, that *every* sequence that generates ν is a sequence of gradients. We have already considered examples of gradient Young measures in the previous section, we note in particular Example 4.10.

We first prove the following technical, but immensely useful, result about gradient Young measures.

Lemma 4.13. *Let $\nu \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m\times d})$, $p \in (1, \infty]$, and let $u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ be an **underlying deformation** of ν, that is, $[\nu] = \nabla u$. Then, there exists a norm-bounded sequence $(u_j) \subset \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ such that*

$$\operatorname{supp}(u_j - u) \Subset \Omega \qquad \textit{and} \qquad \nabla u_j \xrightarrow{\mathbf{Y}} \nu.$$

Furthermore, if $p \in (1, \infty)$, we can in addition require that the sequence (∇u_j) is L^p-equiintegrable.

Proof. Step 1. Since Ω has a Lipschitz boundary, we can extend a generating sequence (∇v_j) for ν to all of $\mathbb{R}^d$ (see Theorem A.25), so from now on we assume $(v_j) \subset \mathrm{W}^{1,p}(\mathbb{R}^d;\mathbb{R}^m)$ with $\sup_j \|v_j\|_{\mathrm{W}^{1,p}} < \infty$.

In the following we need the maximal function $Mf\colon \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$ of $f\colon \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$, see Appendix A.6. With this tool at hand, consider the sequence

$$V_j := M\bigl(|v_j| + |\nabla v_j|\bigr), \qquad j \in \mathbb{N}.$$

By Theorem A.36, (V_j) is uniformly bounded in $\mathrm{L}^p(\Omega)$ and we may select a subsequence (not explicitly labeled) such that (V_j) generates a Young measure $\mu \in \mathbf{Y}^p(\Omega;\mathbb{R}^{m\times d})$.

Step 2. For $p \in (1,\infty)$ we first show the claim concerning equiintegrability. Define for $h \in \mathbb{N}$ the (nonlinear) *truncation* τ_h,

$$\tau_h s := \begin{cases} s & \text{if } |s| \le h, \\ h\frac{s}{|s|} & \text{if } |s| > h, \end{cases} \qquad s \in \mathbb{R}.$$

For fixed $h \in \mathbb{N}$ the sequence $(\tau_h V_j)_j$ is uniformly bounded in $\mathrm{L}^\infty(\Omega)$ and so, by the Young measure representation of limits, we have for every $\varphi \in \mathrm{L}^\infty(\Omega)$ that

$$\begin{aligned} \lim_{h\to\infty}\lim_{j\to\infty}\int_\Omega \varphi(x)|\tau_h V_j(x)|^p \,\mathrm{d}x &= \lim_{h\to\infty}\int_\Omega \varphi(x)\int |\tau_h s|^p \,\mathrm{d}\mu_x(s)\,\mathrm{d}x \\ &= \langle\!\langle \varphi \otimes |\cdot|^p, \mu \rangle\!\rangle, \end{aligned} \tag{4.13}$$

where in the last step we used the monotone convergence theorem. Now choose for every $k \in \mathbb{N}$ a natural number $j(k) > j(k-1)$, where $j(0) := 0$, such that

$$\left| \lim_{j\to\infty}\int_\Omega |\tau_k V_j(x)|^p \,\mathrm{d}x - \int_\Omega |\tau_k V_n(x)|^p \,\mathrm{d}x \right| \le \frac{1}{k} \tag{4.14}$$

for all $n \ge j(k)$. Let $\psi \in \mathrm{L}^\infty(\Omega)$. We may also estimate for $l \le k$,

$$\begin{aligned} \int_\Omega \psi(x)|\tau_k V_{j(k)}(x)|^p \,\mathrm{d}x &\le \|\psi\|_{\mathrm{L}^\infty}\cdot\int_\Omega |\tau_k V_{j(k)}(x)|^p \,\mathrm{d}x \\ &\quad - \int_\Omega \bigl(\|\psi\|_{\mathrm{L}^\infty} - \psi(x)\bigr)|\tau_l V_{j(k)}|^p \,\mathrm{d}x. \end{aligned}$$

Thus, using (4.13) for $\varphi := \mathbb{1}_\Omega$, (4.14), and also the Young measure representation of limits,

$$\begin{aligned} \limsup_{k\to\infty}\int_\Omega \psi(x)|\tau_k V_{j(k)}(x)|^p \,\mathrm{d}x &\le \|\psi\|_{\mathrm{L}^\infty}\cdot\langle\!\langle \mathbb{1}\otimes|\cdot|^p, \mu\rangle\!\rangle \\ &\quad - \int_\Omega \bigl(\|\psi\|_{\mathrm{L}^\infty} - \psi(x)\bigr)\int |\tau_l s|^p \,\mathrm{d}\mu_x(s)\,\mathrm{d}x. \end{aligned}$$

Letting $l \to \infty$, we get by the monotone convergence theorem

$$\limsup_{k\to\infty} \int_\Omega \psi(x)|\tau_k V_{j(k)}(x)|^p \,\mathrm{d}x \le \langle\!\langle \psi \otimes |\bullet|^p, \mu \rangle\!\rangle.$$

Repeating the same argument for $-\psi$, we conclude that

$$|\tau_k V_{j(k)}| \rightharpoonup \big(x \mapsto \langle |\bullet|^p, \mu_x \rangle\big) \quad \text{in } \mathrm{L}^1.$$

Thus, by the Dunford–Pettis Theorem A.12 the functions $W_k := \tau_k V_{j(k)}$ ($k \in \mathbb{N}$) are uniformly L^p-bounded and L^p-equiintegrable.

Theorem A.36 implies that $v_{j(k)}$ is Lipschitz continuous with Lipschitz constant at most Ck on the set

$$S_k := \big\{ x \in \Omega \;:\; V_{j(k)}(x) \le k \big\}$$

and by the Kirszbraun Theorem A.34, we may extend each v_k to a function $w_k \colon \mathbb{R}^d \to \mathbb{R}^m$ that is globally Lipschitz continuous with Lipschitz constant at most Ck. Since $w_k = v_{j(k)}$ in S_k, for the gradients ∇w_k (which exist almost everywhere by Rademacher's Theorem A.30) we have

$$\begin{aligned} |\nabla w_k| &= |\nabla v_{j(k)}| \le V_{j(k)} = W_k && \text{a.e. in } S_k, \\ |\nabla w_k| &\le Ck = CW_k && \text{a.e. in } \Omega \setminus S_k. \end{aligned}$$

Consequently, $|\nabla w_k| \le CW_k$ almost everywhere in Ω and thus $\{\nabla w_k\}_k$ inherits the L^p-equiintegrability from $\{W_k\}_k$. Moreover, by the Markov inequality,

$$|\Omega \setminus S_k| \le \frac{\|V_k\|_{\mathrm{L}^p}^p}{k^p} \to 0 \qquad \text{as } k \to \infty.$$

Therefore, for all $\varphi \in \mathrm{C}_0(\Omega)$ and all $h \in \mathrm{C}_0(\mathbb{R}^m)$,

$$\int_\Omega |\varphi(x)h(\nabla w_k(x)) - \varphi(x)h(\nabla v_k(x))| \,\mathrm{d}x \le \|\varphi\|_\infty \cdot \|h\|_\infty \cdot |\Omega \setminus S_k| \to 0.$$

Since all such φ, h determine the Young measure (see the proof of the Fundamental Theorem 4.1), we have shown that the L^p-equiintegrable sequence (∇w_k) generates the same Young measure ν as (∇v_j).

Step 3. It remains to perform the boundary adjustment. Since $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ embeds compactly into the space $\mathrm{L}^p(\Omega;\mathbb{R}^m)$ by the Rellich–Kondrachov Theorem A.28, we have $w_k \to u$ in L^p. Let $(\rho_j) \subset \mathrm{C}_c^\infty(\Omega;[0,1])$ be a sequence of cut-off functions with the property that for the sets $G_j := \{ x \in \Omega \;:\; \rho_j(x) = 1 \}$ it holds that $|\Omega \setminus G_j| \to 0$ as $j \to \infty$. For

$$u_{j,k} := \rho_j w_k + (1-\rho_j)u \in \mathrm{W}^{1,p}_u(\Omega;\mathbb{R}^m)$$

we observe

$$\nabla u_{j,k} = \rho_j \nabla w_k + (1-\rho_j)\nabla u + (w_k - u) \otimes \nabla \rho_j.$$

For all $\varphi \in \mathrm{C}_0(\Omega)$ and $h \in \mathrm{C}_0(\mathbb{R}^m)$, we have

$$\int_\Omega |\varphi(x)h(\nabla w_k(x)) - \varphi(x)h(\nabla u_{j,k}(x))| \, \mathrm{d}x \le |\Omega \setminus G_j| \cdot \|\varphi\|_\infty \cdot \|h\|_\infty \to 0$$

as $j \to \infty$, uniformly in k. As we have remarked before, these φ, h determine the Young measure, so we can now select a diagonal sequence $u_j = u_{j,k(j)}$ such that (∇u_j) generates ν and satisfies all requirements from the statement of the lemma. It is easy to see that the equiintegrability is not affected by the cut-off procedure. □

It is the task of Problem 4.9 to show that the preceding lemma cannot hold in the case $p = 1$.

4.5 Homogeneous Gradient Young Measures

We next discuss some properties of **homogeneous gradient Young measures** $\nu \in \mathbf{GY}^p(B(0,1); \mathbb{R}^{m\times d})$, for which ν_x is almost everywhere constant in x. We simply write ν for any ν_x and $[\nu] = [\nu_x]$.

It is a particular consequence of the following *averaging principle* that the domain in the definition of homogeneous gradient Young measures can be chosen to be any bounded Lipschitz domain $D \subset \mathbb{R}^d$.

Lemma 4.14. *Let $\nu \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m\times d})$, where $p \in [1,\infty]$, such that $[\nu] = \nabla u$ for some $u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ with linear boundary values. Then, for any bounded Lipschitz domain $D \subset \mathbb{R}^d$ there exists a homogeneous gradient Young measure $\overline{\nu} \in \mathbf{GY}^p(D; \mathbb{R}^{m\times d})$ such that*

$$\int h \, \mathrm{d}\overline{\nu} = \fint_\Omega \int h \, \mathrm{d}\nu_x \, \mathrm{d}x \tag{4.15}$$

for all continuous $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ with p-growth if $p < \infty$ (no growth condition if $p = \infty$). This result remains valid if $\Omega = (-1/2, 1/2)^d$, the d-dimensional unit cube, and u has periodic boundary values.

Proof. We only treat the case $p \in [1,\infty)$, the case $p = \infty$ is in fact easier.

Let $(u_j) \subset \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ with $u_j|_{\partial\Omega} = Fx$ (in the sense of trace) for a fixed matrix $F \in \mathbb{R}^{m\times d}$ and such that $\nabla u_j \overset{\mathbf{Y}}{\to} \nu$. This sequence exists by Lemma 4.13 and the fact that $u|_{\partial\Omega} = Fx$ for some $F \in \mathbb{R}^{m\times d}$ (we denote by "Fx" the linear map $x \mapsto Fx$). In particular, $\sup_j \|\nabla u_j\|_{\mathrm{L}^p} < \infty$. For every $j \in \mathbb{N}$ choose a Vitali cover of D consisting of rescaled disjoint copies of Ω, see Theorem A.15, i.e.,

$$D = Z^{(j)} \cup \bigcup_{k=1}^{\infty} \Omega(a_k^{(j)}, r_k^{(j)}), \qquad |Z^{(j)}| = 0,$$

with $a_k^{(j)} \in D$, $0 < r_k^{(j)} \le 1/j$ ($k \in \mathbb{N}$), and $\Omega(a, r) := a + r\Omega$. Then define

$$v_j(y) := r_k^{(j)} u_j\biggl(\frac{y - a_k^{(j)}}{r_k^{(j)}}\biggr) + F a_k^{(j)} \qquad \text{if } y \in \Omega(a_k^{(j)}, r_k^{(j)}) \ (k \in \mathbb{N}).$$

We have $v_j \in \mathrm{W}^{1,p}(D; \mathbb{R}^m)$ (it is easy to see that there are no jumps over the gluing boundaries) and

$$\nabla v_j(y) = \nabla u_j\biggl(\frac{y - a_k^{(j)}}{r_k^{(j)}}\biggr) \qquad \text{if } y \in \Omega(a_k^{(j)}, r_k^{(j)}) \ (k \in \mathbb{N}).$$

We can then use a change of variables to compute for all $\varphi \in \mathrm{C}_0(D)$ and all continuous $h \colon \mathbb{R}^{m \times d} \to \mathbb{R}$ with p-growth that

$$\begin{aligned} \int_D \varphi(y) h(\nabla v_j(y)) \, \mathrm{d}y &= \sum_{k=1}^{\infty} \int_{\Omega(a_k^{(j)}, r_k^{(j)})} \varphi(y) \, h\biggl(\nabla u_j\biggl(\frac{y - a_k^{(j)}}{r_k^{(j)}}\biggr)\biggr) \, \mathrm{d}y \\ &= \sum_{k=1}^{\infty} (r_k^{(j)})^d \varphi(a_k^{(j)}) \int_{\Omega} h(\nabla u_j(x)) \, \mathrm{d}x + \mathrm{O}\biggl(\frac{1}{j}\biggr)|D|, \end{aligned}$$

where we also used that φ is uniformly continuous. Letting $j \to \infty$ and using that the Riemann sums converge to the integral,

$$\lim_{j \to \infty} \sum_{k=1}^{\infty} (r_k^{(j)})^d \varphi(a_k^{(j)}) = \frac{1}{|\Omega|} \int_D \varphi(x) \, \mathrm{d}x,$$

we arrive at

$$\lim_{j \to \infty} \int_D \varphi(y) h(\nabla v_j(y)) \, \mathrm{d}y = \int_D \varphi(x) \, \mathrm{d}x \cdot \fint_{\Omega} \int h(A) \, \mathrm{d}\nu_x \, \mathrm{d}x. \tag{4.16}$$

For $\varphi = \mathbb{1}$ and $h(A) := |A|^p$, this gives

$$\sup_{j \in \mathbb{N}} \|\nabla v_j\|_{\mathrm{L}^p}^p = \sup_{j \in \mathbb{N}} \|\nabla u_j\|_{\mathrm{L}^p}^p < \infty.$$

Thus, there exists a $\overline{\nu} \in \mathbf{GY}^p(D; \mathbb{R}^m)$ such that $\nabla v_j \overset{\mathbf{Y}}{\to} \overline{\nu}$ (up to selecting a subsequence). Using Lemma 4.13 we may moreover assume that (∇u_j), and hence also (∇v_j), is L^p-equiintegrable. Then, (4.16) implies

$$\int_D \int \varphi(y)h(A)\, \mathrm{d}\overline{\nu}_y(A)\, \mathrm{d}y = \int_D \varphi(x)\, \mathrm{d}x \cdot \fint_\Omega \int h(A)\, \mathrm{d}\nu_x(A)\, \mathrm{d}x$$

for all φ, h as above. This implies in particular that $(\overline{\nu}_y) = \overline{\nu}$ is homogeneous. For $\varphi = \mathbb{1}$, we get (4.15).

The additional claim about $\Omega = (-1/2, 1/2)^d$ and an underlying deformation u with periodic boundary values follows analogously, since in this case we can "glue" generating functions via a staircase construction; this is the task of Problem 4.10. □

Applying the preceding averaging principle to an elementary gradient Young measure, we get the following result, often called the *Riemann–Lebesgue lemma*.

Lemma 4.15. *Let $u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$, $p \in [1, \infty]$, have linear boundary values. Then, there exists a homogeneous gradient Young measure $\overline{\delta[\nabla u]} \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m\times d})$ such that*

$$\int h \, \mathrm{d}\overline{\delta[\nabla u]} = \fint_\Omega h(\nabla u(x))\, \mathrm{d}x \tag{4.17}$$

for all continuous $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ with p-growth if $p < \infty$ (no growth condition if $p = \infty$). This result remains valid if $\Omega = (-1/2, 1/2)^d$ and u has periodic boundary values.

Notes and Historical Remarks

Laurence Chisholm Young originally introduced the objects that are now called Young measures as "generalized curves/surfaces" in the late 1930s and early 1940s, see [280–282], to treat problems in the calculus of variations and optimal control theory that could not be solved using classical methods. His book [283] explains these objects and their applications in great detail (in particular, the "sailing against the wind" example from Section 1.6 is adapted from there). The theory of Young measures is now very mature and there are several monographs [57, 222, 235] that give overviews of the theory from different points of view.

In Chapter 7 we will consider relaxation problems formulated using Young measures. Further, as we will see in Chapters 8 and 9, Young measures provide a convenient framework to describe fine phase mixtures in the theory of microstructure. A second avenue of development—somewhat different from Young's original intention—is to use Young measures as a *technical tool* only. This approach is in fact quite old and was probably first adopted in a series of articles by McShane from the 1940s [184–186]. There, the author first finds a Young measure solution to a variational problem, then proves additional properties of the obtained minimizing Young measure, and finally concludes that these properties entail that the generalized solution is in fact classical.

Several people contributed to Young measure theory from the 1970s onward, including Berliocchi & Lasry [39], Balder [24], Ball [27] and Kristensen [164], among many others. An important breakthrough in this respect was the characterization of the class of Young measures generated by sequences of gradients in the early 1990s by Kinderlehrer and Pedregal [157, 158], see Theorem 7.15. Their result places gradient Young measures in duality with quasiconvex functions (to be defined in the next chapter) via Jensen-type inequalities; another work in this direction is Sychev's article [259]. Young measures can also be used to show regularity, see the recent work by Dolzmann & Kristensen [102]. Carstensen & Roubíček [56] considered numerical approximations.

Young measure theory was opened up to many new applications in the late 1970s and early 1980s, when Tartar [267, 268, 270] and Murat [209–211] developed the theory of compensated compactness and were able to settle many open problems in the theory of hyperbolic conservation laws; another important contributor here was DiPerna, see, for example, [98]. A key point of this strategy is to use the good compactness properties of Young measures to pass to limits in nonlinear quantities and then to deduce from pointwise and differential constraints on the generating sequences that the Young measure collapses to a point mass, corresponding to a classical function (so no oscillation phenomena occurred). Moreover, in this situation weak convergence improves to convergence in measure (or even in norm), hence the name compensated *compactness*. We discuss compensated compactness theory in Section 8.8.

The disintegration result from Theorem 4.4 is essentially contained in the result from probability theory that *regular conditional probabilities* exist, see, for instance, Theorem 89.1 in [234]. The result as stated also holds for vector-measures, see Theorem 2.28 of [15]. A stronger version of the Scorza Dragoni Theorem 4.5 can be found in Theorem 6.35 of [122]. Lemma 4.13 is a version of the well-known *decomposition lemma* from [125], another version is in [163].

In the case $p = 1$ the theory of (classical) Young measures is not very satisfactory and some important results such as Lemma 4.11 and Lemma 4.13 do not hold (see Problem 4.9 for a counterexample to Lemma 4.13 in the case $p = 1$). The fundamental reason for this deficiency is that in L^1 norm-bounded sequences are not weakly precompact. A partial remedy can be found by weakening the notion of convergence to be employed in L^1. Then, one can use Chacon's *biting lemma*:

Lemma 4.16 (Chacon 1980 [53]). *Let* $(V_j) \subset \mathrm{L}^1(\Omega; \mathbb{R}^N)$ *be a norm-bounded sequence generating the Young measure* $\nu \in \mathbf{Y}^1(\Omega; \mathbb{R}^N)$. *Define* $V(x) := [\nu_x]$. *Then, there exists an increasing sequence* $\Omega_k \subset \Omega$ *with* $|\Omega_k| \uparrow |\Omega|$ *such that* $V_j \rightharpoonup V$ *in* $\mathrm{L}^1(\Omega_k; \mathbb{R}^N)$ *for all k; this is called* ***biting convergence***.

More information on this topic can be found in [35] and Chapter 6 of [222]. We will return to the topic of Young measures generated by merely L^1-bounded sequences and develop a much more satisfying theory in Chapter 12.

Problems

4.1. Show that the family $(\nu_x)_{x\in\Omega}$ constructed in the proof of Theorem 4.4 is weakly* measurable. *Hint:* Use the Scorza Dragoni Theorem 4.5.

4.2. Prove that a sequence of measurable maps $V_j : \Omega \to \mathbb{R}^N$ satisfying only the tightness condition

$$\lim_{h\uparrow\infty} \sup_{j\in\mathbb{N}} |\{|V_j| \geq h\}| = 0$$

also generates a Young measure (in a suitable sense).

4.3. Show that for every $p \in [1, \infty)$ any weakly* measurable parametrized measure $(\nu_x)_{x\in\Omega} \subset \mathscr{M}^1(\mathbb{R}^N)$ with $\langle |\bullet|^p, \nu_x\rangle \in \mathrm{L}^p(\Omega)$ (as a function of x) can be generated by a sequence $(V_j) \subset \mathrm{L}^p(\Omega; \mathbb{R}^N)$. *Hint:* Approximate a general measure by linear combinations of Dirac masses and use a gluing argument.

4.4. Take $\Omega := (0, 1)$ and let $u_j(x) = \sin(2\pi jx)$ for $j \in \mathbb{N}$. Show that the sequence (u_j) generates the homogeneous Young measure $\nu \in \mathbf{Y}^\infty((0, 1))$ with

$$\nu_x = \frac{1}{\pi\sqrt{1-y^2}} \mathscr{L}^1_y \llcorner (-1, 1) \qquad \text{for a.e. } x \in (0, 1).$$

4.5. Let $a, b \in \mathbb{R}^m$ with $a \neq b$ and let $\theta \in (0, 1)$.

(i) Set $\Omega := (0, 1)$. Let $\nu = (\nu_x)_{x\in\Omega} \subset \mathscr{M}^1(\mathbb{R}^m)$ be the Young measure with

$$\nu_x = \theta\delta_a + (1-\theta)\delta_b \qquad \text{for a.e. } x \in \Omega.$$

Construct a generating sequence $(V_j) \subset \mathrm{L}^\infty(\Omega; \mathbb{R}^m)$ of ν.

(ii) Let $Q := (0, 1)^d$ and define

$$A := a \otimes \mathrm{e}_1, \qquad B := b \otimes \mathrm{e}_1 \in \mathbb{R}^{m\times d}.$$

Construct $(u_j) \subset \mathrm{W}^{1,\infty}(Q; \mathbb{R}^m)$, based on V_j from the previous problem, such that the sequence (∇u_j) generates the gradient Young measure $\mu = (\mu_y)_{y\in Q} \in \mathbf{Y}^\infty(Q; \mathbb{R}^{m\times d})$ given as

$$\mu_y = \theta\delta_A + (1-\theta)\delta_B \qquad \text{for a.e. } y \in Q.$$

4.6. Assume for the sequence from the previous problem, part (ii), that $u_j \overset{*}{\rightharpoonup} a$ in $\mathrm{W}^{1,\infty}$ for $a(y) := Fy$ with $F := \theta A + (1-\theta)B$. Based on this, construct a sequence $(v_j) \subset \mathrm{W}^{1,\infty}(Q; \mathbb{R}^m)$ such that $v_j \in \mathrm{C}(\overline{Q})$ for all $j \in \mathbb{N}$, (∇v_j) generates μ, and $v_j|_{\partial Q} = Fx$.

4.7. Let $A, B, C \in \mathbb{R}^{m \times d}$ such that for some $b, c \in \mathbb{R}^m$,

$$B - A = b \otimes \mathrm{e}_1 \qquad \text{and} \qquad C - A = c \otimes \mathrm{e}_1.$$

Let $\theta_A, \theta_B, \theta_C \in (0,1)$ be such that $\theta_A + \theta_B + \theta_C = 1$. Show that $\nu = (\nu_x)_{x \in \Omega}$ ($\Omega \subset \mathbb{R}^d$ a bounded Lipschitz domain that you can choose as you like) with

$$\nu_x := \theta_A \delta_A + \theta_B \delta_B + \theta_C \delta_C \qquad \text{for a.e. } x \in \Omega$$

is a (homogenous) $\mathrm{W}^{1,\infty}$-gradient Young measure with barycenter $[\nu] = \theta_A A + \theta_B B + \theta_C C$.

4.8. Let $\nu \in \mathbf{Y}^p(\Omega; \mathbb{R}^N)$ be a Young measure with generating sequence $(V_j) \subset \mathrm{L}^p(\Omega; \mathbb{R}^N)$. Show that for every closed set $E \subset \mathbb{R}^N$ it holds that $\nu_x(E)$ is the asymptotic fraction of values of v_j in E, that is,

$$\nu_x(E) = \lim_{r \downarrow 0} \lim_{j \to \infty} \frac{|\{ x \in B(x,r) : v_j(x) \in E \}|}{\omega_d r^d},$$

where $\omega_d = |B(0,1)|$. Also show that this formula fails in general when E is not closed.

4.9. Show that in $\Omega := (0,2)^2$ for the map $V \colon \Omega \to \mathbb{R}^2$ given as

$$V(x) := \begin{cases} 0 & \text{if } x \in (0,1)^2, \\ \mathrm{e}_1 & \text{if } x \in (0,2)^2 \setminus (0,1)^2, \end{cases}$$

there cannot exist a $u \in \mathrm{W}^{1,1}((0,2)^2)$ such that $\nabla u = V$. However, prove that the Young measure $\nu \in \mathbf{Y}^1((0,2)^2; \mathbb{R}^2)$ defined by

$$\nu_x := \mathbb{1}_{(0,1)^2}(x)\delta_0 + \mathbb{1}_{(0,2)^2 \setminus (0,1)^2}(x)\delta_{\mathrm{e}_1}, \qquad x \in (0,2)^2,$$

is in $\mathbf{GY}^1((0,2)^2; \mathbb{R}^2)$ by exhibiting a norm-bounded sequence $(u_j) \subset \mathrm{W}^{1,1}((0,2)^2)$ with $\nabla u_j \overset{\mathbf{Y}}{\to} \nu$. Conclude that Lemma 4.13 cannot be extended to cover the case $p = 1$.

4.10. Prove Lemma 4.14 in the case $\Omega = (-1/2, 1/2)^d$ and underlying deformation u with periodic boundary values.

Chapter 5
Quasiconvexity

We saw in the Tonelli–Serrin Theorem 2.6 that convexity of the integrand (in the gradient variable) implies the weak lower semicontinuity of the corresponding integral functional. Moreover, we proved in Proposition 2.9 that if $d = 1$ or $m = 1$, then convexity of the integrand is also necessary for weak lower semicontinuity. In the vectorial case ($d, m > 1$), however, it turns out that one can find weakly lower semicontinuous integral functionals whose integrands are non-convex. The following is the most fundamental one: Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain as usual and with $p \in [d, \infty)$ define

$$\mathscr{F}[u] := \int_\Omega \det \nabla u(x) \,\mathrm{d}x, \qquad u \in \mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^d).$$

Then, one can argue using the wedge product and Stokes' theorem that

$$\mathscr{F}[u] = \int_\Omega du^1 \wedge \cdots \wedge du^d = \int_{\partial\Omega} u^1 \wedge du^2 \wedge \cdots \wedge du^d = 0,$$

because $u \in \mathrm{W}_0^{1,d}(\Omega; \mathbb{R}^d)$ is zero on the boundary $\partial\Omega$ (this can also be computed in a more elementary way, see Lemma 5.8 below). Thus, $\mathscr{F}$ is in fact constant on $\mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^d)$, hence trivially weakly lower semicontinuous. Using slightly more sophisticated arguments (to be made precise in this chapter), we will also show that $\mathscr{F}$ is weakly continuous on the whole space $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^d)$ if $p \in [d, \infty)$.

However, the determinant function is far from being convex if $d \geq 2$; for instance, we can easily write a matrix with positive determinant as the convex combination of two singular matrices. We can also find examples not involving singular matrices: For

$$A := \begin{pmatrix} -1 & -2 \\ 2 & 1 \end{pmatrix}, \qquad B := \begin{pmatrix} 1 & -2 \\ 2 & -1 \end{pmatrix}, \qquad \frac{1}{2}A + \frac{1}{2}B = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix},$$

we have $\det A = \det B = 3$, but $\det(A/2 + B/2) = 4$.

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8_5

Furthermore, convexity of the integrand is not compatible with one of the most fundamental principles of continuum mechanics: Assume that our integrand $f = f(A)$ is **frame-indifferent**, that is,

$$f(QA) = f(A) \qquad \text{for all } Q \in \mathrm{SO}(d),\ A \in \mathbb{R}^{d\times d},$$

where $\mathrm{SO}(d)$ is the set of $(d \times d)$-orthogonal matrices with determinant 1 (rotations if $d = 2$ or $d = 3$). Furthermore, suppose that every purely compressive or purely expansive deformation costs energy, i.e.,

$$f(\alpha\,\mathrm{Id}) > f(\mathrm{Id}) \qquad \text{for all } \alpha \neq 1, \tag{5.1}$$

which is very reasonable in applications. Then, f cannot be convex: Let us for simplicity assume $d = 2$. Set, for a fixed $\gamma \in (0, 2\pi)$,

$$Q := \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix} \in \mathrm{SO}(2).$$

Then, if f was convex, we would get

$$f((\cos\gamma)\,\mathrm{Id}) \leq \frac{1}{2}\bigl(f(Q) + f(Q^T)\bigr) = f(\mathrm{Id}),$$

contradicting (5.1). Sharper arguments are available, but the essential conclusion is the same: convexity is not suitable for many variational problems originating from continuum mechanics.

This chapter introduces Charles B. Morrey Jr.'s concept of *quasiconvexity*, which remedies the above shortcomings of convexity for vector-valued variational problems. After exploring some basic properties of quasiconvex functions, we show how this concept neatly combines with the Young measure theory from the previous chapter to yield an essentially optimal lower semicontinuity theorem if we assume standard growth bounds. We also take a brief look at some regularity results for quasiconvex variational problems.

5.1 Quasiconvexity

Since it was introduced by Morrey in the 1950s, the following notion has become one of the cornerstones of the modern calculus of variations: A locally bounded Borel-measurable function $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ is called **quasiconvex** if

$$h(A) \leq \fint_{B(0,1)} h(A + \nabla\psi(z))\,\mathrm{d}z \tag{5.2}$$

for all $A \in \mathbb{R}^{m\times d}$ and all $\psi \in \mathrm{W}^{1,\infty}_0(B(0,1); \mathbb{R}^m)$.

Before we come to the mathematical analysis, let us give a physical interpretation of quasiconvexity: For $d = m = 3$ suppose that

$$\mathscr{F}[y] := \int_{B(0,1)} h(\nabla y(x)) \,\mathrm{d}x, \qquad y \in \mathrm{W}^{1,\infty}(B(0,1); \mathbb{R}^3),$$

models the physical energy of an elastically deformed body, whose deformation from the reference configuration $\Omega := B(0,1)$ is given as $y\colon B(0,1) \to \mathbb{R}^3$ (see Section 1.7 for more details on this model). A special class of deformations are the *affine* ones, $a(x) = y_0 + Ax$ for some $y_0 \in \mathbb{R}^3$, $A \in \mathbb{R}^{3\times 3}$. Then, quasiconvexity of f entails that

$$\mathscr{F}[a] = \int_{B(0,1)} h(A) \,\mathrm{d}x \le \int_{B(0,1)} h(A + \nabla\psi(x)) \,\mathrm{d}x = \mathscr{F}[a + \psi]$$

for all $\psi \in \mathrm{W}_0^{1,\infty}(B(0,1); \mathbb{R}^3)$. This means that the affine deformation a is always energetically favorable over the *internally distorted* deformation $a + \psi$, which is very often a reasonable assumption for real materials. This interpretation also holds for any other bounded Lipschitz domain Ω by Lemma 5.2 below.

To justify the name, we also need to convince ourselves that quasiconvexity is indeed a notion of *convexity*: For $A \in \mathbb{R}^{m\times d}$ and $V \in \mathrm{L}^1(B(0,1); \mathbb{R}^{m\times d})$ with $\fint_{B(0,1)} V(x) \,\mathrm{d}x = 0$ define the probability measure $\mu \in \mathscr{M}^1(\mathbb{R}^{m\times d})$ via its action as follows (recall that $\mathscr{M}(\mathbb{R}^{m\times d}) \cong \mathrm{C}_0(\mathbb{R}^{m\times d})^*$ by the Riesz Representation Theorem A.21):

$$\langle h, \mu \rangle := \fint_{B(0,1)} h(A + V(x)) \,\mathrm{d}x \qquad \text{for } h \in \mathrm{C}_0(\mathbb{R}^{m\times d}).$$

This μ is easily seen to be an element of the dual space to $\mathrm{C}_0(\mathbb{R}^{m\times d})$ and in fact μ is a probability measure: For the boundedness we observe $|\langle h, \mu\rangle| \le \|h\|_\infty$, whereas the positivity $\langle h, \mu\rangle \ge 0$ for $h \ge 0$ and the normalization $\langle \mathbb{1}, \mu\rangle = 1$ are clear. The barycenter $[\mu]$ of μ is

$$[\mu] = \langle \mathrm{id}, \mu \rangle = A + \fint_{B(0,1)} V(x) \,\mathrm{d}x = A.$$

Therefore, if h is convex, we get from Jensen's inequality (see Lemma A.18),

$$h(A) = h([\mu]) \le \langle h, \mu \rangle = \fint_{B(0,1)} h(A + V(x)) \,\mathrm{d}x.$$

In particular, (5.2) holds if we set $V(x) := \nabla\psi(x)$ for any $\psi \in \mathrm{W}_0^{1,\infty}(B(0,1); \mathbb{R}^m)$. Thus, we have shown:

Proposition 5.1. *All convex functions $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ are quasiconvex.*

Two basic properties of quasiconvexity are collected in the following lemma.

Lemma 5.2. *The following statements are true:*

(i) In the definition of quasiconvexity we can replace the domain $B(0,1)$ by any bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$.

(ii) If h has p-growth, i.e.,

$$|h(A)| \le M(1+|A|^p), \qquad A \in \mathbb{R}^{m\times d},$$

for some $p \in [1,\infty)$, $M > 0$, then in the definition (5.2) of quasiconvexity we can replace testing with all $\psi \in \mathrm{W}_0^{1,\infty}(\Omega;\mathbb{R}^m)$ by testing with all $\psi \in \mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^m)$.

Proof. Ad (i). To see the first statement, we will prove the following claim: Let $\widetilde{\Omega}$ be a bounded Lipschitz domain. If $\psi \in \mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^m)$ then there exists a map $\tilde{\psi} \in \mathrm{W}_0^{1,p}(\widetilde{\Omega};\mathbb{R}^m)$ such that for all $A \in \mathbb{R}^{m\times d}$ it holds that

$$\fint_{\Omega} h(A+\nabla\psi)\,\mathrm{d}x = \fint_{\widetilde{\Omega}} h(A+\nabla\tilde{\psi})\,\mathrm{d}y \tag{5.3}$$

for all measurable $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$, if one of these integrals exists and is finite. Clearly, for $\widetilde{\Omega} := B(0,1)$ this will imply that the definition of quasiconvexity is independent of the domain.

To see (5.3), take a Vitali cover of $\widetilde{\Omega}$ with rescaled disjoint copies of Ω, see Theorem A.15, i.e.,

$$\widetilde{\Omega} = Z \cup \bigcup_{k=1}^{\infty} \Omega(a_k, r_k), \qquad |Z| = 0,$$

with $a_k \in \Omega$, $r_k > 0$, $\Omega(a_k, r_k) := a_k + r_k\Omega$ $(k \in \mathbb{N})$. Then define

$$\tilde{\psi}(y) := r_k\psi\left(\frac{y-a_k}{r_k}\right) \qquad \text{if } y \in \Omega(a_k, r_k)\ (k \in \mathbb{N}).$$

We compute for any measurable $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$,

$$\begin{aligned}
\int_{\widetilde{\Omega}} h(A+\nabla\tilde{\psi})\,\mathrm{d}y &= \sum_{k=1}^{\infty} \int_{\Omega(a_k,r_k)} h\left(A+\nabla\psi\left(\frac{y-a_k}{r_k}\right)\right)\mathrm{d}y \\
&= \sum_{k=1}^{\infty} r_k^d \int_{\Omega} h(A+\nabla\psi)\,\mathrm{d}x \\
&= \frac{|\widetilde{\Omega}|}{|\Omega|} \int_{\Omega} h(A+\nabla\psi)\,\mathrm{d}x
\end{aligned}$$

since $\sum_k r_k^d |\Omega| = |\widetilde{\Omega}|$. This shows (5.3). In particular, $\tilde{\psi} \in W_0^{1,p}(\widetilde{\Omega}; \mathbb{R}^m)$.

Ad (ii). The second assertion follows since $W_0^{1,\infty}(B(0,1); \mathbb{R}^m)$ is dense in the space $W_0^{1,p}(B(0,1); \mathbb{R}^m)$ and under a p-growth assumption for all $A \in \mathbb{R}^{m \times d}$ the integral functional

$$\psi \mapsto \fint_\Omega h(A + \nabla\psi) \, dx, \qquad \psi \in W_0^{1,p}(B(0,1); \mathbb{R}^m),$$

is well-defined and $W^{1,p}$-continuous by Pratt's Theorem A.10 (see the proof of Theorem 2.13 for a similar argument). □

An even weaker notion of convexity than quasiconvexity is the following one: A locally bounded Borel-measurable function $h\colon \mathbb{R}^{m \times d} \to \mathbb{R}$ is called **rank-one convex** if it is convex along any rank-one line, that is,

$$h(\theta A + (1-\theta)B) \leq \theta h(A) + (1-\theta)h(B) \tag{5.4}$$

for all $A, B \in \mathbb{R}^{m \times d}$ with $\operatorname{rank}(A - B) \leq 1$ and all $\theta \in (0,1)$. In this context, recall that a matrix $F \in \mathbb{R}^{m \times d}$ has rank one if and only if $F = a \otimes b = ab^T$ for some $a \in \mathbb{R}^m \setminus \{0\}$, $b \in \mathbb{R}^d \setminus \{0\}$. We remark that the local boundedness of h is in fact automatic if (5.4) holds, see, for instance, the proof of Lemma 2.3 in [162].

Proposition 5.3. *If $h\colon \mathbb{R}^{m \times d} \to \mathbb{R}$ is quasiconvex, then it is rank-one convex.*

Proof. Let $A, B \in \mathbb{R}^{m \times d}$ with $B - A = a \otimes n$ for $a \in \mathbb{R}^m \setminus \{0\}$ and $n \in \mathbb{S}^{d-1}$, the unit sphere in $\mathbb{R}^d$. Denote by Q_n a unit-volume cube ($|Q_n| = 1$) centered at the origin and with two faces orthogonal to n. We also let $\theta \in (0,1)$.

Step 1. Set $F := \theta A + (1-\theta)B$ and define the sequence of test functions $u_j \in W_0^{1,\infty}(Q_n; \mathbb{R}^m)$ as follows:

$$u_j(x) := Fx + \frac{1}{j}\varphi_0\bigl(jx \cdot n - \lfloor jx \cdot n \rfloor\bigr)a, \qquad x \in Q_n$$

and

$$\varphi_0(t) := \begin{cases} -(1-\theta)t & \text{if } t \in [0,\theta], \\ \theta t - \theta & \text{if } t \in (\theta, 1], \end{cases}$$

see Figure 2.1 (on p. 30) for φ_0 and Figure 5.1 for u_j. The sequence (u_j) is called a **laminate** in direction n.

We calculate

$$\nabla u_j(x) = \begin{cases} F - (1-\theta)a \otimes n = A & \text{if } jx \cdot n - \lfloor jx \cdot n \rfloor \in (0,\theta), \\ F + \theta a \otimes n = B & \text{if } jx \cdot n - \lfloor jx \cdot n \rfloor \in (\theta, 1). \end{cases}$$

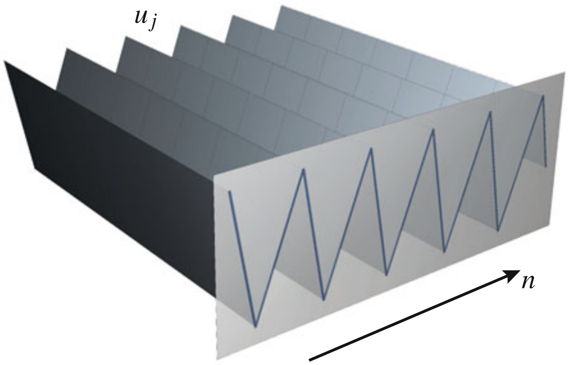


Fig. 5.1 The laminate u_j

Thus,

$$\lim_{j\to\infty} \fint_{Q_n} h(\nabla u_j(x))\,\mathrm{d}x = \theta h(A) + (1-\theta)h(B).$$

Also notice that $u_j \overset{*}{\rightharpoonup} Fx$ in $\mathrm{W}^{1,\infty}$ since φ_0 is uniformly bounded. We will show below that we may replace the sequence (u_j) with a sequence $(v_j) \subset \mathrm{W}^{1,\infty}_{Fx}(Q_n;\mathbb{R}^m)$ (i.e., with the additional property that $v_j|_{\partial Q} = Fx$), but such that still

$$\lim_{j\to\infty} \fint_{Q_n} h(\nabla v_j(x))\,\mathrm{d}x = \theta h(A) + (1-\theta)h(B).$$

By quasiconvexity (also see Lemma 5.2), for all $j \in \mathbb{N}$ it holds that

$$h(F) \le \fint_{Q_n} h(\nabla v_j(z))\,\mathrm{d}z.$$

Thus, we may conclude that

$$h(\theta A + (1-\theta)B) \le \theta h(A) + (1-\theta)h(B)$$

and h is indeed rank-one convex.

Step 2. It remains to construct the sequence (v_j), for which we employ a standard cut-off construction: Take a sequence $(\rho_j) \subset \mathrm{C}^\infty_c(Q_n;[0,1])$ of cut-off functions such that with $G_j := \{\, x \in \Omega : \rho_j(x) = 1 \,\}$ it holds that $|Q_n \setminus G_j| \to 0$ as $j \to \infty$. Set

$$v_{j,k}(x) := \rho_j(x)u_k(x) + (1-\rho_j(x))Fx, \qquad x \in \Omega,$$

which lies in $\mathrm{W}^{1,\infty}(Q_n;\mathbb{R}^m)$ and satisfies $v_{j,k}(x) = Fx$ near ∂Q_n. Also,

$$\nabla v_{j,k}(x) = \rho_j(x)\nabla u_k(x) + (1-\rho_j(x))F + (u_k(x) - Fx) \otimes \nabla \rho_j(x).$$

Since the space $\mathrm{W}^{1,\infty}(Q_n;\mathbb{R}^m)$ embeds compactly into the space $\mathrm{L}^\infty(Q_n;\mathbb{R}^m)$ by the Rellich–Kondrachov Theorem A.28 (or the classical Arzéla–Ascoli theorem), we have $u_k \to Fx$ uniformly. Thus, for fixed j,

$$\limsup_{k\to\infty} \|\nabla v_{j,k}\|_{\mathrm{L}^\infty} \le \|\nabla u_k\|_{\mathrm{L}^\infty} + |F| < \infty$$

because the ∇u_k are uniformly L^∞-bounded. Therefore, we can for every $j \in \mathbb{N}$ choose $k(j) \in \mathbb{N}$ such that $\|\nabla v_{j,k(j)}\|_{\mathrm{L}^\infty}$ is bounded by a constant that is independent of j. As h is assumed to be locally bounded, this implies that there exists a constant $C > 0$ (again independent of j) with

$$\|h(\nabla u_{k(j)})\|_{\mathrm{L}^\infty} + \|h(\nabla v_{j,k(j)})\|_{\mathrm{L}^\infty} \le C.$$

Hence, for $v_j := v_{j,k(j)}$, we may estimate

$$\begin{aligned}\lim_{j\to\infty} \int_{Q_n} |h(\nabla v_j) - h(\nabla u_{k(j)})| \,\mathrm{d}x &\le \lim_{j\to\infty} \int_{Q_n\setminus G_j} |h(\nabla u_{k(j)})| + |h(\nabla v_{j,k(j)})| \,\mathrm{d}x \\ &\le \lim_{j\to\infty} C|Q_n \setminus G_j| \\ &= 0.\end{aligned}$$

This shows that in Step 1 we may indeed replace (u_j) by (v_j). □

Since for $d = 1$ or $m = 1$ rank-one convexity obviously is equivalent to convexity, the same holds true for quasiconvexity. However, quasiconvexity is weaker than classical convexity if $d, m \ge 2$. The determinant function and, more generally, minors are quasiconvex, as will be proved in the next section, but these minors (except for (1×1)-minors) are not convex. The following is a standard example. We will see further non-trivial examples in the following chapters.

Example 5.4. (*Alibert–Dacorogna–Marcellini 1988* [7, 78]) For $d = m = 2$ and $\gamma \in \mathbb{R}$ define

$$h_\gamma(A) := |A|^2\big(|A|^2 - 2\gamma \det A\big), \qquad A \in \mathbb{R}^{2\times 2}.$$

For this function it is known that

- h_γ is convex if and only if $|\gamma| \le \dfrac{2\sqrt{2}}{3} \approx 0.94$,
- h_γ is rank-one convex if and only if $|\gamma| \le \dfrac{2}{\sqrt{3}} \approx 1.15$,
- h_γ is quasiconvex if and only if $|\gamma| \le \gamma_{\mathrm{QC}}$ for some $\gamma_{\mathrm{QC}} \in \left(1, \dfrac{2}{\sqrt{3}}\right]$.

It is currently unknown whether $\gamma_{\mathrm{QC}} = 2/\sqrt{3}$. We do not prove these statements here, see Section 5.3.8 in [76] for the details.

Later, we will see in Example 7.10 that rank-one convexity in general does not imply quasiconvexity. However, for quadratic forms rank-one convexity and quasiconvexity are equivalent, see Problem 5.7.

We end this section with the following observations concerning the growth and continuity properties of rank-one convex (or quasiconvex) functions.

Lemma 5.5. *If $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ is rank-one convex and there are $M > 0$, $p \in [1, \infty)$ such that*

$$h(A) \le M(1 + |A|^p), \qquad A \in \mathbb{R}^{m\times d},$$

then h has p-growth.

Proof. Let $R > 0$ and choose $F_1 \in \mathbb{R}^{m\times d}$ such that $h(F_1) = \inf_{|A|\le R} h(A)$. Let $F_1, \dots, F_{2^{md}}$ be the matrices that are obtained from F_1 by flipping the sign of any number of entries (they do not all have to be distinct). The two matrices of the collection that only differ in the flipped sign at position (i, j) lie on the rank-one line $\mathbb{R}(\mathrm{e}_i \otimes \mathrm{e}_j)$ and average to the zero matrix. Thus, applying the rank-one convexity md times, we have

$$h(0) \le \frac{1}{2^{md}} \sum_{k=1}^{2^{md}} h(F_k).$$

Then,

$$2^{md} h(0) \le (2^{md} - 1) \sup_{|A|\le R} h(A) + \inf_{|A|\le R} h(A),$$

from which we conclude that

$$-h(A) \le M(2^{md} - 1)(1 + R^p) - 2^{md} h(0) \qquad \text{if } |A| \le R.$$

Hence,

$$-h(A) \le \tilde{M}(1 + |A|^p), \qquad A \in \mathbb{R}^{m\times d},$$

for some $\tilde{M} > 0$, and h has been shown to have p-growth. □

Lemma 5.6. *If $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ is rank-one convex, then it is locally Lipschitz continuous. If additionally h has p-growth with growth constant $M > 0$, then*

$$|h(A) - h(B)| \le CM(1 + |A|^{p-1} + |B|^{p-1})|A - B|, \qquad A, B \in \mathbb{R}^{m\times d}, \tag{5.5}$$

where $C = C(d, m) > 0$ is a dimensional constant. In particular, a rank-one convex h with linear growth ($p = 1$) is (globally) Lipschitz continuous.

Proof. For any $F \in \mathbb{R}^{m\times d}$ and $r > 0$, we will prove the quantitative bound

$$\mathrm{lip}(h; B(F, r)) \le \sqrt{\min\{d, m\}} \cdot \frac{\mathrm{osc}(h; \overline{B(F, 6r)})}{3r}, \tag{5.6}$$

where

$$\operatorname{lip}(h;\, B(F,r)) := \sup_{\substack{A,B\in B(F,r)\\ A\neq B}} \frac{|h(A)-h(B)|}{|A-B|}$$

is the Lipschitz constant of h on the ball $B(F,r) \subset \mathbb{R}^{m\times d}$, and

$$\operatorname{osc}(h;\, \overline{B(F,r)}) := \sup_{\substack{A,B\in \overline{B(F,r)}\\ A\neq B}} |h(A)-h(B)|$$

is called the **oscillation** of h on $\overline{B(F,r)}$. By the local boundedness of h, which is part of our definition of rank-one convexity, the oscillation is bounded on every ball. Thus the Lipschitz constant is locally finite.

To show (5.6), let $A, B \in B(F,r)$ and assume first that $\operatorname{rank}(A-B) \leq 1$. Define $M \in \mathbb{R}^{m\times d}$ as the intersection of $\partial B(F,2r)$ with the ray starting at B and going through A. Then, because h is convex along this ray,

$$\frac{|h(A)-h(B)|}{|A-B|} \leq \frac{|h(M)-h(B)|}{|M-B|} \leq \frac{\operatorname{osc}(h;\, \overline{B(F,2r)})}{r} =: \alpha(2r). \tag{5.7}$$

For general $A, B \in B(F,r)$, use the (real) singular value decomposition (see Appendix A.1) to write

$$B - A = \sum_{i=1}^{\min\{d,m\}} \sigma_i P(\mathrm{e}_i \otimes \mathrm{e}_i) Q^T,$$

where $\sigma_i \geq 0$ is the i'th singular value, and $P \in \mathbb{R}^{m\times m}$, $Q \in \mathbb{R}^{d\times d}$ are orthogonal matrices. Set

$$A_k := A + \sum_{i=1}^{k-1} \sigma_i P(\mathrm{e}_i \otimes \mathrm{e}_i) Q^T, \qquad k = 1, \ldots, \min\{d,m\}+1,$$

for which we have $A_1 = A$ and $A_{\min\{d,m\}+1} = B$. We obtain (recall that we are employing the Frobenius norm $|M| = \sqrt{\sum_i \sigma_i(M)^2}$)

$$|A_k - F| \leq |A-F| + \sqrt{\sum_{i=1}^{k-1} \sigma_i^2} \leq |A-F| + |B-A| < 3r$$

and

$$\sum_{k=1}^{\min\{d,m\}} |A_k - A_{k+1}|^2 = \sum_{k=1}^{\min\{d,m\}} \sigma_k^2 = |A-B|^2.$$

Applying (5.7) to $A_k, A_{k+1} \in B(F, 3r)$, $k = 1, \ldots, \min\{d, m\}$, we get

$$\begin{aligned}
|h(A) - h(B)| &\le \sum_{k=1}^{\min\{d,m\}} |h(A_k) - h(A_{k+1})| \\
&\le \alpha(6r) \sum_{k=1}^{\min\{d,m\}} |A_k - A_{k+1}| \\
&\le \alpha(6r)\sqrt{\min\{d, m\}} \cdot \left[\sum_{k=1}^{\min\{d,m\}} |A_k - A_{k+1}|^2 \right]^{1/2} \\
&= \alpha(6r)\sqrt{\min\{d, m\}} \cdot |A - B|.
\end{aligned}$$

This is (5.6).

If we additionally assume that h has p-growth, then

$$\operatorname{osc}(h; \overline{B(0, R)}) \le M(1 + R^p), \qquad R > 0,$$

and so, with $F := 0$, $r := \max\{|A|, |B|\}$ the estimate (5.5) follows from (5.6). $\square$

Remark 5.7. An improved argument (see Lemma 2.2 in [33]), where one orders the singular values in a favorable way, allows one to establish the better estimate

$$\operatorname{lip}(h; B(F, r)) \le \sqrt{\min\{d, m\}} \cdot \frac{\operatorname{osc}(h; \overline{B(F, 2r)})}{r}.$$

5.2 Null-Lagrangians

The determinant is quasiconvex, but it is only one representative of a larger class of canonical examples of quasiconvex, but not convex, functions: In this section, we will investigate the properties of minors (subdeterminants) as integrands. Let for $r \in \{1, 2, \ldots, \min\{d, m\}\}$,

$$I \in P(m, r) := \left\{ (i_1, i_2, \ldots, i_r) \in \{1, \ldots, m\}^r \;:\; i_1 < i_2 < \cdots < i_r \right\}$$

and $J \in P(d, r)$ be **ordered multi-indices**. Then, a $(r \times r)$**-minor** $M\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ is a function of the form

$$M(A) = M^I_J(A) := \det\bigl(A^I_J\bigr),$$

where A^I_J is the $(r \times r)$-matrix consisting of the I-rows and J-columns of A; the number r is called the **rank** of the minor M.

The first result of this section shows that all minors are **null-Lagrangians**, which by definition is the class of integrands $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ such that $\int_\Omega h(\nabla u)\,\mathrm{d}x$ only depends on the boundary values of u.

Lemma 5.8. *Let $M\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ be an $(r\times r)$-minor, $r \in \{1,\ldots,\min\{d,m\}\}$. If $u, v \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$, $p\in[r,\infty]$, with $u - v \in \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^m)$, then*

$$\int_\Omega M(\nabla u(x))\,dx = \int_\Omega M(\nabla v(x))\,dx.$$

Proof. In all of the following we will assume that u, v are smooth and $\operatorname{supp}(u-v) \Subset \Omega$, which can be achieved by approximation and a cut-off procedure, see Theorem A.29. We also need the fact that taking the minor M of the gradient commutes with strong convergence, i.e., the strong continuity of $u \mapsto M(\nabla u)$ in $\mathrm{W}^{1,p}$ for $p \geq r$; this follows by Hadamard's inequality $|M(A)| \leq |A|^r$ and Pratt's Theorem A.10 (see the proof of Lemma 2.16 for a similar argument).

All minors of rank one are just the entries of the matrix and the result follows from the Gauss–Green theorem,

$$\int_\Omega \nabla u\,\mathrm{d}x = -\int_{\partial\Omega} u\cdot n\,\mathrm{d}\mathcal{H}^{d-1} = -\int_{\partial\Omega} v\cdot n\,\mathrm{d}\mathcal{H}^{d-1} = \int_\Omega \nabla v\,\mathrm{d}x$$

since $\operatorname{supp}(u-v) \Subset \Omega$. Here, $\mathcal{H}^{d-1}$ is the $(d-1)$-dimensional surface (Hausdorff) measure on $\partial\Omega$ and n is the unit inner normal on $\partial\Omega$.

For higher-rank minors, the crucial observation is that minors of gradients can be written as *divergences*, which we will establish below. So, if $M(\nabla u) = \operatorname{div} G(u, \nabla u)$, then, since $\operatorname{supp}(u-v)\Subset\Omega$,

$$\begin{aligned}\int_\Omega M(\nabla u)\,\mathrm{d}x &= -\int_{\partial\Omega} G(u,\nabla u)\cdot n\,\mathrm{d}\mathcal{H}^{d-1}\\ &= -\int_{\partial\Omega} G(v,\nabla v)\cdot n\,\mathrm{d}\mathcal{H}^{d-1}\\ &= \int_\Omega M(\nabla v)\,\mathrm{d}x\end{aligned}$$

and the result follows.

We first consider the physically most relevant cases $d = m \in \{2,3\}$.

For $d = m = 2$ and $u = (u^1, u^2)^T$, the only second-order minor is the Jacobian determinant and we easily see from the fact that second derivatives of smooth maps commute that

$$\begin{aligned}\det\nabla u &= \partial_1 u^1 \partial_2 u^2 - \partial_2 u^1 \partial_1 u^2\\ &= \partial_1(u^1\partial_2 u^2) - \partial_2(u^1\partial_1 u^2)\\ &= \operatorname{div}\big(u^1\partial_2 u^2, -u^1\partial_1 u^2\big).\end{aligned}$$

For $d = m = 3$, consider a second-order minor $M^{\neg k}_{\neg l}(A)$, i.e., the determinant of A after deleting the k'th row and l'th column. Then, analogously to the situation in two dimensions, we get, using *cyclic* indices $k, l \in \{1, 2, 3\}$,

$$\begin{aligned} M^{\neg k}_{\neg l}(\nabla u) &= \partial_{l+1} u^{k+1} \partial_{l+2} u^{k+2} - \partial_{l+2} u^{k+1} \partial_{l+1} u^{k+2} \\ &= \partial_{l+1}(u^{k+1} \partial_{l+2} u^{k+2}) - \partial_{l+2}(u^{k+1} \partial_{l+1} u^{k+2}). \end{aligned} \tag{5.8}$$

For the three-dimensional Jacobian determinant, we will show

$$\det \nabla u = \sum_{l=1}^{3} \partial_l u^1 \cdot (\operatorname{cof} \nabla u)^1_l = \sum_{l=1}^{3} \partial_l \big(u^1 (\operatorname{cof} \nabla u)^1_l\big), \tag{5.9}$$

where we recall that $(\operatorname{cof} A)^k_l = (-1)^{k+l} M^{\neg k}_{\neg l}(A)$. To see this, use the Cramer formula $(\det A)I = A(\operatorname{cof} A)^T$, which holds for any square matrix A, to get

$$\det \nabla u = \sum_{l=1}^{3} \partial_l u^1 \cdot (\operatorname{cof} \nabla u)^1_l.$$

Then, (5.9) follows from the **Piola identity**

$$\operatorname{div} \operatorname{cof} \nabla u = 0, \tag{5.10}$$

which can be verified directly from the expression (5.8) for $M^{\neg k}_{\neg l}(\nabla u)$.

For general dimensions d, m we use the notation of differential forms to tame the multilinear algebra involved in the proof (this is not absolutely necessary, one can also argue in an elementary way by induction, but this is quite cumbersome). So, let M be an $(r \times r)$-minor. Reordering $x^1, \ldots, x^d$ and $u^1, \ldots, u^m$, we can assume without loss of generality that M is a *principal* minor, i.e., M is the determinant of the top-left $(r \times r)$-submatrix. Then,

$$\begin{aligned} M(\nabla u)\, dx^1 \wedge \cdots \wedge dx^d &= du^1 \wedge \cdots \wedge du^r \wedge dx^{r+1} \wedge \cdots \wedge dx^d \\ &= d(u^1 \wedge du^2 \wedge \cdots \wedge du^r \wedge dx^{r+1} \wedge \cdots \wedge dx^d). \end{aligned}$$

Thus, the general Stokes theorem gives

$$\begin{aligned} \int_\Omega M(\nabla u)\, dx^1 \wedge \cdots \wedge dx^d &= \int_\Omega d(u^1 \wedge du^2 \wedge \cdots \wedge du^r \wedge dx^{r+1} \wedge \cdots \wedge dx^d) \\ &= \int_{\partial\Omega} u^1 \wedge du^2 \wedge \cdots \wedge du^r \wedge dx^{r+1} \wedge \cdots \wedge dx^d. \end{aligned}$$

Therefore, $\int_\Omega M(\nabla u)\, dx^1 \wedge \cdots \wedge dx^d$ only depends on the values of u around $\partial\Omega$. □

As an immediate consequence, we have:

Corollary 5.9. *All $(r \times r)$-minors $M : \mathbb{R}^{m\times d} \to \mathbb{R}$ are* ***quasiaffine****, that is, both M and $-M$ are quasiconvex.*

Proof. Let $F \in \mathbb{R}^{m\times d}$ and let $\psi \in \mathrm{W}_0^{1,\infty}(B(0,1);\mathbb{R}^m)$. Then, by the preceding lemma,

$$M(F) = \fint_{B(0,1)} M(F + \nabla\psi(z))\,\mathrm{d}z.$$

This already implies the claim. □

Minors also enjoy a surprising weak continuity property:

Lemma 5.10. *Let $M : \mathbb{R}^{m\times d} \to \mathbb{R}$ be an $(r \times r)$-minor, $r \in \{1,\ldots,\min\{d,m\}\}$, and let $(u_j) \subset \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$, where $p \in (r,\infty]$. If*

$$u_j \rightharpoonup u \quad \text{in } \mathrm{W}^{1,p} (\overset{*}{\rightharpoonup} \text{ in } \mathrm{L}^\infty \text{ if } p = \infty),$$

then

$$M(\nabla u_j) \rightharpoonup M(\nabla u) \quad \text{in } \mathrm{L}^{p/r} (\overset{*}{\rightharpoonup} \text{ if } p = \infty).$$

Proof. We will only prove this lemma in the case $p < \infty$, $d = m \in \{2,3\}$, where we employ the special structure of minors as divergences, as exhibited in the proof of Lemma 5.8.

Let $M_{\neg l}^{\neg k}$ be a (2×2)-minor in three dimensions; in two dimensions there is only one (2×2)-minor, the determinant, but we still use the same notation. We rely on (5.8) to observe that with cyclic indices $k, l \in \{1,2,3\}$,

$$\int_\Omega M_{\neg l}^{\neg k}(\nabla u_j)\psi\,\mathrm{d}x = -\int_\Omega (u_j^{k+1}\partial_{l+2}u_j^{k+2})\partial_{l+1}\psi - (u_j^{k+1}\partial_{l+1}u_j^{k+2})\partial_{l+2}\psi\,\mathrm{d}x$$

for all $\psi \in \mathrm{C}_c^\infty(\Omega)$ and then by density also for all $\psi \in \mathrm{L}^{p/2}(\Omega)^* \cong \mathrm{L}^{p/(p-2)}(\Omega)$. Since $u_j \rightharpoonup u$ in $\mathrm{W}^{1,p}$, we have $u_j \to u$ in L^p. The above expressions under the integral consists of products of one L^p-strongly and one L^p-weakly continuous factor as well as a fixed $\mathrm{L}^{p/(p-2)}$-function. Hence by Hölder's inequality, the integral converges as $j \to \infty$ to

$$\int_\Omega M_{\neg l}^{\neg k}(\nabla u)\psi\,\mathrm{d}x.$$

For $d = m = 3$, we additionally need to consider the determinant. However, as a consequence of the above argument in two dimensions, $\operatorname{cof}\nabla u_j \rightharpoonup \operatorname{cof}\nabla u$ in $\mathrm{L}^{p/2}$. Then, (5.9) implies

$$\int_\Omega \det\nabla u_j\,\psi\,\mathrm{d}x = -\sum_{l=1}^{3}\int_\Omega \big[u_j^1(\operatorname{cof}\nabla u_j)_l^1\big]\partial_l\psi\,\mathrm{d}x$$

for all $\psi \in L^{p/3}(\Omega)^* \cong L^{p/(p-3)}(\Omega)$. By a similar reasoning as before this expression converges to

$$-\sum_{l=1}^{3} \int_{\Omega} \big[u^1 (\operatorname{cof} \nabla u)_l^1\big] \partial_l \psi \, dx = \int_{\Omega} \det \nabla u \, \psi \, dx.$$

In the general case, one proceeds by induction, see Problem 5.8. □

It can also be shown that *any* quasiaffine function can be written as an affine function of all the minors. This characterization of quasiaffine functions is due to Ball [25], a different proof (also including further characterizing statements) can be found in Theorem 5.20 of [76].

5.3 A Jensen-Type Inequality for Gradient Young Measures

The connection between Young measure theory and quasiconvexity is furnished by the following *Jensen-type inequality*:

Lemma 5.11. *Let $\nu \in \mathbf{GY}^p(B(0,1); \mathbb{R}^{m \times d})$, where $p \in (1, \infty]$, be a homogeneous gradient Young measure. Then, for all quasiconvex functions $h \colon \mathbb{R}^{m \times d} \to \mathbb{R}$ with p-growth (no growth condition if $p = \infty$) it holds that*

$$h([\nu]) \le \int h \, d\nu. \tag{5.11}$$

Notice that if h is *convex* the conclusion of this lemma is trivially true by the classical Jensen inequality (and also holds for general Young measures, not just the gradient Young measures).

Proof. Set $F := [\nu]$ and let $(u_j) \subset W^{1,p}_{Fx}(B(0,1); \mathbb{R}^m)$ with $\nabla u_j \overset{\mathbf{Y}}{\to} \nu$ and (∇u_j) L^p-equiintegrable (if $p < \infty$), the latter two conditions being realizable by Lemma 4.13. Then, from the definition of quasiconvexity, we get

$$h(F) \le \fint_{\Omega} h(\nabla u_j(x)) \, dx$$

for every $j \in \mathbb{N}$. Passing to the Young measure limit as $j \to \infty$ on the right-hand side, for which we note that the family $\{h(\nabla u_j)\}_j$ is equiintegrable by the growth assumption on h, we arrive at

$$h(F) \le \fint_{\Omega} \int h \, d\nu \, dx = \int h \, d\nu,$$

which is the sought inequality. □

This result will be of crucial importance in proving weak lower semicontinuity in Section 5.5. It is remarkable that the converse also holds, i.e., the validity of (5.11) for all quasiconvex h with p-growth (no growth condition if $p = \infty$) characterizes the class of homogeneous gradient L^p-Young measures in the class of all homogeneous L^p-Young measures. This assertion and its extension to non-homogeneous Young measures is the content of the Kinderlehrer–Pedregal Theorem 7.15.

Corollary 5.12. *Let $p \in (1, \infty]$ and let $\nu \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m \times d})$ be a homogeneous gradient Young measure. Then, for all quasiaffine functions $h \colon \mathbb{R}^{m\times d} \to \mathbb{R}$ with p-growth it holds that*

$$h([\nu]) = \int h \, d\nu.$$

In particular, the preceding corollary applies to the determinant and, more generally, minors, see Corollary 5.9.

5.4 Rigidity for Gradients

Is every Young measure also a gradient Young measure? For *inhomogeneous* Young measures the answer is clearly negative since the barycenter of a gradient Young measure must be a gradient (i.e., curl-free), so the elementary Young measure $\delta[V]$ for V with $\operatorname{curl} V \not\equiv 0$ provides an immediate counterexample.

The question of whether all *homogeneous* Young measures are gradient Young measures is more intricate since then the barycenter is constant and hence trivially a gradient. Still, there are homogeneous Young measures that are not gradient Young measures, but proving that no generating sequence of gradients can be found is often not straightforward. One possibility is to show that there is a quasiconvex function such that the Jensen-type inequality of Lemma 5.11 fails. This strategy is used to good effect in Chapters 8, 9.

Here we consider a more elementary argument: Let $A, B \in \mathbb{R}^{m\times d}$ with $A \neq B$ and $\theta \in (0, 1)$. Consider the homogeneous Young measure

$$\nu := \theta\delta_A + (1-\theta)\delta_B \in \mathbf{Y}^\infty(B(0,1); \mathbb{R}^{m\times d}). \tag{5.12}$$

We know from Example 4.10 that for $\operatorname{rank}(A - B) \leq 1$, ν is a gradient Young measure. The case $\operatorname{rank}(A - B) \geq 2$ can be investigated via the following *rigidity* result.

Theorem 5.13 (Ball–James 1987 [30]). *Let $\Omega \subset \mathbb{R}^d$ be open, bounded, and connected. Suppose also that $A, B \in \mathbb{R}^{m\times d}$.*

*(i) Suppose that $u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ satisfies the **exact two-gradient inclusion***

$$\nabla u \in \{A, B\} \quad \text{a.e. in } \Omega.$$

(a) *If* $\operatorname{rank}(A - B) \geq 2$*, then* $\nabla u = A$ *a.e. or* $\nabla u = B$ *a.e.*
(b) *If* $B - A = a \otimes n$ *for* $a \in \mathbb{R}^m$*,* $n \in \mathbb{S}^{d-1}$ *and* Ω *additionally is assumed to be convex, then there exists a Lipschitz function* $h \colon \mathbb{R} \to \mathbb{R}$ *with* $h' \in \{0, 1\}$ *almost everywhere and a constant vector* $v_0 \in \mathbb{R}^m$ *such that*

$$u(x) = v_0 + Ax + h(x \cdot n)a.$$

(ii) *Assume* $\operatorname{rank}(A - B) \geq 2$ *and suppose that the sequence* $(u_j) \subset \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ *satisfies the* ***approximate two-gradient inclusion****,*

$$\operatorname{dist}(\nabla u_j, \{A, B\}) \to 0 \quad \textit{in measure,}$$

that is, for every $\varepsilon > 0$*,*

$$|\{x \in \Omega : \operatorname{dist}(\nabla u_j(x), \{A, B\} > \varepsilon\}| \to 0 \qquad \textit{as } j \to \infty,$$

and that (u_j) *converges weakly* to a limit* $u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$*. Then,*

$$\nabla u_j \to \nabla u = A \quad \textit{in measure} \qquad \textit{or} \qquad \nabla u_j \to \nabla u = B \quad \textit{in measure.}$$

Proof *Ad (i) (a).* Assume after a translation that $B = 0$ and thus $\operatorname{rank} A \geq 2$. Then, $\nabla u = Ag$ for a scalar function $g \colon \Omega \to \mathbb{R}$. Mollifying u (see Appendix A.5), we may assume that $g \in \mathrm{C}^\infty(\Omega)$.

The idea of the proof is that the curl of ∇u vanishes, expressed as follows: for all $i, j = 1, \ldots, d$ and $k = 1, \ldots, m$, it holds that

$$\partial_i [\nabla u]^k_j = \partial_i \partial_j u^k = \partial_j \partial_i u^k = \partial_j [\nabla u]^k_i.$$

For our special $\nabla u = Ag$, this reads as

$$A^k_j \partial_i g = A^k_i \partial_j g. \tag{5.13}$$

Under the assumption of (i) (a), we claim that $\nabla g = 0$. If otherwise $\xi(x) := \nabla g(x) \neq 0$ for some $x \in \Omega$, then set $a_k(x) := A^k_j / \xi_j(x)$ $(k = 1, \ldots, m)$ for any j such that $\xi_j(x) \neq 0$, which is well-defined by the relation (5.13). We have

$$A^k_j = a_k(x)\xi_j(x), \qquad \text{i.e.,} \qquad A = a(x) \otimes \xi(x).$$

This, however, is impossible if $\operatorname{rank} A \geq 2$. Hence, $\nabla g = 0$ and u is an affine function since Ω is connected; this property is also stable under mollification.

Ad (i) (b). As in (i) (a) we assume $\nabla u = Ag$ $(B = 0)$, where now $A = a \otimes n$. Pick any $v \in \mathbb{R}^n$ that is orthogonal to n. Then,

$$\frac{\mathrm{d}}{\mathrm{d}t} u(x + tv)\bigg|_{t=0} = \nabla u(x)v = [an^T v]g(x) = 0.$$

This implies that u is constant in direction v. As v was an arbitrary vector orthogonal to n and Ω is assumed convex, $u(x)$ can only depend on $x \cdot n$. This implies the claim.

Ad (ii). Assume once more that $B = 0$ and that there exists a (2×2)-minor M with $M(A) \neq 0$. By assumption, for the sets

$$D_j := \left\{ x \in \Omega \,:\, |\nabla u_j(x) - A| < \frac{|A|}{2} \right\},$$

we have

$$\nabla u_j - A\mathbb{1}_{D_j} \to 0 \quad \text{in measure.}$$

Let us also assume that we have selected a subsequence such that

$$\mathbb{1}_{D_j} \overset{*}{\rightharpoonup} \chi \quad \text{in } \mathrm{L}^\infty.$$

In the following we use that for uniformly L^∞-bounded sequences convergence in measure implies weak* convergence in L^∞. Indeed, for any $w \in \mathrm{L}^1(\Omega)$ and any $\varepsilon > 0$ we have

$$\int_\Omega (\nabla u_j - A\mathbb{1}_{D_j}) w \,\mathrm{d}x \leq \|\nabla u_j - A\mathbb{1}_{D_j}\|_{\mathrm{L}^\infty} \int_{\{|\nabla u_j - A\mathbb{1}_{D_j}| > \varepsilon\}} w \,\mathrm{d}x + \varepsilon \|w\|_{\mathrm{L}^1}$$
$$\to 0 + \varepsilon \|w\|_{\mathrm{L}^1} \qquad \text{as } j \to \infty.$$

Since $\varepsilon > 0$ was arbitrary, we obtain $\nabla u_j - A\mathbb{1}_{D_j} \overset{*}{\rightharpoonup} 0$ in L^∞. Thus,

$$\nabla u_j \overset{*}{\rightharpoonup} \nabla u = A\chi \quad \text{in } \mathrm{L}^\infty.$$

Then, by the weak* continuity of minors proved in Lemma 5.10,

$$\text{w*-lim}_{j\to\infty} M(\nabla u_j) = M(A\chi) = M(A)\chi^2.$$

On the other hand, by a similar reasoning as above, we also have $M(\nabla u_j) - M(A)\mathbb{1}_{D_j} \overset{*}{\rightharpoonup} 0$ in L^∞ and thus

$$\text{w*-lim}_{j\to\infty} M(\nabla u_j) = M(A) \cdot \text{w*-lim}_{j\to\infty} \mathbb{1}_{D_j} = M(A)\chi.$$

Since $M(A) \neq 0$, we conclude that $\chi = \chi^2$ and hence that there exists a set $D \subset \Omega$ such that $\chi = \mathbb{1}_D$ and $\nabla u = A\mathbb{1}_D$. Since $\|\mathbb{1}_{D_j}\|_{\mathrm{L}^2} \to \|\mathbb{1}_D\|_{\mathrm{L}^2}$ (this follows from $\mathbb{1}_{D_j} \overset{*}{\rightharpoonup} \mathbb{1}_D$ in L^∞), the Radon–Riesz Theorem A.14 implies that $\mathbb{1}_{D_j} \to \mathbb{1}_D$ in L^2 and then also in measure. Thus, combining the above convergence assertions, we arrive at

$$\nabla u_j \to A\mathbb{1}_D = \nabla u \quad \text{in measure.}$$

Part (i) (a) of the present theorem then implies that $\nabla u = A$ or $\nabla u = 0 = B$ almost everywhere in Ω. As we assumed weak* convergence of our original sequence (u_j), the limit of the selected subsequence is unique and the result holds. □

With this result at hand it is easy to see that our example (5.12) cannot be a gradient Young measure if $\operatorname{rank}(A - B) \geq 2$: Assume that there is a sequence $(u_j) \subset \mathrm{W}^{1,\infty}(B(0,1); \mathbb{R}^m)$ with $\nabla u_j \overset{\mathbf{Y}}{\to} \nu$. By Lemma 4.12 it follows that

$$\operatorname{dist}(\nabla u_j, \{A, B\}) \to 0 \quad \text{in measure.}$$

Then, however, from statement (ii) of the Ball–James rigidity theorem we get $\nabla u_j \to A$ or $\nabla u_j \to B$ in measure, either one of which yields a contradiction (again by Lemma 4.12).

5.5 Lower Semicontinuity

We now turn to the central subject of this chapter, namely to minimization problems of the form

$$\begin{cases} \text{Minimize} \quad \mathscr{F}[u] := \displaystyle\int_\Omega f(x, \nabla u(x)) \,\mathrm{d}x \\ \text{over all} \quad u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m) \text{ with } u|_{\partial\Omega} = g, \end{cases}$$

where $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain, $p \in (1, \infty)$, the Carathéodory integrand $f \colon \Omega \times \mathbb{R}^{m \times d} \to \mathbb{R}$ has p-growth, i.e.,

$$|f(x, A)| \leq M(1 + |A|^p), \qquad (x, A) \in \Omega \times \mathbb{R}^{m \times d},$$

for some $M > 0$, and $g \in \mathrm{W}^{1-1/p,p}(\partial\Omega; \mathbb{R}^m)$ specifies the boundary values. In Chapter 2 we solved this problem in the convex case via the Direct Method, a coercivity result, and, crucially, Tonelli's Lower Semicontinuity Theorem 2.6. In this section, we recycle the Direct Method and the coercivity result, but extend lower semicontinuity to *quasi*convex integrands; some motivation for this was given at the beginning of the chapter.

Let us first consider how we could approach the proof of lower semicontinuity (it should be clear that the proof via Mazur's lemma that we used for the convex lower semicontinuity theorem, does not extend). Suppose that we have a sequence $(u_j) \subset \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ with $u_j \rightharpoonup u$ in $\mathrm{W}^{1,p}$. We want to show the weak lower semicontinuity of our functional $\mathscr{F}$. If we assume that the (norm-bounded) sequence (∇u_j) generates the gradient Young measure $\nu \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m \times d})$, which is true up to selecting a subsequence, and that the sequence of integrands $(f(x, \nabla u_j(x)))_j$ is equiintegrable, then we have a limit:

$$\mathscr{F}[u_j] \to \int_\Omega \int f(x, A) \, \mathrm{d}\nu_x(A) \, \mathrm{d}x \qquad \text{as } j \to \infty.$$

This is useful, because it now suffices to show the Jensen-type inequality

$$\int f(x, A) \, \mathrm{d}\nu_x(A) \ge f(x, \nabla u(x))$$

for almost every $x \in \Omega$, which we have already seen for *homogeneous* gradient Young measures in Lemma 5.11. The only issue is that here we need to "localize" in $x \in \Omega$ and make ν_x a gradient Young measure in its own right. This is accomplished via the fundamental *blow-up technique* (also called the *localization technique*):

Proposition 5.14 *Let $\nu = (\nu_x)_x \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m\times d})$, where $p \in [1, \infty)$, be a gradient Young measure. Then, for almost every $x_0 \in \Omega$ the probability measure ν_{x_0} is a homogeneous gradient Young measure in its own right, $\nu_{x_0} \in \mathbf{GY}^p(B(0, 1); \mathbb{R}^{m\times d})$.*

Proof Take a countable collection $\{\varphi_k \otimes h_k\}_{k\in\mathbb{N}}$ as in Lemma 4.7. Let $x_0 \in \Omega$ be a Lebesgue point of all the functions $x \mapsto \langle h_k, \nu_x\rangle$, $k \in \mathbb{N}$, that is,

$$\lim_{r\downarrow 0} \int_{B(0,1)} \bigl|\langle h_k, \nu_{x_0+ry}\rangle - \langle h_k, \nu_{x_0}\rangle\bigr| \, \mathrm{d}y = 0.$$

By Theorem A.20, almost every point in Ω has this property. Then, at such a point x_0, set

$$v_j^{(r)}(y) := \frac{u_j(x_0 + ry) - [u_j]_{B(x_0,r)}}{r}, \qquad y \in B(0, 1),$$

where $[u]_{B(x_0,r)} := \fint_{B(x_0,r)} u \, \mathrm{d}x$. We get

$$\begin{aligned}\int_{B(0,1)} \varphi_k(y) h_k(\nabla v_j^{(r)}(y)) \, \mathrm{d}y &= \int_{B(0,1)} \varphi_k(y) h_k(\nabla u_j(x_0 + ry)) \, \mathrm{d}y \\ &= \frac{1}{r^d} \int_{B(x_0,r)} \varphi_k\Bigl(\frac{x - x_0}{r}\Bigr) h_k(\nabla u_j(x)) \, \mathrm{d}x\end{aligned}$$

after a change of variables. Letting first $j \to \infty$ and then $r \downarrow 0$, we obtain

$$\begin{aligned}\lim_{r\downarrow 0} \lim_{j\to\infty} \int_{B(0,1)} \varphi_k(y) h_k(\nabla v_j^{(r)}(y)) \, \mathrm{d}y &= \lim_{r\downarrow 0} \frac{1}{r^d} \int_{B(x_0,r)} \varphi_k\Bigl(\frac{x - x_0}{r}\Bigr) \bigl\langle h_k, \nu_x\bigr\rangle \, \mathrm{d}x \\ &= \lim_{r\downarrow 0} \int_{B(0,1)} \varphi_k(y) \bigl\langle h_k, \nu_{x_0+ry}\bigr\rangle \, \mathrm{d}x \\ &= \int_{B(0,1)} \varphi_k(y) \bigl\langle h_k, \nu_{x_0}\bigr\rangle \, \mathrm{d}y,\end{aligned}$$

where the last convergence follows from the Lebesgue point property of x_0. Moreover,

$$\int_{B(0,1)} |\nabla v_j^{(r)}|^p \, dy = \int_{B(0,1)} |\nabla u_j(x_0 + ry)|^p \, dy = \frac{1}{r^d} \int_{B(x_0,r)} |\nabla u_j(x)|^p \, dx$$

and the last integral is uniformly bounded in j (for fixed r). Denote by $\lambda \in \mathscr{M}^+(\overline{\Omega})$ the weak* limit of the measures $|\nabla u_j|^p \mathscr{L}^d \llcorner \Omega$, which exists after taking a subsequence. If we require of x_0 additionally that

$$\limsup_{r \downarrow 0} \frac{\lambda(\overline{B(x_0, r)})}{r^d} < \infty,$$

which holds at $\mathscr{L}^d$-almost every $x_0 \in \Omega$ (see the Besicovitch Differentiation Theorem A.23), then

$$\limsup_{r \downarrow 0} \lim_{j \to \infty} \int_{B(0,1)} |\nabla v_j^{(r)}|^p \, dy < \infty.$$

Since also $[v_j^{(r)}]_{B(0,1)} = 0$, the Poincaré inequality from Theorem A.26 (ii) yields that there exists a diagonal sequence $w_n := v_{j(n)}^{r(n)}$ ($n \in \mathbb{N}$) that is uniformly bounded in the space $W^{1,p}(B(0,1); \mathbb{R}^m)$ and that is such that for all $k \in \mathbb{N}$,

$$\lim_{n \to \infty} \int_{B(0,1)} \varphi_k(y) h_k(\nabla w_n(y)) \, dy = \int_{B(0,1)} \varphi_k(y) \langle h_k, \nu_{x_0} \rangle \, dy.$$

Therefore, $\nabla w_n \xrightarrow{\mathbf{Y}} \nu_{x_0}$ by Lemma 4.7, where we understand ν_{x_0} as a homogeneous (gradient) Young measure on $B(0,1)$. □

Remark 5.15 The preceding result also remains true for $p = \infty$, but this needs Zhang's Lemma 7.18, which we will prove in Chapter 7. The proof of this fact is the task of Problem 7.9.

Our main weak lower semicontinuity theorem is then a straightforward application of the theory developed so far. The first result of this type is due to Charles B. Morrey, Jr. from 1952 (under additional technical assumptions), but our Young measure approach allows us to prove a fairly general result, which was first established by Acerbi & Fusco (using different methods).

Theorem 5.16 (Morrey 1952 & Acerbi–Fusco 1984 [1, 195]). *Let* $p \in (1, \infty)$ *and let* $f \colon \Omega \times \mathbb{R}^{m \times d} \to [0, \infty)$ *be a Carathéodory integrand with* p*-growth and such that*

$$f(x, \bullet) \text{ is quasiconvex for almost every } x \in \Omega.$$

Then, the functional $\mathscr{F}$ *is weakly lower semicontinuous on* $W^{1,p}(\Omega; \mathbb{R}^m)$.

Proof. Let $(u_j) \subset W^{1,p}(\Omega; \mathbb{R}^m)$ with $u_j \rightharpoonup u$ in $W^{1,p}$. Assume that (∇u_j) generates the gradient Young measure $\nu = (\nu_x)_x \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m \times d})$, for which it holds that $[\nu] = \nabla u$. This is only true up to a (not explicitly labeled) subsequence, but if we can establish the lower semicontinuity for every such subsequence it follows that the result also holds for the original sequence.

From Proposition 4.6 we get

$$\liminf_{j\to\infty} \int_\Omega f(x, \nabla u_j(x))\,\mathrm{d}x \geq \langle\!\langle f, \nu \rangle\!\rangle = \int_\Omega \int f(x, A)\,\mathrm{d}\nu_x(A)\,\mathrm{d}x.$$

Now, for almost every $x \in \Omega$ we can consider ν_x as a homogeneous Young measure in $\mathbf{GY}^p(B(0,1); \mathbb{R}^{m\times d})$ by the blow-up technique from Proposition 5.14. Thus, the Jensen-type inequality from Lemma 5.11 reads

$$\int f(x, A)\,\mathrm{d}\nu_x(A) \geq f(x, \nabla u(x)) \qquad \text{for a.e. } x \in \Omega.$$

Combining, we arrive at

$$\liminf_{j\to\infty} \mathscr{F}[u_j] \geq \mathscr{F}[u],$$

which is what we wanted to show. □

Regarding the question of lower semicontinuity for non-positive integrands, see Problem 5.6.

At this point it is worthwhile to reflect on the role of Young measures in the proof of the preceding result, namely that they allowed us to split the argument into two parts: First, we passed to the (lower) limit in the functional via the Young measure. Second, we established a Jensen-type inequality, which then yielded the lower semicontinuity inequality. It is remarkable that the Young measure preserves exactly the right amount of information to serve as an intermediate object.

We can now sum up and prove the existence of a solution for our minimization problem:

Theorem 5.17. *Let $f\colon \Omega \times \mathbb{R}^{m\times d} \to [0,\infty)$ be a Carathéodory integrand such that*

(i) f has p-growth, where $p \in (1,\infty)$;
(ii) f satisfies the p-coercivity estimate $\mu|A|^p \leq f(x, A)$ for some $\mu > 0$;
(iii) f is quasiconvex in its second argument.

Then, the associated functional $\mathscr{F}$ has a minimizer over $\mathrm{W}^{1,p}_g(\Omega; \mathbb{R}^m)$, where $g \in \mathrm{W}^{1-1/p,p}(\partial\Omega; \mathbb{R}^m)$.

Proof. This follows directly by combining the Direct Method from Theorem 2.3 with the coercivity result in Proposition 2.5 and Morrey's Theorem 5.16. □

The following result shows that quasiconvexity is also necessary for weak lower semicontinuity; we only state and show this for x-independent integrands, but we note that it also holds for x-dependent integrands by a localization argument.

Proposition 5.18. *Let $f\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ be continuous and have p-growth. If the associated functional*

$$\mathscr{F}[u] := \int_\Omega f(\nabla u(x))\,dx, \qquad u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m),$$

is weakly lower semicontinuous with or without fixed boundary values, then f is quasiconvex.

Proof. We may assume that $B(0,1) \Subset \Omega$; otherwise we can translate and rescale the domain. Let $A \in \mathbb{R}^{m\times d}$ and $\psi \in \mathrm{W}^{1,\infty}_0(B(0,1);\mathbb{R}^m)$. We need to show

$$f(A) \le \fint_{B(0,1)} f(A + \nabla\psi(z))\,\mathrm{d}z.$$

Take for every $j \in \mathbb{N}$ a Vitali cover of $B(0,1)$ consisting of disjoint balls, see Theorem A.15,

$$B(0,1) = Z^{(j)} \cup \bigcup_{k=1}^{\infty} B(a_k^{(j)}, r_k^{(j)}), \qquad |Z^{(j)}| = 0,$$

with $a_k^{(j)} \in B(0,1)$, $0 < r_k^{(j)} \le 1/j$ ($k \in \mathbb{N}$). Also fix a smooth function $h\colon \Omega \setminus B(0,1) \to \mathbb{R}^m$ with $h(x) = Ax$ for $x \in \partial B(0,1)$ and $h|_{\partial\Omega}$ equal to the prescribed boundary values if there are any. Define

$$u_j(x) := \begin{cases} Ax + r_k^{(j)}\psi\left(\dfrac{x - a_k^{(j)}}{r_k^{(j)}}\right) & \text{if } x \in B(a_k^{(j)}, r_k^{(j)})\ (k \in \mathbb{N}), \\ h(x) & \text{if } x \in \Omega \setminus B(0,1), \end{cases} \qquad x \in \Omega.$$

Then, since ψ is uniformly bounded, it is not hard to see that $u_j \rightharpoonup u$ in $\mathrm{W}^{1,p}$ for

$$u(x) = \begin{cases} Ax & \text{if } x \in B(0,1), \\ h(x) & \text{if } x \in \Omega \setminus B(0,1), \end{cases} \qquad x \in \Omega.$$

Thus, the lower semicontinuity yields, after cancelling the constant part of the functional on $\Omega \setminus B(0,1)$,

$$
\begin{aligned}
\int_{B(0,1)} f(A)\,\mathrm{d}x &\le \liminf_{j\to\infty} \int_{B(0,1)} f(\nabla u_j(x))\,\mathrm{d}x \\
&= \liminf_{j\to\infty} \sum_{k=1}^{\infty} \int_{B(a_k^{(j)},r_k^{(j)})} f\left(A + \nabla\psi\left(\frac{x-a_k^{(j)}}{r_k^{(j)}}\right)\right)\mathrm{d}x \\
&= \liminf_{j\to\infty} \sum_{k=1}^{\infty} (r_k^{(j)})^d \int_{B(0,1)} f(A+\nabla\psi(y))\,\mathrm{d}y \\
&= \int_{B(0,1)} f(A+\nabla\psi(y))\,\mathrm{d}y
\end{aligned}
$$

since $\sum_k (r_k^{(j)})^d = 1$. This is nothing else than quasiconvexity. □

5.6 Integrands with *u*-Dependence

One very useful feature of our Young measure approach is that it allows us to derive a lower semicontinuity result for u-dependent integrands with minimal additional effort. So consider

$$
\begin{cases}
\text{Minimize} & \mathscr{F}[u] := \displaystyle\int_\Omega f(x,u(x),\nabla u(x))\,\mathrm{d}x \\
\text{over all} & u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m) \text{ with } u|_{\partial\Omega} = g,
\end{cases}
$$

where $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain, $p \in (1,\infty)$, and the Carathéodory integrand $f\colon \Omega\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\to\mathbb{R}$ satisfies the p-growth bound

$$
|f(x,v,A)| \le M(1+|v|^p+|A|^p), \qquad (x,v,A)\in\Omega\times\mathbb{R}^m\times\mathbb{R}^{m\times d}, \tag{5.14}
$$

for some $M > 0$, and $g \in \mathrm{W}^{1-1/p,p}(\partial\Omega;\mathbb{R}^m)$.

The idea is to consider Young measures generated by the pairs $(u_j, \nabla u_j) \in \mathbb{R}^{m+md}$.

Lemma 5.19. *Let $(u_j) \subset \mathrm{L}^p(\Omega;\mathbb{R}^M)$ and $(V_j) \subset \mathrm{L}^p(\Omega;\mathbb{R}^N)$ be norm-bounded sequences such that for some $u \in \mathrm{L}^p(\Omega;\mathbb{R}^M)$, $\nu \in \mathbf{Y}^p(\Omega;\mathbb{R}^N)$ it holds that*

$$
u_j \to u \quad \textit{pointwise a.e.} \qquad \textit{and} \qquad V_j \overset{\mathbf{Y}}{\to} \nu.
$$

Then, $(u_j, V_j) \overset{\mathbf{Y}}{\to} \mu = (\mu_x) \in \mathbf{Y}^p(\Omega;\mathbb{R}^{M+N})$ with

$$
\mu_x = \delta_{u(x)} \otimes \nu_x \qquad \textit{for a.e. } x\in\Omega,
$$

that is,

$$\int_{\Omega} f(x, u_j(x), V_j(x))\,dx \to \int_{\Omega} \big\langle f(x, u(x), \bullet), \nu_x(A)\big\rangle\,dx \tag{5.15}$$

for all Carathéodory integrands $f : \Omega \times \mathbb{R}^M \times \mathbb{R}^N \to \mathbb{R}$ *satisfying the* p*-growth bound* (5.14).

Proof. By a similar density argument as in the proof of Proposition 5.14, it suffices to show the convergence (5.15) for $f(x, v, A) = \varphi(x)\psi(v)h(A)$, where $\varphi \in \mathrm{C}_0(\Omega)$, $\psi \in \mathrm{C}_0(\mathbb{R}^M)$, $h \in \mathrm{C}_0(\mathbb{R}^N)$. We already know from the assumptions that

$$h(V_j) \overset{*}{\rightharpoonup} \big(x \mapsto \langle h, \nu_x\rangle\big) \quad \text{in } \mathrm{L}^\infty.$$

Furthermore, $\psi(u_j) \to \psi(u)$ almost everywhere and thus (strongly) in L^1 since ψ is bounded. Since the product of an L^∞-weakly* converging sequence and an L^1-strongly converging sequence converges itself weakly* in the sense of measures, we deduce that

$$\int_{\Omega} \varphi(x)\psi(u_j(x))h(V_j(x))\,\mathrm{d}x \to \int_{\Omega} \varphi(x)\psi(u(x))\langle h, \nu_x\rangle\,\mathrm{d}x,$$

which is (5.15) for our special f. This already finishes the proof. □

The trick of the preceding lemma is that in our situation, where $V_j = \nabla u_j \overset{\mathbf{Y}}{\to} \nu \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m\times d})$, it allows us to "freeze" $u(x)$ in the integrand. Then, we can apply the Jensen-type inequality from Lemma 5.11 just as we did in Morrey's Theorem 5.16.

Theorem 5.20 (Acerbi–Fusco 1984 [1]). *Let* $p \in (1, \infty)$ *and let* $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}$ *be a Carathéodory integrand with* p*-growth, i.e.* (5.14) *holds. Assume furthermore that*

$$f(x, v, \bullet) \text{ is quasiconvex for every fixed } (x, v) \in \Omega \times \mathbb{R}^m.$$

Then, the functional $\mathscr{F}$ *corresponding to* f *is weakly lower semicontinuous on* $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$*. If additionally* f *satisfies the* p*-coercivity estimate*

$$\mu|A|^p \le f(x, v, A), \qquad (x, v, A) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times d},$$

for some $\mu > 0$*, then there exists a minimizer of* $\mathscr{F}$ *over* $\mathrm{W}^{1,p}_g(\Omega; \mathbb{R}^m)$*, where* $g \in \mathrm{W}^{1-1/p,p}(\partial\Omega; \mathbb{R}^m)$.

Remark 5.21 It is not difficult to extend the previous theorem to integrands f satisfying the more general upper growth condition

$$0 \le f(x, v, A) \le M(1 + |v|^q + |A|^p), \qquad (x, v, A) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times d},$$

for some $M > 0$ and $q \in [1, p/(d-p))$. By the Sobolev Embedding Theorem A.27, $u_j \to u$ in L^q for all such q and thus Lemma 5.19 and then Theorem 5.20 can be suitably generalized. Also, we may only require quasiconvexity of $f(x, v, \bullet)$ for $(x, v) \in (\Omega \setminus Z) \times \mathbb{R}^m$, where $|Z| \subset \Omega$ is a negligible set.

5.7 Regularity of Minimizers

At the end of Section 3.2 we discussed the failure of regularity for minimizers of vector-valued problems. Upon closer inspection, however, it turns out that in all counterexamples the points where regularity fails form a relatively closed "small" set. This is not a coincidence, as we will see momentarily.

Another issue was that all the regularity theorems discussed so far have required (strong) *convexity* of the integrand. However, our discussion in this chapter has shown that convexity is not a good notion for vector-valued problems. To remedy this, we need a new notion, which will take over from strong convexity in regularity theory: A locally bounded Borel-measurable function $h \colon \mathbb{R}^{m\times d} \to \mathbb{R}$ is called **strongly quasiconvex** if there exists a $\gamma > 0$ such that

$$A \mapsto h(A) - \gamma |A|^2 \quad \text{is quasiconvex.}$$

Equivalently, we may require that

$$\gamma \int_{B(0,1)} |\nabla\psi(z)|^2 \,\mathrm{d}z \le \int_{B(0,1)} h(A + \nabla\psi(z)) - h(A) \,\mathrm{d}z$$

for all $A \in \mathbb{R}^{m\times d}$ and all $\psi \in \mathrm{W}_0^{1,\infty}(B(0,1); \mathbb{R}^m)$.

The most well-known regularity result in this situation is due to Evans (there is significant overlap with work by Acerbi & Fusco [2]):

Theorem 5.22 (Evans 1986 [109]). *Let $f \colon \mathbb{R}^{m\times d} \to \mathbb{R}$ be twice continuously differentiable, strongly quasiconvex, and assume that there exists an $M > 0$ such that*

$$\mathrm{D}^2 f(A)[B, B] \le M|B|^2, \qquad A, B \in \mathbb{R}^{m\times d}.$$

*Let $u \in \mathrm{W}_g^{1,2}(\Omega; \mathbb{R}^m)$ be a minimizer of $\mathscr{F}$ over the set $\mathrm{W}_g^{1,2}(\Omega; \mathbb{R}^m)$, where $g \in \mathrm{W}^{1/2,2}(\partial\Omega; \mathbb{R}^m)$. Then, there exists a relatively closed **singular set** $\Sigma_u \subset \Omega$ with $|\Sigma_u| = 0$ such that*

$$u \in \mathrm{C}^{1,\alpha}_{\mathrm{loc}}(\Omega \setminus \Sigma_u)$$

for all $\alpha \in (0, 1)$.

This theorem is called a *partial regularity* result because the regularity does not hold everywhere. It should be noted that while the scalar regularity theory was essentially a theory for PDEs, and hence applies to *all* solutions of the Euler–Lagrange equations, this is not the case for the present result: There is *no* regularity theory for critical points of the Euler–Lagrange equation for a quasiconvex or even polyconvex (see the next chapter) integral functional. This was shown by Müller & Šverák in 2003 [207] (for the quasiconvex case) and Székelyhidi Jr. in 2004 [264] (for the polyconvex case); we quote the first result later in Theorem 9.15.

We finally remark that sometimes better estimates on the "smallness" of Σ_u than merely $|\Sigma_u| = 0$ are available. In fact, for strongly convex integrands it can be shown that the (Hausdorff-)dimension of the singular set is at most $d - 2$, see Chapter 2 of [137]. In the strongly quasiconvex case much less is known, but at least for minimizers that happen to be $W^{1,\infty}$ it was established in 2007 by Kristensen & Mingione that the dimension of the singular set is strictly less than d, see [167]. Many other questions are open.

Notes and Historical Remarks

The notion of quasiconvexity was first introduced in Morrey's seminal paper [195]. Lemma 5.6 is originally due to Morrey [196]; we follow the presentation in [33]. The results about null-Lagrangians, in particular Lemmas 5.8 and 5.10, go back to Morrey [196] and Ball [25]. The pivotal proof idea that certain combinations of derivatives might have good convergence properties even if the individual derivatives do not, is also the starting point for the theory of *compensated compactness* (see Section 8.8). A more general result on why convexity is inadmissible for realistic problems in nonlinear elasticity can be found in Section 4.8 of [64].

The convexity properties of quadratic forms have received considerable attention because they correspond to *linear* Euler–Lagrange equations. In this case, quasiconvexity and rank-one convexity are the same, see Problem 5.7. Moreover, for quadratic forms, even polyconvexity (see the next chapter) is equivalent to rank-one convexity if $d = 2$ or $m = 2$, but this does not hold for $d, m \geq 3$. These results together with pointers to the literature can be found in Section 5.3.2 of [76].

The result that rank-one convex functions are locally Lipschitz continuous, Lemma 5.6, is well-known for convex functions, see, for example, Corollary 2.4 in [106] and an adapted version for rank-one convex (even separately convex) functions is in Theorem 2.31 of [76]. Our proof with a quantitative bound is from Lemma 2.2 in [33]. A more general version of this statement can be found in Lemma 2.3 of [162].

The Ball–James Rigidity Theorem 5.13 is from [30]. We will see much more general rigidity results in Chapter 8.

It is possible to prove Morrey's Theorem 5.16 without the use of Young measures, see, for instance, Chapter 8 in [76] for such an approach. However, many of the ideas are essentially the same, they are just carried out directly without the Young measure intermediary (which obscures them somewhat). More on lower semicontinuity and Young measures can be found in the book [222].

All results in this chapter are formulated for Carathéodory integrands, but many continue to hold for $f: \Omega \times \mathbb{R}^N \to \mathbb{R}$ that are Borel-measurable and lower semicontinuous in the second argument, so called **normal integrands**, see [39] and [122].

An argument by Kružík [170], which was refined by Müller, shows the curious fact that for a quasiconvex $h: \mathbb{R}^{m\times d} \to \mathbb{R}$ with $m \geq 3$, $d \geq 2$ the function $A \mapsto h(A^T)$ may not be quasiconvex. The proof can be found in Section 4.7 of [203]; it is based on Šverák's example of a rank-one convex function that is not quasiconvex (for the same dimensions as above), which we will present in Example 7.10 in Chapter 7.

For minimization problems where the integrand can take negative values one needs to look carefully at the negative part of the integrand, see Problem 5.6. If the integrand has critical negative growth, then lower semicontinuity only holds if the boundary values are fixed along a sequence or if one imposes *quasiconvexity at the boundary*, see [37] for a recent survey article discussing this topic.

Problems

5.1 For non-convex domains, statement (i) (b) of the Ball–James Rigidity Theorem 5.13 is false. Construct a counterexample.

5.2 Define, with $D := (0,1)^d \subset \mathbb{R}^d$,

$$\mathrm{W}^{1,\infty}_{\mathrm{per}}(D;\mathbb{R}^m) := \left\{ u \in \mathrm{W}^{1,\infty}(\mathbb{R}^d;\mathbb{R}^m) : u(x+\mathrm{e}_i) = u(x),\ x \in \mathbb{R}^d,\ i = 1,\ldots,d \right\}.$$

A locally bounded Borel-measurable function $h: \mathbb{R}^{m\times d} \to \mathbb{R}$ is called **periodic quasiconvex** if

$$h(A) \leq \int_D h(A + \nabla\psi(z))\,\mathrm{d}z \qquad \text{for all } A \in \mathbb{R}^{m\times d} \text{ and all } \psi \in \mathrm{W}^{1,\infty}_{\mathrm{per}}(D;\mathbb{R}^m).$$

Show that periodic quasiconvexity and the usual quasiconvexity are equivalent. *Hint:* Let $\psi \in \mathrm{W}^{1,\infty}_{\mathrm{per}}(D;\mathbb{R}^m)$ and define for $k \in \mathbb{N}$ the function $\psi_k: \mathbb{R}^d \to \mathbb{R}^m$ as

$$\psi_k(x) := \frac{1}{k}\psi(kx), \qquad x \in \mathbb{R}^d.$$

Prove that

$$\int_D h(A + \nabla\psi(z))\,\mathrm{d}z = \int_D h(A + \nabla\psi_k(z))\,\mathrm{d}z \qquad \text{for all } A \in \mathbb{R}^{m\times d},\ k \in \mathbb{N},$$

and that $\psi_k \in \mathrm{W}^{1,\infty}_{\mathrm{per}}(D;\mathbb{R}^m)$. You will also need a cut-off argument close to the boundary ∂D.

5.3 Denote by $B := \overline{B(0,1)}$ the closed unit ball in $\mathbb{R}^d$.

(i) Let $w\colon B \to \partial B$ be smooth (a "retraction"). Use $|w(x)|^2 = 1$ for every $x \in B$ to show that $\det \nabla w = 0$ in B.
(ii) Use the fact that the determinant is a null-Lagrangian to conclude that there exists at least one $x \in \partial\Omega$ with $w(x) \neq x$. *Hint:* Use an argument by contradiction.
(iii) Derive a smooth version of the *Brouwer fixed point theorem*: Let $u\colon B \to B$ be smooth. Then u has a fixed point $x_* \in B$, that is, $u(x_*) = x_*$. *Hint:* Argue by contradiction and consider the ray emanating from $u(x)$ and passing through x for all $x \in \Omega$ and reduce to (ii).
(iv) Extend the proof to also apply to merely continuous u.

5.4 Let $f\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ be Borel-measurable and strongly quasiconvex, that is, there exists a $\gamma > 0$ such that $A \mapsto f(A) - \gamma|A|^2$ is quasiconvex. Assume furthermore that $|f(A)| \le M(1 + |A|^2)$ for some $M > 0$ and all $A \in \mathbb{R}^{m\times d}$. Show that the functional

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

attains its minimum on $\mathrm{W}^{1,2}_{Fx}(\Omega;\mathbb{R}^m)$ for any $F \in \mathbb{R}^{m\times d}$.

5.5 Show that for $\Omega := (-1,1)^2 \subset \mathbb{R}^2$ the functional

$$\mathscr{F}[u] := \int_\Omega \det(\nabla u_j(x))\,\mathrm{d}x$$

is not weakly lower semicontinuous on $\mathrm{W}^{1,2}(\Omega;\mathbb{R}^2)$ by considering the sequence

$$u_j(x,y) := \frac{(1-|x_2|)^j}{\sqrt{j}}\big(\sin(jx),\cos(jx)\big).$$

5.6 Show that we may extend Morrey's Theorem 5.16 to Carathéodory integrands $f\colon \Omega \times \mathbb{R}^{m\times d} \to \mathbb{R}$ that take negative values as long as

$$-M^{-1}|A|^q - M \le f(A) \le M(1+|A|^p), \qquad A \in \mathbb{R}^{m\times d},$$

where $q \in (0,p)$ and $M > 0$. *Hint:* Observe that the family of negative parts $\{f(\nabla u_j)\}_j$ is equiintegrable.

5.7 Prove that every quadratic form $q\colon \mathbb{R}^{m\times d} \to \mathbb{R}$, that is, $q(A) = b(A,A)$ for a bilinear $b\colon \mathbb{R}^{m\times d} \times \mathbb{R}^{m\times d} \to \mathbb{R}$, is quasiconvex if and only if it is rank-one convex. *Hint:* Use Plancherel's identity (A.4).

5.8 Complete the proof of Lemma 5.10 for higher dimensions. *Hint:* Use the multilinear algebra formulation with differential forms and an induction over the dimension.

5.9 Show that a weaker version of Lemma 5.10 is true if $r = p$, where we only have the convergence of the minors in the sense of distributions.

5.10 A locally bounded Borel-measurable function $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ is called $\mathrm{W}^{1,p}$**-closed-quasiconvex** if

$$h(F) \leq \fint_{B(0,1)} h(A)\,\mathrm{d}\nu(A)$$

for all $F \in \mathbb{R}^{m\times d}$ and for all homogeneous $\mathrm{W}^{1,p}$-gradient Young measures $\nu \in \mathbf{GY}^p(B(0,1);\mathbb{R}^{m\times d})$ with $[\nu] = F$. Show that for all continuous h that satisfy the p-growth condition $|h(A)| \leq M(1+|A|^p)$, $\mathrm{W}^{1,p}$-closed-quasiconvexity is equivalent to $\mathrm{W}^{1,p}$-quasiconvexity.

Chapter 6
Polyconvexity

At the beginning of the previous chapter we saw that convexity cannot hold concurrently with frame-indifference (and a mild non-degeneracy condition). Thus, we were led to consider quasiconvex integrands. However, while quasiconvexity is of tremendous importance in the theory of the calculus of variations, Morrey's Theorem 5.16 has one major drawback: we needed to require the p-growth bound

$$|f(x, A)| \le M(1 + |A|^p), \qquad (x, A) \in \Omega \times \mathbb{R}^{m \times d},$$

for some $M > 0$ and $p \in (1, \infty)$. Unfortunately, this is not a realistic assumption for nonlinear elasticity theory because it ignores the requirement that infinite compressions should cost infinite energy, as we saw in Section 1.7. Indeed, realistic integrands for hyperelastic energy functionals have the property that

$$f(A) \to +\infty \quad \text{as} \quad \det A \downarrow 0$$

and

$$f(A) = +\infty \quad \text{if} \quad \det A \le 0.$$

For instance, the family of matrices

$$A_\alpha := \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix}, \qquad \alpha > 0,$$

satisfies $\det A_\alpha \downarrow 0$ as $\alpha \downarrow 0$, but $|A_\alpha|$ remains uniformly bounded. Thus, the above p-growth bound cannot hold.

The question of whether Morrey's Theorem 5.16 for quasiconvex integrands can be extended to integrands with the above growth is currently a major unsolved problem, see [28]. For the time being, we have to confine ourselves to a more restrictive notion of convexity if we want to allow for the above "elastic" growth. This type of

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8_6

convexity was introduced by John M. Ball in [25] and is called *polyconvexity*. Ball's theorem for the first time made it possible to prove the existence of minimizers for a realistic class of stored-energy functionals in nonlinear elasticity theory, including the Mooney–Rivlin and Ogden materials.

We will focus on the three-dimensional theory because it is by far the most physically relevant. This restriction eases the notational burden considerably; however, any number of dimensions can be treated in a similar way, for which we refer to the comprehensive treatment in [76].

After proving Ball's existence theorem, we also briefly discuss the question of injectivity, which is very relevant for applications.

6.1 Polyconvexity

A function $h\colon \mathbb{R}^{3\times 3} \to \mathbb{R}\cup\{+\infty\}$ (now we allow the value $+\infty$) is called **polyconvex** if it can be written in the form

$$h(A) = H(A, \operatorname{cof} A, \det A), \qquad A \in \mathbb{R}^{3\times 3},$$

where $H\colon \mathbb{R}^{3\times 3} \times \mathbb{R}^{3\times 3} \times \mathbb{R} \to \mathbb{R}\cup\{+\infty\}$ is *convex* (as a function on $\mathbb{R}^{3\times 3} \times \mathbb{R}^{3\times 3} \times \mathbb{R} \cong \mathbb{R}^{19}$). Here, $\operatorname{cof} A$ denotes the cofactor matrix as defined in Appendix A.1.

While convexity obviously implies polyconvexity, the converse is clearly false, as the determinant function shows.

Proposition 6.1. *A polyconvex function $h\colon \mathbb{R}^{3\times 3} \to \mathbb{R}$ (not taking the value $+\infty$) is quasiconvex.*

Proof. Let h be as in the definition of polyconvexity. For $A \in \mathbb{R}^{3\times 3}$ and $w \in \mathrm{W}^{1,\infty}_{Ax}(B(0,1);\mathbb{R}^3)$ we get, using Jensen's inequality (see Lemma A.18),

$$\begin{aligned}
\fint_{B(0,1)} h(\nabla w)\,\mathrm{d}x &= \fint_{B(0,1)} H(\nabla w, \operatorname{cof} \nabla w, \det \nabla w)\,\mathrm{d}x \\
&\geq H\biggl(\fint_{B(0,1)} \nabla w\,\mathrm{d}x, \fint_{B(0,1)} \operatorname{cof} \nabla w\,\mathrm{d}x, \fint_{B(0,1)} \det \nabla w\,\mathrm{d}x\biggr) \\
&= H(A, \operatorname{cof} A, \det A) \\
&= h(A),
\end{aligned}$$

where for the penultimate equality we used the fact that minors are null-Lagrangians as proved in Lemma 5.8. □

We will later see in Example 7.7 that the converse of this proposition is not true.

Example 6.2 (*Compressible neo-Hookean materials*). Functions $f\colon \mathbb{R}^{3\times 3} \to \mathbb{R}$ of the form

$$f(A) := a|A|^2 + \Gamma(\det A)$$

with $a > 0$ and $\Gamma \colon \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ convex, are clearly polyconvex.

Example 6.3 (*Compressible Mooney–Rivlin materials*). Functions $f \colon \mathbb{R}^{3\times 3} \to \mathbb{R}$ of the form

$$f(A) := a|A|^2 + b|\operatorname{cof} A|^2 + \Gamma(\det A)$$

with $a, b > 0$ and

$$\Gamma(d) = \begin{cases} \alpha d^2 - \beta \log d & \text{if } d > 0, \\ +\infty. & \text{if } d \le 0, \end{cases}$$

for some $\alpha, \beta > 0$, are polyconvex. This is obvious once we realize that Γ is convex. See [64, 65] for details.

Example 6.4 (*Ogden materials*). Functions $f \colon \mathbb{R}^{3\times 3} \to \mathbb{R} \cup \{+\infty\}$ of the form

$$f(A) := \sum_{i=1}^{M} a_i \operatorname{tr}\bigl[(A^T A)^{\gamma_i/2}\bigr] + \sum_{j=1}^{N} b_j \operatorname{tr} \operatorname{cof}\bigl[(A^T A)^{\delta_j/2}\bigr] + \Gamma(\det A), \qquad A \in \mathbb{R}^{3\times 3},$$

where $M, N \in \mathbb{N}$, $a_i > 0$, $\gamma_i \ge 1$, $b_j > 0$, $\delta_j \ge 1$, and $\Gamma \colon \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is a convex function with $\Gamma(d) \to +\infty$ as $d \downarrow 0$ and $\Gamma(d) = +\infty$ for $d \le 0$, can be shown to be polyconvex, see Problem 6.6. These stored energy functionals correspond to so-called *Ogden materials* and occur in a wide range of elasticity applications, see [64] for details.

It can also be proved that in three dimensions convex functions of certain combinations of the singular values of a matrix are polyconvex, see Problem 6.10.

6.2 Existence of Minimizers

Let $\Omega \subset \mathbb{R}^3$ be a bounded Lipschitz domain. In this section we will prove the existence of a minimizer of the variational problem

$$\begin{cases} \text{Minimize} \quad \mathscr{F}[u] := \displaystyle\int_\Omega f(x, \nabla u(x)) - b(x) \cdot u(x) \,\mathrm{d}x \\ \text{over all} \qquad u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^3) \text{ with } \det \nabla u > 0 \text{ a.e. and } u|_{\partial\Omega} = g, \end{cases} \tag{6.1}$$

where $f \colon \Omega \times \mathbb{R}^{3\times 3} \to \mathbb{R} \cup \{+\infty\}$ is a Carathéodory integrand (with extended-real values, but the definition is analogous) and $f(x, \bullet)$ is polyconvex for almost every $x \in \Omega$. Thus,

$$f(x, A) = F(x, A, \operatorname{cof} A, \det A), \qquad (x, A) \in \Omega \times \mathbb{R}^{3\times 3},$$

for $F\colon \Omega \times \mathbb{R}^{3\times 3} \times \mathbb{R}^{3\times 3} \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ with $F(x, \bullet, \bullet, \bullet)$ jointly convex and continuous for almost every $x \in \Omega$. As the only upper growth assumptions on f we impose

$$\begin{cases} f(x, A) \to +\infty & \text{as} \quad \det A \downarrow 0, \\ f(x, A) = +\infty & \text{if} \quad \det A \le 0. \end{cases}$$

Furthermore, as usual we suppose that $g \in \mathrm{W}^{1-1/p,p}(\partial\Omega; \mathbb{R}^3)$ and $b \in \mathrm{L}^q(\Omega; \mathbb{R}^3)$, where $1/p + 1/q = 1$. The exponent $p \in (1, \infty)$ will remain unspecified for now. Later, when we impose conditions on the *coercivity* of f, we will also specify p.

The first existence result is relatively straightforward:

Theorem 6.5. *If in addition to the above assumptions it holds that*

$$\mu|A|^p \le f(x, A), \qquad (x, A) \in \Omega \times \mathbb{R}^{3\times 3}, \tag{6.2}$$

for some $\mu > 0$ *and* $p \in (3, \infty)$, *then the minimization problem* (6.1) *has at least one solution in the space*

$$\mathscr{A} := \bigl\{\, u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^3) \;:\; \det \nabla u > 0 \text{ a.e. and } u|_{\partial\Omega} = g \,\bigr\}$$

whenever this set is non-empty.

Proof. We employ the usual Direct Method. For a minimizing sequence $(u_j) \subset \mathscr{A}$ for $\mathscr{F}$ we first show that there exists a constant $C > 0$ such that

$$\|\nabla u_j\|_{\mathrm{L}^p}^p \ge \frac{1}{C}\|u_j\|_{\mathrm{W}^{1,p}}^p - C. \tag{6.3}$$

For this, fix $u_0 \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^3)$ with $u_0|_{\partial\Omega} = g$. Then, the Poincaré inequality from Theorem A.26 (i) in conjunction with the elementary inequality $(a + b)^p \le 2^{p-1}(a^p + b^p)$ for $a, b \ge 0$ implies (with a constant $C = C(\Omega, p, u_0) > 0$ that may change from line to line)

$$\begin{aligned} \|\nabla u_j\|_{\mathrm{L}^p}^p &\ge \frac{1}{C}\|\nabla(u_j - u_0)\|_{\mathrm{L}^p}^p - C \\ &\ge \frac{1}{C}\|u_j - u_0\|_{\mathrm{W}^{1,p}}^p - C \\ &\ge \frac{1}{C}\|u_j\|_{\mathrm{W}^{1,p}}^p - C. \end{aligned}$$

This is (6.3).

We then get from the coercivity estimate (6.2) and Young's inequality that for any $\delta > 0$ it holds that

$$\begin{aligned}
&\int_\Omega f(x, \nabla u_j(x)) - b(x) \cdot u_j(x) \,\mathrm{d}x \\
&\qquad \geq \mu \|\nabla u_j\|_{\mathrm{L}^p}^p - \|b\|_{\mathrm{L}^q} \cdot \|u_j\|_{\mathrm{L}^p} \\
&\qquad \geq \frac{\mu}{C} \|u_j\|_{\mathrm{W}^{1,p}}^p - C - \frac{1}{\delta^q q} \|b\|_{\mathrm{L}^q}^q - \frac{\delta^p}{p} \|u_j\|_{\mathrm{W}^{1,p}}^p.
\end{aligned}$$

Choosing $\delta = (p\mu/(2C))^{1/p}$, one derives (for a different constant $C > 0$)

$$\sup_{j\in\mathbb{N}} \|u_j\|_{\mathrm{W}^{1,p}} \leq C\Big(\sup_{j\in\mathbb{N}} \mathscr{F}[u_j] + 1\Big).$$

Thus, we may select a subsequence (not explicitly labeled) such that $u_j \rightharpoonup u_*$ in $\mathrm{W}^{1,p}$.

By Lemma 5.10 and $p > 3$,

$$\begin{aligned}
\det \nabla u_j &\rightharpoonup \det \nabla u_* \quad \text{in } \mathrm{L}^{p/3} \qquad \text{and} \\
\operatorname{cof} \nabla u_j &\rightharpoonup \operatorname{cof} \nabla u_* \quad \text{in } \mathrm{L}^{p/2}.
\end{aligned}$$

Thus, an argument entirely analogous to the proof of the Tonelli–Serrin Theorem 2.6 yields that the main part of $\mathscr{F}$,

$$v \mapsto \int_\Omega f(x, \nabla v) \,\mathrm{d}x = \int_\Omega F(x, \nabla v, \operatorname{cof} \nabla v, \det \nabla v) \,\mathrm{d}x,$$

is weakly lower semicontinuous on $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^3)$. Indeed, for $v_j \rightharpoonup v$ in $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^3)$ set

$$\begin{aligned}
V_j &:= (\nabla v_j, \operatorname{cof} \nabla v_j, \det \nabla v_j), \\
V &:= (\nabla v, \operatorname{cof} \nabla v, \det \nabla v),
\end{aligned}$$

for which it holds that $V_j \rightharpoonup V$ in $\mathrm{L}^p \times \mathrm{L}^{p/2} \times \mathrm{L}^{p/3}$. Then we may argue as in the proof of the Tonelli–Serrin Theorem 2.6 via Mazur's Lemma A.4 to see that

$$\int_\Omega F(x, V(x)) \,\mathrm{d}x \leq \liminf_{j\to\infty} \int_\Omega F(x, V_j(x)) \,\mathrm{d}x.$$

In this context we also note that F is continuous with values in $[0, \infty]$ by assumption. Thus, also using that the second part of $\mathscr{F}$ is weakly continuous by Lemma 2.16,

$$\mathscr{F}[u_*] \leq \liminf_{j\to\infty} \mathscr{F}[u_j] = \inf_{\mathscr{A}} \mathscr{F} < \infty.$$

In particular, $\det \nabla u_* > 0$ almost everywhere and $u_*|_{\partial\Omega} = g$ by the weak continuity of the trace. Hence, $u_* \in \mathscr{A}$ and the proof is finished. □

The preceding theorem's major drawback is that $p > 3$ has to be assumed. In applications in elasticity theory, however, a more realistic form of f is

$$f(A) = \frac{\lambda}{2}(\operatorname{tr} E)^2 + \mu \operatorname{tr} E^2 + \mathrm{O}(|E|^2), \qquad E = \frac{1}{2}(A^T A - I), \tag{6.4}$$

where $\lambda, \mu > 0$ are the *Lamé constants*. Except for the last term $\mathrm{O}(|E|^2)$, which vanishes as $|E| \downarrow 0$, this energy corresponds to a so-called *St. Venant–Kirchhoff material*. It was shown by Ciarlet & Geymonat [65] (also see Theorem 4.10-2 in [64]) that for any such Lamé constants, there exists a polyconvex function f of compressible Mooney–Rivlin form such that (6.4) holds, that is,

$$f(A) = a|A|^2 + b|\operatorname{cof} A|^2 + \Gamma(\det A) + c,$$

with

$$\Gamma(d) = \begin{cases} \alpha d^2 - \beta \log d & \text{if } d > 0, \\ +\infty & \text{if } d \leq 0, \end{cases}$$

where $a, b, \alpha, \beta > 0$ and $c \in \mathbb{R}$. Clearly, such f has only 2-growth in $|A|$. Thus, we need an existence theorem for functions with these growth properties.

The core of a refined lower semicontinuity argument will be an improvement of Lemma 5.10:

Lemma 6.6. *Let $p, q, r \in [1, \infty)$ with*

$$p \geq 2, \qquad \frac{1}{p} + \frac{1}{q} \leq 1, \qquad r \geq 1$$

and assume that the sequence $(u_j) \subset \mathrm{W}^{1,p}(\Omega; \mathbb{R}^3)$ satisfies

$$\begin{cases} u_j \rightharpoonup u & \text{in } \mathrm{W}^{1,p}, \\ \operatorname{cof} \nabla u_j \rightharpoonup H & \text{in } \mathrm{L}^q, \\ \det \nabla u_j \rightharpoonup d & \text{in } \mathrm{L}^r \end{cases}$$

for some $H \in \mathrm{L}^q(\Omega; \mathbb{R}^{3\times 3})$, $d \in \mathrm{L}^r(\Omega)$. Then, $H = \operatorname{cof} \nabla u$ and $d = \det \nabla u$.

Proof. The idea is to use *distributional* versions of the cofactors and determinant of a gradient and to show that they agree with the usual definitions for sufficiently regular functions. To this end we will use the representation of minors as divergences already employed in Lemmas 5.8, 5.10.

Step 1. From (5.8) we get that for all $\varphi = (\varphi^1, \varphi^2, \varphi^3)^T \in \mathrm{C}^1(\Omega; \mathbb{R}^3)$,

$$(\operatorname{cof} \nabla \varphi)^k_l = (-1)^{k+l} \big[\partial_{l+1}(\varphi^{k+1} \partial_{l+2} \varphi^{k+2}) - \partial_{l+2}(\varphi^{k+1} \partial_{l+1} \varphi^{k+2}) \big],$$

where $k, l \in \{1, 2, 3\}$ are cyclic indices. Thus, for all $\psi \in \mathrm{C}^\infty_c(\Omega)$,

$$\begin{aligned}
&\int_{\Omega} (\operatorname{cof} \nabla\varphi)^k_l \psi \, \mathrm{d}x \\
&\quad = -(-1)^{k+l} \int_{\Omega} (\varphi^{k+1} \partial_{l+2} \varphi^{k+2}) \partial_{l+1} \psi - (\varphi^{k+1} \partial_{l+1} \varphi^{k+2}) \partial_{l+2} \psi \, \mathrm{d}x \\
&\quad =: \big\langle (\operatorname{Cof} \nabla\varphi)^k_l, \psi \big\rangle. \qquad (6.5)
\end{aligned}$$

We call the functional $\operatorname{Cof} \nabla\varphi$ the **distributional cofactors**.

To investigate the continuity properties of $\operatorname{Cof} \nabla\varphi$, we consider

$$\mathscr{G}[\varphi] := \int_{\Omega} (\varphi^k \partial_l \varphi^m) \partial_n \psi \, \mathrm{d}x, \qquad \varphi \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^3),$$

for some $k, l, m, n \in \{1, 2, 3\}$ and fixed $\psi \in \mathrm{C}^\infty_c(\Omega)$. We can estimate this using the Hölder inequality as follows:

$$|\mathscr{G}[\varphi]| \le \|\varphi\|_{\mathrm{L}^s} \|\nabla\varphi\|_{\mathrm{L}^p} \|\nabla\psi\|_\infty$$

whenever

$$\frac{1}{s} + \frac{1}{p} \le 1. \qquad (6.6)$$

In particular, this is true for $s = p \ge 2$. Since $\mathrm{C}^1(\Omega; \mathbb{R}^3)$ is dense in $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^3)$, we have that (6.5) also holds for $\varphi \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^3)$.

Moreover, the above bound yields for any sequence $(\varphi_j) \subset \mathrm{W}^{1,p}(\Omega; \mathbb{R}^3)$ that

$$\mathscr{G}[\varphi_j] \to \mathscr{G}[\varphi] \qquad \text{if} \qquad \begin{cases} \varphi_j \to \varphi & \text{in } \mathrm{L}^s, \\ \nabla\varphi_j \rightharpoonup \nabla\varphi & \text{in } \mathrm{L}^p. \end{cases}$$

If $\varphi_j \rightharpoonup \varphi$ in $\mathrm{W}^{1,p}$, then the first convergence on the right-hand side follows from the Rellich–Kondrachov Theorem A.28 if

$$s < \begin{cases} \frac{3p}{3-p} & \text{if } p < 3, \\ \infty & \text{if } p \ge 3. \end{cases} \qquad (6.7)$$

We can always choose s such that it simultaneously satisfies (6.6) and (6.7), as a quick calculation shows. Thus,

$$\big\langle \operatorname{Cof} \nabla\varphi_j, \psi \big\rangle \to \big\langle \operatorname{Cof} \nabla\varphi, \psi \big\rangle \qquad \text{if} \qquad \varphi_j \rightharpoonup \varphi \quad \text{in } \mathrm{W}^{1,p}. \qquad (6.8)$$

Step 2. If $\varphi = (\varphi^1, \varphi^2, \varphi^3)^T \in \mathrm{C}^2(\Omega; \mathbb{R}^3)$, then we know from (5.9) that

$$\det \nabla\varphi = \sum_{l=1}^{3} \partial_l \varphi^1 (\operatorname{cof} \nabla\varphi)^1_l = \sum_{l=1}^{3} \partial_l \big(\varphi^1 (\operatorname{cof} \nabla\varphi)^1_l\big).$$

Thus, we have for all $\psi \in \mathrm{C}_c^\infty(\Omega)$ that

$$\int_\Omega (\det \nabla\varphi)\psi \,\mathrm{d}x = \sum_{l=1}^3 \int_\Omega \partial_l \varphi^1 (\operatorname{cof} \nabla\varphi)_l^1 \psi \,\mathrm{d}x$$
$$= -\sum_{l=1}^3 \int_\Omega \bigl(\varphi^1 (\operatorname{cof} \nabla\varphi)_l^1\bigr)\partial_l \psi \,\mathrm{d}x. \tag{6.9}$$

The key idea now is that by Hölder's inequality the last integral is well-defined and finite if only

$$\varphi \in \mathrm{L}^p(\Omega;\mathbb{R}^3) \quad \text{with} \quad \operatorname{cof} \nabla\varphi \in \mathrm{L}^q(\Omega;\mathbb{R}^3), \quad \text{where} \quad \frac{1}{p} + \frac{1}{q} \le 1.$$

Analogously to the argument for the cofactor matrix, this motivates us to define the **distributional determinant** $\operatorname{Det} \nabla\varphi$ as the linear functional on $\mathrm{C}_c^\infty(\Omega)$ given as

$$\bigl\langle \operatorname{Det} \nabla\varphi, \psi \bigr\rangle := -\sum_{l=1}^3 \int_\Omega \bigl(\varphi^1 (\operatorname{cof} \nabla\varphi)_l^1\bigr)\partial_l \psi \,\mathrm{d}x, \qquad \psi \in \mathrm{C}_c^\infty(\Omega).$$

From (6.9) we see that if $\varphi \in \mathrm{C}^1(\Omega;\mathbb{R}^3)$, then "det = Det", i.e.,

$$\int_\Omega (\det \nabla\varphi)\psi \,\mathrm{d}x = \sum_{l=1}^3 \int_\Omega \partial_l \varphi^1 (\operatorname{cof} \nabla\varphi)_l^1 \psi \,\mathrm{d}x = \bigl\langle \operatorname{Det} \nabla\varphi, \psi \bigr\rangle \tag{6.10}$$

for all $\psi \in \mathrm{C}_c^\infty(\Omega)$. We want to show that this equality remains valid if merely $\varphi \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^3)$ with $\operatorname{cof} \nabla\varphi \in \mathrm{L}^q(\Omega;\mathbb{R}^{3\times 3})$, where $1/p + 1/q \le 1$.

For the moment fix $\psi \in \mathrm{C}_c^\infty(\Omega)$. Define for $\varphi \in \mathrm{C}^1(\Omega;\mathbb{R}^3)$ and $W \in \mathrm{C}^1(\Omega;\mathbb{R}^{3\times 3})$,

$$\mathscr{Z}[\varphi, W] := \sum_{l=1}^3 \int_\Omega \partial_l \varphi^1 W_l^1 \psi + \varphi^1 W_l^1 \partial_l \psi \,\mathrm{d}x.$$

Observe that

$$\mathscr{Z}[\varphi, W] = \sum_{l=1}^3 \int_\Omega W_l^1 \partial_l (\varphi^1 \psi) \,\mathrm{d}x. \tag{6.11}$$

Hölder's inequality furthermore implies

$$|\mathscr{Z}[\varphi, W]| \le C\|\varphi\|_{\mathrm{W}^{1,p}} \|W\|_{\mathrm{L}^q} \|\psi\|_{\mathrm{W}^{1,\infty}} \qquad \text{whenever} \qquad \frac{1}{p} + \frac{1}{q} \le 1.$$

To establish (6.10) for $\varphi \in W^{1,p}(\Omega; \mathbb{R}^3)$ with $\operatorname{cof} \nabla\varphi \in L^q(\Omega; \mathbb{R}^{3\times 3})$, where $1/p + 1/q \le 1$, we need to show $\mathscr{Z}[\varphi, \operatorname{cof} \nabla\varphi] = 0$ for all such φ.

If $v \in C^1(\Omega; \mathbb{R}^3)$, then we get from the Piola identity (5.10) that

$$\operatorname{div} \operatorname{cof} \nabla v = 0.$$

Thus, using (6.11), we have $\mathscr{Z}[\varphi, \operatorname{cof} \nabla v] = 0$ for $\varphi, v \in C^1(\Omega; \mathbb{R}^3)$. Moreover, the map $v \mapsto \operatorname{cof} \nabla v$ is continuous from $W^{1,p}(\Omega; \mathbb{R}^3)$ to $L^1(\Omega; \mathbb{R}^{3\times 3})$ since $|\operatorname{cof} A| \le C|A|^2$ and $p \ge 2$ (one can, for instance, argue using pointwise almost everywhere convergence and Pratt's Theorem). By the density of $C^1(\Omega; \mathbb{R}^3)$ in $W^{1,p}(\Omega; \mathbb{R}^3)$ this continuity then yields

$$\mathscr{Z}[\varphi, \operatorname{cof} \nabla v] = 0 \qquad \text{for } \varphi \in C^1(\Omega; \mathbb{R}^3) \text{ and } v \in W^{1,p}(\Omega; \mathbb{R}^3).$$

On the other hand, $\mathscr{Z}[\,\bullet\,, W]$ is continuous in the first argument with respect to strong convergence in $W^{1,p}$ if $W \in L^q(\Omega; \mathbb{R}^{3\times 3})$ and $1/p + 1/q \le 1$. Thus, another approximation yields that $\mathscr{Z}[\varphi, \operatorname{cof} \nabla v] = 0$ for all $\varphi \in W^{1,p}(\Omega; \mathbb{R}^3)$ and $v \in W^{1,p}(\Omega; \mathbb{R}^3)$ with $\operatorname{cof} \nabla v \in L^q(\Omega; \mathbb{R}^{3\times 3})$. In particular,

$$\mathscr{Z}[\varphi, \operatorname{cof} \nabla\varphi] = 0 \qquad \text{for } \varphi \in W^{1,p}(\Omega; \mathbb{R}^3) \text{ with } \operatorname{cof} \nabla\varphi \in L^q(\Omega; \mathbb{R}^{3\times 3}).$$

Therefore, (6.10) ("det = Det") holds for all such $\varphi \in W^{1,p}(\Omega; \mathbb{R}^3)$.

We see from the definition of the distributional determinant that for a sequence $(\varphi_j) \subset W^{1,p}(\Omega; \mathbb{R}^3)$ we have

$$\langle \operatorname{Det} \nabla\varphi_j, \psi \rangle \to \langle \operatorname{Det} \nabla\varphi, \psi \rangle \qquad \text{if} \qquad \begin{cases} \varphi_j \to \varphi & \text{in } L^s, \\ \operatorname{cof} \nabla\varphi_j \rightharpoonup \operatorname{cof} \nabla\varphi & \text{in } L^q \end{cases}$$

whenever

$$\frac{1}{s} + \frac{1}{q} \le 1. \tag{6.12}$$

For sequences $\varphi_j \rightharpoonup \varphi$ in $W^{1,p}$ the first convergence on the right-hand side follows, as before, from the Rellich–Kondrachov Theorem A.28 if

$$s < \begin{cases} \frac{3p}{3-p} & \text{if } p < 3, \\ \infty & \text{if } p \ge 3. \end{cases} \tag{6.13}$$

However, since we assumed

$$\frac{1}{p} + \frac{1}{q} \le 1,$$

we can always choose s such that it satisfies (6.12) and (6.13) simultaneously, as can be seen by some elementary algebra. Thus,

$$\langle \operatorname{Det} \nabla \varphi_j, \psi \rangle \to \langle \operatorname{Det} \nabla \varphi, \psi \rangle \quad \text{if} \quad \begin{cases} \varphi_j \rightharpoonup \varphi & \text{in } \mathrm{W}^{1,p}, \\ \operatorname{cof} \nabla \varphi_j \rightharpoonup \operatorname{cof} \nabla \varphi & \text{in } \mathrm{L}^q, \end{cases} \tag{6.14}$$

where p, q satisfy the assumptions of the lemma.

Step 3. Assume that, as in the statement of the lemma, we are given a sequence $(u_j) \subset \mathrm{W}^{1,p}(\Omega; \mathbb{R}^3)$ such that for $H \in \mathrm{L}^q(\Omega; \mathbb{R}^{3\times 3})$, $d \in \mathrm{L}^r(\Omega)$ it holds that

$$\begin{cases} u_j \rightharpoonup u & \text{in } \mathrm{W}^{1,p}, \\ \operatorname{cof} \nabla u_j \rightharpoonup H & \text{in } \mathrm{L}^q, \\ \det \nabla u_j \rightharpoonup d & \text{in } \mathrm{L}^r. \end{cases}$$

Then, (6.5), (6.10) imply that for all $\psi \in \mathrm{C}_c^\infty(\Omega)$,

$$\langle \operatorname{Cof} \nabla u_j, \psi \rangle \to \langle H, \psi \rangle \quad \text{and} \quad \langle \operatorname{Det} \nabla u_j, \psi \rangle \to \langle d, \psi \rangle.$$

On the other hand, from (6.8) and (6.5) again,

$$\langle \operatorname{Cof} \nabla u_j, \psi \rangle \to \langle \operatorname{Cof} \nabla u, \psi \rangle = \int_\Omega (\operatorname{cof} \nabla u) \psi \, \mathrm{d}x,$$

whereby

$$\int_\Omega (\operatorname{cof} \nabla u - H) \psi \, \mathrm{d}x = 0$$

for all $\psi \in \mathrm{C}_c^\infty(\Omega)$. The Fundamental Lemma 3.10 then gives immediately

$$\operatorname{cof} \nabla u = H \in \mathrm{L}^q(\Omega; \mathbb{R}^{3\times 3}).$$

Since we have just shown that $\operatorname{cof} \nabla u_j \rightharpoonup \operatorname{cof} \nabla u$ in L^q, (6.10) and (6.14) imply

$$\langle \operatorname{Det} \nabla u_j, \psi \rangle \to \langle \operatorname{Det} \nabla u, \psi \rangle = \int_\Omega (\det \nabla u) \psi \, \mathrm{d}x.$$

Thus, by a similar argument as above,

$$\det \nabla u = d \in \mathrm{L}^r(\Omega).$$

This finishes the proof. □

With this tool at hand, we can now prove the main existence result for integral functionals with polyconvex integrands:

Theorem 6.7 (Ball 1977 [25]). *Let* $p, q, r \in [1, \infty)$ *with*

$$p \geq 2, \qquad \frac{1}{p} + \frac{1}{q} \leq 1, \qquad r > 1$$

such that in addition to the assumptions at the beginning of this section it holds that

$$f(x, A) \geq \mu\big(|A|^p + |\operatorname{cof} A|^q + |\det A|^r\big), \qquad (x, A) \in \Omega \times \mathbb{R}^{3\times 3}, \tag{6.15}$$

for some $\mu > 0$*. Then, the minimization problem* (6.1) *has at least one solution in the space*

$$\mathscr{A} := \Big\{ u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^3) \;:\; \operatorname{cof} \nabla u \in \mathrm{L}^q(\Omega; \mathbb{R}^{3\times 3}), \det \nabla u \in \mathrm{L}^r(\Omega), \\ \det \nabla u > 0 a.e., \text{ and } u|_{\partial\Omega} = g \Big\}$$

whenever this set is non-empty.

Proof. This follows in a completely analogous way to the proof of Theorem 6.5, but now we select a subsequence of a minimizing sequence $(u_j) \subset \mathscr{A}$ such that for some $u_* \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^3)$, $H \in \mathrm{L}^q(\Omega; \mathbb{R}^{3\times 3})$, $d \in \mathrm{L}^r(\Omega)$ we have

$$\begin{cases} u_j \rightharpoonup u_* & \text{in } \mathrm{W}^{1,p}, \\ \operatorname{cof} \nabla u_j \rightharpoonup H & \text{in } \mathrm{L}^q, \\ \det \nabla u_j \rightharpoonup d & \text{in } \mathrm{L}^r, \end{cases}$$

which is possible by the usual weak compactness results in conjunction with the coercivity assumption (6.15). Lemma 6.6 yields

$$\begin{cases} u_j \rightharpoonup u_* & \text{in } \mathrm{W}^{1,p}, \\ \operatorname{cof} \nabla u_j \rightharpoonup \operatorname{cof} \nabla u_* & \text{in } \mathrm{L}^q, \\ \det \nabla u_j \rightharpoonup \det \nabla u_* & \text{in } \mathrm{L}^r, \end{cases}$$

and we may argue as in Theorem 6.5 to conclude that u_* is a minimizer over $\mathscr{A}$. □

Example 6.8. For the example from Section 1.7 we can now show that a minimizer exists for the problem

$$\begin{cases} \text{Minimize} & \mathscr{F}[y] := \displaystyle\int_\Omega W(\nabla y(x)) - b(x) \cdot y(x) \,\mathrm{d}x \\ \text{over all} & y \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^3) \text{ with } \operatorname{cof} \nabla y \in \mathrm{L}^q(\Omega; \mathbb{R}^{3\times 3}), \det \nabla y \in \mathrm{L}^r(\Omega), \\ & \det \nabla y > 0 \text{ a.e., and } y|_{\partial\Omega} = g, \end{cases}$$

if W is any one of the polyconvex integrands exhibited in Examples 6.2, 6.3, or 6.4, $b \in \mathrm{L}^s(\Omega; \mathbb{R}^3)$, and $g \in \mathrm{W}^{1-1/p,p}(\partial\Omega; \mathbb{R}^3)$ with $p, q, r, s \in [1, \infty)$ such that

$$p \geq 2, \qquad \frac{1}{p} + \frac{1}{q} \leq 1, \qquad \frac{1}{p} + \frac{1}{s} \leq 1, \qquad r > 1.$$

This follows from Theorem 6.7 in conjunction with Lemma 2.16 (for the strong continuity of the second part of $\mathscr{F}$).

6.3 Global Injectivity

For reasons of *physical admissibility* we often want to additionally prove that we can find a minimizer that is **injective almost everywhere**, that is, $u\colon \Omega \to \mathbb{R}^3$ is such that

$$\mathscr{H}^0(u^{-1}(x')) = 1 \qquad \text{for a.e. } x' \in u(\Omega),$$

where $\mathscr{H}^0$ is the counting measure. Note that if the deformed configuration has *self-contact*, then we cannot expect full injectivity.

There are several approaches to this delicate question, for example via the *topological degree*. Here, we present a classical argument by Ciarlet & Nečas [67]. It is important to notice that it is only realistic to expect injectivity for $p > d$. For lower exponents, complex effects such as cavitation and (microscopic) fracture have to be considered. This is already indicated by the fact that Sobolev functions in $\mathrm{W}^{1,p}$ for $p \leq d$ are not necessarily continuous.

We will prove the following basic theorem:

Theorem 6.9. *In the situation of Theorem 6.5, in particular $p > 3$, the minimization problem (6.1) has at least one solution in the space*

$$\mathscr{A} := \left\{ u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^3) \;:\; \det \nabla u > 0 \textit{ a.e., } u \textit{ is injective a.e., } u|_{\partial\Omega} = g \right\}$$

whenever this set is non-empty.

Proof. Let $(u_j) \subset \mathscr{A}$ be a minimizing sequence with $u_j \rightharpoonup u_*$ in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^3)$. The existence proof is analogous to that of Theorem 6.5; we only need to show in addition that u_* is injective almost everywhere.

In the following we will use the fact that even if only $v \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^3)$ it holds that

$$\int_\Omega |\det \nabla v| \,\mathrm{d}x = \int_{v(\Omega)} \mathscr{H}^0(v^{-1}(x')) \,\mathrm{d}x',$$

where we denote by $\mathscr{H}^0$ the counting measure. See, for example, [180] or [46] for a proof.

For our choice $p > 3$, the space $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^3)$ embeds continuously into $\mathrm{C}(\Omega;\mathbb{R}^3)$, so we have that

$$u_j \to u_* \quad \text{uniformly.}$$

Let $U \subset \mathbb{R}^3$ be any precompact open set with $u_*(\Omega) \Subset U$ (note that $u_*(\Omega)$ is bounded). Then, $u_j(\Omega) \subset U$ for j sufficiently large by the uniform convergence. Hence, for such j,

$$|u_j(\Omega)| \le |U|.$$

Since $\det \nabla u_j$ converges weakly to $\det \nabla u_*$ in $\mathrm{L}^{p/3}$ by Lemma 5.10 and all u_j are injective almost everywhere by definition of the space $\mathscr{A}$,

$$\begin{aligned}\int_\Omega \det \nabla u_* \, \mathrm{d}x &= \lim_{j\to\infty} \int_\Omega \det \nabla u_j \, \mathrm{d}x \\ &= \lim_{j\to\infty} \int_{u_j(\Omega)} \mathscr{H}^0(u_j^{-1}(x')) \, \mathrm{d}x' \\ &= \lim_{j\to\infty} |u_j(\Omega)| \\ &\le |U|.\end{aligned}$$

Letting $|U| \downarrow |\overline{u_*(\Omega)}| = |u_*(\Omega)|$ (the last equality follows since $|\partial u_*(\Omega)| = 0$, which again is proved in [180]), we get the **Ciarlet–Nečas non-interpenetration condition**

$$\int_\Omega \det \nabla u_* \, \mathrm{d}x \le |u_*(\Omega)|.$$

Then, we have the estimate

$$|u_*(\Omega)| \le \int_{u_*(\Omega)} \mathscr{H}^0(u_*^{-1}(x')) \, \mathrm{d}x' = \int_\Omega \det \nabla u_* \, \mathrm{d}x \le |u_*(\Omega)|,$$

where we used that $\det \nabla u_* > 0$ almost everywhere (which follows as in the proof of Theorem 6.5). Thus,

$$\mathscr{H}^0(u_*^{-1}(x')) = 1 \qquad \text{for a.e. } x' \in u(\Omega),$$

and we have shown the almost everywhere injectivity of u_*. □

Unfortunately, injectivity almost everywhere does not exclude all unphysical examples. For example, a countable dense set may be mapped into one point. Injectivity *everywhere* is a harder problem. A well-known result in this direction is the following:

Theorem 6.10 (Ball 1981 [26]). *In the situation of Theorem 6.5, assume furthermore that*

$$\mu\left(|A|^p + \frac{|\operatorname{cof} A|^p}{(\det A)^{p-1}}\right) \le f(x, A), \qquad (x, A) \in \Omega \times \mathbb{R}^{3\times 3},$$

for some $\mu > 0$ *and* $p \in (3, \infty)$*, and that there exists an injective map* $u_0 \in \mathrm{C}(\overline{\Omega}; \mathbb{R}^3)$ *such that* $u_0(\Omega)$ *is also a bounded Lipschitz domain. Then, the minimization problem* (6.1) *has at least one solution* u_* *in the space*

$$\mathscr{A} := \left\{ u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^3) \;:\; \det \nabla u > 0 \text{ a.e. and } u|_{\partial\Omega} = u_0|_{\partial\Omega} \right\}$$

such that the following assertions hold:

(i) u_* *is a homeomorphism of* Ω *onto* $u_*(\Omega)$*;*
(ii) $u_*(\overline{\Omega}) = u_0(\overline{\Omega})$*;*
(iii) $u_*^{-1} \in \mathrm{W}^{1,p}(u_0(\Omega); \mathbb{R}^3)$*;*
(iv) $\nabla u_*^{-1}(u_*(x)) = (\nabla u_*(x))^{-1}$ *for almost every* $x \in \Omega$*.*

Notes and Historical Remarks

Theorem 6.7 is a refined version of Ball's original result [25] due to Ball, Currie & Olver [29]; also see [30, 31] for further reading. This theorem and its applications to elasticity theory are described in great detail in [64].

Many questions about polyconvex integral functionals remain open to this day. In particular, the regularity of solutions and the validity of the Euler–Lagrange equations are largely unknown in the general case. Note that the regularity theory from the previous chapter is not in general applicable, at least if we do not assume the upper p-growth. These questions are even open for the more restricted situation of nonlinear elasticity theory. See [28] for a survey on the current state of the art and a collection of challenging open problems. We note that since the publication of [28] counterexamples to uniqueness have been found, see [245].

As for the almost injectivity (for $p > 3$), this is in fact sometimes automatic, as shown by Ball [26], but the arguments do not apply to all situations. Tang [266] extended this to $p > 2$, but since then one has to deal with non-continuous functions, it is not even obvious how to define $u(\Omega)$.

Theorem 6.10 is from [26]. More general results can be found in [248]. The questions of injectivity, invertibility, and regularity are intimately connected with *cavitation* and *fracture* phenomena, see, for instance, [204] and the recent [148], which also contains a large bibliography.

Finally, we mention in passing the alternative so-called intrinsic approach to elasticity, as pioneered by Ciarlet, see [66].

Problems

6.1 Define $h \colon \mathbb{R}^{3\times 3} \to \mathbb{R}$ via

$$h(A) := \left(|A|^6 + |\operatorname{cof} A|^6\right)^{1/2} + g(\det A), \qquad A \in \mathbb{R}^{3\times 3},$$

where $g\colon \mathbb{R} \to [0, \infty]$ is convex and continuous. Prove that h is polyconvex.

6.2 Show that the function

$$h(A) := \begin{cases} |A| \ln(1 + |A|) + \dfrac{1}{\det A} & \text{if } \det A > 0, \\ +\infty & \text{if } \det A \leq 0, \end{cases} \qquad A \in \mathbb{R}^{3\times 3},$$

is polyconvex.

6.3 Show that for $u \in (\mathrm{W}^{1,3} \cap \mathrm{C}^2)(\Omega; \mathbb{R}^3)$ it holds that

$$\det \nabla u(x) = \sum_{l=1}^{3} \partial_l (u^1 (\operatorname{cof} \nabla u(x))_l^1), \qquad x \in \Omega.$$

Hint: Use the Piola identity.

6.4 Let $u_j, u \in (\mathrm{W}^{1,3} \cap \mathrm{C}^2)(\Omega; \mathbb{R}^3)$, $j \in \mathbb{N}$, with

$$\operatorname{cof} \nabla u_j,\ \operatorname{cof} \nabla u \in \mathrm{L}^3(\Omega; \mathbb{R}^{3\times 3}).$$

Prove that if

$$u_j \rightharpoonup u \quad \text{in } \mathrm{W}^{1,3} \qquad \text{and} \qquad \operatorname{cof} \nabla u_j \rightharpoonup \operatorname{cof} \nabla u \quad \text{in } \mathrm{L}^3,$$

then $\det \nabla u_j, \det \nabla u \in \mathrm{L}^{3/2}(\Omega)$ and $\det \nabla u_j \rightharpoonup \det \nabla u$ in $\mathrm{L}^{3/2}$.

6.5 In this problem we will construct a rank-one convex function that is not polyconvex.

(i) Find $A_1, A_2, A_3 \in \mathbb{R}^{2\times 2}$ and $\theta_1, \theta_2, \theta_3 \in (0, 1)$ such that simultaneously

- (a) $\theta_1 + \theta_2 + \theta_3 = 1$;
- (b) $\displaystyle\sum_{k=1}^{3} \theta_i \det A_i = \det \left[\sum_{k=1}^{3} \theta_i A_i \right]$;
- (c) $\det(A_1 - A_2) \neq 0$, $\det(A_1 - A_3) \neq 0$, $\det(A_2 - A_3) \neq 0$;
- (d) $\displaystyle\sum_{k=1}^{3} \theta_i A_i \notin \{A_1, A_2, A_3\}$.

(ii) Define with the A_1, A_2, A_3 from (i) the function $f\colon \mathbb{R}^{2\times 2} \to \mathbb{R} \cup \{+\infty\}$ as

$$f(A) := \begin{cases} 0 & \text{if } A \in \{A_1, A_2, A_3\}, \\ +\infty & \text{otherwise.} \end{cases}$$

Show that f is rank-one convex (extending the definition of rank-one convexity in a suitable way to $(\mathbb{R} \cup \{+\infty\})$-valued functions).

(iii) Show that f is not polyconvex, i.e., there exists no convex function $F\colon \mathbb{R}^{2\times 2} \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ such that

$$f(A) = F(A, \det A), \qquad A \in \mathbb{R}^{2\times 2}.$$

6.6 Show that the function f from Example 6.4 (Ogden materials) is polyconvex. *Hint:* Use Cramer's rule to see that $\operatorname{cof}(AB) = \operatorname{cof}(A)\operatorname{cof}(B)$.

6.7 Set

$$\mathrm{GL}(d) := \left\{ A \in \mathbb{R}^{d\times d} \ \mathbf{:}\ \det A \neq 0 \right\} \qquad \text{and}$$
$$\mathrm{GL}^+(d) := \left\{ A \in \mathbb{R}^{d\times d} \ \mathbf{:}\ \det A > 0 \right\}.$$

Prove that the convex hull $\mathrm{GL}^+(3)^{**}$ of $\mathrm{GL}^+(3)$ is equal to $\mathrm{GL}(3)$.

6.8 With the notation from the previous problem, define

$$U := \left\{ (A, \operatorname{cof} A, \det A) \in \mathrm{GL}(3) \times \mathrm{GL}(3) \times \mathbb{R} \ \mathbf{:}\ A \in \mathrm{GL}^+(3) \right\}$$

and show that

$$U^{**} = \mathrm{GL}(3) \times \mathrm{GL}(3) \times (0, \infty).$$

Hint: Show first:

(i) $(A, H, \delta) \in U$ and $G \in \mathrm{GL}^+(3)$ implies that $(GA, (\operatorname{cof} G)H, (\det G)\delta) \in U^{**}$;
(ii) $(\pm\mathrm{Id}, 0, \delta) \in U$ and $(0, \pm\mathrm{Id}, \delta) \in U^{**}$ for all $\delta > 0$;
(iii) $(A, 0, \delta), (0, H, \delta) \in U^{**}$ for all $A, H \in \mathrm{GL}(3)$ and all $\delta > 0$ (see the previous problem).

6.9 Let $\Phi\colon [0,\infty)^d \to \mathbb{R}$ be symmetric, jointly convex, and increasing (in every variable). Denote by $\sigma_1(A), \dots, \sigma_d(A) \geq 0$ the singular values of a matrix $A \in \mathbb{R}^{d\times d}$. Then, show that

$$g(A) := \Phi(\sigma_1(A), \dots, \sigma_d(A)), \qquad A \in \mathbb{R}^{d\times d},$$

is convex.

6.10 Let $h\colon \mathbb{R}^{3\times 3} \to \mathbb{R} \cup \{+\infty\}$ be of the form

$$h(A) := \begin{cases} \Theta\big(\sigma_1, \sigma_2, \sigma_3, \sigma_1\sigma_2, \sigma_2\sigma_3, \sigma_3\sigma_1, \sigma_1\sigma_2\sigma_3\big) & \text{if } \det A > 0, \\ +\infty & \text{if } \det A \leq 0, \end{cases}$$

where $\sigma_1, \sigma_2, \sigma_3$ are the three singular values of A and the function $\Theta: [0, \infty)^7 \times (0, \infty) \to \mathbb{R}$ is jointly convex, increasing in the first six variables, and

$$\Theta(x_1, x_2, x_3, y_1, y_2, y_3, z) = \Theta(x_{\gamma(1)}, x_{\gamma(2)}, x_{\gamma(3)}, y_{\eta(1)}, y_{\eta(2)}, y_{\eta(3)}, z)$$

for all permutations $\gamma, \eta: \{1, 2, 3\} \to \{1, 2, 3\}$ and all $x_1, x_2, x_3, y_1, y_2, y_3 \in [0, \infty)$, $z \in (0, \infty)$. Show that h is polyconvex. *Hint:* Use the previous problem.

Chapter 7
Relaxation

Consider the functional

$$\mathscr{F}[u] := \int_0^1 |u(x)|^2 + \bigl(|u'(x)|^2 - 1\bigr)^2 \,\mathrm{d}x, \qquad u \in \mathrm{W}_0^{1,4}(0,1).$$

The gradient part of the integrand, $a \mapsto (a^2 - 1)^2$, see Figure 7.1, has two distinct minima, which makes it a **double-well potential**. Approximate minimizers of $\mathscr{F}$ try to satisfy $u' \in \{-1, 1\}$ as closely as possible, while at the same time staying close to zero because of the first term. These contradicting requirements lead to minimizing sequences that develop faster and faster oscillations similar to the ones shown in Figure 5.1. It should be intuitively clear that no classical function can be a minimizer of $\mathscr{F}$.

In this situation, we have essentially two options, both of which we will consider in this chapter: First, if we only care about the infimal *value* of $\mathscr{F}$, we can compute the *relaxation* $\mathscr{F}_*$ of $\mathscr{F}$, which by definition is the largest lower semicontinuous functional below $\mathscr{F}$. It turns out that, under reasonable assumptions, $\mathscr{F}_*$ is also an integral functional and its integrand is the *quasiconvex envelope* of the integrand of $\mathscr{F}$. However, the minimizer of $\mathscr{F}_*$ may not say much about the minimizing sequence of our original $\mathscr{F}$ since all oscillations (and concentrations in some cases) have been "averaged out".

Second, we can focus on the minimizing sequences themselves and try to find a generalized limit object to a minimizing sequence that encapsulates "interesting" information. The natural candidates for such limit objects are (gradient) Young measures. In fact, applications to the relaxation of integral functionals were the original motivation for introducing them. Young measure theory allows one to replace a minimization problem over a Sobolev space by a generalized minimization problem over (gradient) Young measures. This generalized minimization problem *always* has a solution.

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8_7

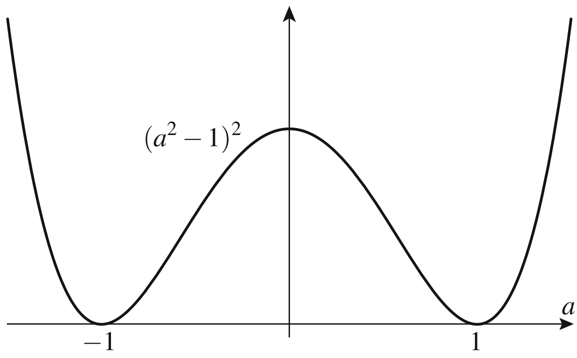


Fig. 7.1 A double-well potential

When formulating minimization problems over a class of gradient Young measures, the question naturally arises whether one can characterize this subset of Young measures. The *Kinderlehrer–Pedregal theorem* provides such a characterization by placing gradient Young measures in duality with quasiconvex functions. This duality also further emphasizes the central place of (gradient) Young measures in the modern calculus of variations.

In applications, the emerging oscillations in minimization sequences for non-quasiconvex integral functionals correspond to *microstructure*, which is very important, for instance, in material science. A finer investigation of these phenomena will be carried out in Chapters 8 and 9.

7.1 Quasiconvex Envelopes

We have seen in Chapter 5 that integral functionals with quasiconvex integrands are weakly lower semicontinuous in $\mathrm{W}^{1,p}$, where the exponent $p \in (1, \infty)$ is determined by growth properties of the integrand. If the integrand is *not* quasiconvex, then we would like to compute the functional's relaxation. Because of the close connection between weak lower semicontinuity and quasiconvexity, we can expect that the relaxation of an integral functional should also be an integral functional with a quasiconvex integrand that is related to the integrand of the original functional.

In this spirit, we define the **quasiconvex envelope** $Qh \colon \mathbb{R}^{m\times d} \to \mathbb{R} \cup \{-\infty\}$ of a locally bounded Borel-function $h \colon \mathbb{R}^{m\times d} \to \mathbb{R}$ as

$$Qh(A) := \inf \left\{ \fint_{B(0,1)} h(A + \nabla\psi(z))\,\mathrm{d}z \;:\; \psi \in \mathrm{W}^{1,\infty}_0(B(0,1);\mathbb{R}^m) \right\}, \qquad (7.1)$$

where $A \in \mathbb{R}^{m\times d}$. Clearly, $Qh \le h$. By a similar covering argument as the one employed in Lemma 5.2 one can see that in the above formula one may replace the unit ball $B(0,1)$ by any bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$. Furthermore, if h has

p-growth, we may replace the space $\mathrm{W}_0^{1,\infty}(B(0,1);\mathbb{R}^m)$ by $\mathrm{W}_0^{1,p}(B(0,1);\mathbb{R}^m)$ via a density argument. Finally, arguing as in the proof of Proposition 5.18, we may restrict the class of ψ in the above infimum to those satisfying $\|\psi\|_{\mathrm{L}^\infty} \le \varepsilon$ for any $\varepsilon > 0$.

Lemma 7.1. *For continuous $h\colon \mathbb{R}^{m\times d} \to [0,\infty)$ with p-growth, $p \in [1,\infty)$, the quasiconvex envelope Qh is quasiconvex.*

Proof. For any $\psi \in \mathrm{W}_0^{1,\infty}(B(0,1);\mathbb{R}^m)$ and any $F \in \mathbb{R}^{m\times d}$ we need to show

$$\fint_{B(0,1)} Qh(F + \nabla\psi(z))\,\mathrm{d}z \ge Qh(F). \tag{7.2}$$

We first note that Qh has p-growth since $0 \le Qh(A) \le h(A) \le M(1 + |A|^p)$. We can then use an approximation argument in conjunction with Theorem 2.13 to see that it suffices to show the inequality (7.2) for countably piecewise affine $\psi \in \mathrm{W}_0^{1,\infty}(B(0,1);\mathbb{R}^m)$. Here we use the $\mathrm{W}^{1,p}$-density of countably piecewise affine functions in $\mathrm{W}_0^{1,\infty}(B(0,1);\mathbb{R}^m)$ under given boundary values (see Theorem A.29). Suppose that $\psi(x) = v_k + A_k x$ ($v_k \in \mathbb{R}^m$, $A_k \in \mathbb{R}^{m\times d}$) for $x \in D_k$ from a disjoint collection of Lipschitz subdomains $D_k \subset B(0,1)$ ($k \in \mathbb{N}$) with $|B(0,1) \setminus \bigcup_k D_k| = 0$.

Fix $\varepsilon > 0$. By the definition of Qh, for every k we can find $\phi_k \in \mathrm{W}_0^{1,\infty}(D_k;\mathbb{R}^m)$ such that

$$Qh(F + A_k) \ge \fint_{D_k} h(F + A_k + \nabla\phi_k(z))\,\mathrm{d}z - \varepsilon.$$

Let $\phi \in \mathrm{W}_0^{1,\infty}(B(0,1);\mathbb{R}^m)$ be defined as

$$\phi(x) := v_k + A_k x + \phi_k(x) \qquad \text{if } x \in D_k \ \ (k \in \mathbb{N}).$$

Then,

$$\begin{aligned}
\int_{B(0,1)} Qh(F + \nabla\psi(z))\,\mathrm{d}z &= \sum_{k=1}^{\infty} |D_k|\, Qh(F + A_k) \\
&\ge \sum_{k=1}^{\infty} \biggl(\int_{D_k} h(F + A_k + \nabla\phi_k(z))\,\mathrm{d}z - \varepsilon |D_k| \biggr) \\
&= \int_{B(0,1)} h(F + \nabla\phi(z))\,\mathrm{d}z - \varepsilon\omega_d \\
&\ge \omega_d \bigl(Qh(F) - \varepsilon \bigr),
\end{aligned}$$

where the last step follows from the definition of Qh. Now let $\varepsilon \downarrow 0$ to conclude that (7.2) holds. □

Lemma 7.2. *For continuous* $h\colon \mathbb{R}^{m\times d} \to [0,\infty)$ *with* p*-growth it holds that*

$$Qh(A) = \sup\left\{g(A) : g \text{ quasiconvex and } g \le h\right\}, \qquad A \in \mathbb{R}^{m\times d}. \tag{7.3}$$

Proof. Denote the right-hand side of (7.3) by Q^*h. By the preceding lemma we have $Qh \le Q^*h$ since Qh itself is quasiconvex. On the other hand, for every quasiconvex g with $g \le h$ it must hold for all $A \in \mathbb{R}^{m\times d}$ that

$$\begin{aligned} g(A) &\le \inf\left\{ \fint_D g(A + \nabla\psi(z))\,\mathrm{d}z \;:\; \psi \in \mathrm{W}_0^{1,\infty}(D;\mathbb{R}^m) \right\} \\ &\le \inf\left\{ \fint_D h(A + \nabla\psi(z))\,\mathrm{d}z \;:\; \psi \in \mathrm{W}_0^{1,\infty}(D;\mathbb{R}^m) \right\} \\ &= Qh(A), \end{aligned}$$

by (7.1). Thus, also $Q^*h \le Qh$. This finishes the proof. □

On a side note, we can use the notion of the quasiconvex envelope to introduce a class of non-trivial quasiconvex functions.

Lemma 7.3. *Let* $F \in \mathbb{R}^{m\times d}$ *with* rank $F \ge 2$ *and let* $p \in (1,\infty)$. *Define*

$$h(A) := \operatorname{dist}(A, \{-F, F\})^p, \qquad A \in \mathbb{R}^{m\times d}.$$

Then, the quasiconvex envelope Qh *of* h *is not convex (at zero). Moreover,* Qh *has* p*-growth.*

Remark 7.4. The result remains true for $p = 1$, see Problem 11.3.

Proof. We will show that $Qh(0) > 0$. Then, if Qh was convex at zero, we would have

$$Qh(0) \le \frac{1}{2}\bigl(Qh(-F) + Qh(F)\bigr) \le \frac{1}{2}\bigl(h(-F) + h(F)\bigr) = 0,$$

a contradiction.

Assume to the contrary that $Qh(0) = 0$. Then, by (7.1) there would exist a sequence $(\psi_j) \subset \mathrm{W}_0^{1,\infty}(B(0,1);\mathbb{R}^m)$ with

$$\fint_{B(0,1)} h(\nabla\psi_j)\,\mathrm{d}z \to 0. \tag{7.4}$$

Set $L := \operatorname{span}\{F\}$ and let $\mathbf{P}\colon \mathbb{R}^{m\times d} \to L^\perp$ be the orthogonal projection onto the orthogonal complement of L. It is straightforward to see that $|\mathbf{P}(A)|^p \le h(A)$ for all $A \in \mathbb{R}^{m\times d}$. Therefore,

$$\mathbf{P}(\nabla\psi_j) \to 0 \quad \text{in } \mathrm{L}^p. \tag{7.5}$$

In the following we will employ the Fourier transform and Fourier multipliers as recalled in Appendix A.6. We will prove below that we may "invert" $\mathbf{P}$ in the sense that if

$$\mathbf{P}(\widehat{\nabla w}) = \widehat{R} \tag{7.6}$$

for some $w \in \mathrm{W}^{1,p}(\mathbb{R}^d; \mathbb{R}^m)$, $R \in \mathrm{L}^p(\mathbb{R}^d; L^\perp)$, then

$$\widehat{\nabla w}(\xi) = \mathbf{M}(\xi)\widehat{R}(\xi) = \mathbf{M}(\xi)\mathbf{P}(\widehat{\nabla w}(\xi)), \qquad \xi \in \mathbb{R}^d \setminus \{0\}, \tag{7.7}$$

for some family of linear operators $\mathbf{M}(\xi) \colon \mathbb{R}^{m\times d} \to \mathbb{R}^{m\times d}$ that depends smoothly and positively 0-homogeneously on ξ. Here, we identified $\mathbf{P}$ with its complexification (that is, $\mathbf{P}(A + \mathrm{i}B) = \mathbf{P}(A) + \mathrm{i}\mathbf{P}(B)$ for $A, B \in \mathbb{R}^{m\times d}$).

For $p = 2$, Plancherel's identity $\|g\|_{\mathrm{L}^2} = \|\widehat{g}\|_{\mathrm{L}^2}$ together with (7.5), (7.7) then implies

$$\begin{aligned}
\|\nabla\psi_j\|_{\mathrm{L}^2} &= \|\widehat{\nabla\psi_j}\|_{\mathrm{L}^2} \\
&= \|\mathbf{M}(\xi)\mathbf{P}(\widehat{\nabla\psi_j}(\xi))\|_{\mathrm{L}^2} \\
&\le \|\mathbf{M}\|_\infty \|\mathbf{P}(\widehat{\nabla\psi_j}(\xi))\|_{\mathrm{L}^2} \\
&= \|\mathbf{M}\|_\infty \|\mathbf{P}(\nabla\psi_j)\|_{\mathrm{L}^2} \\
&\to 0.
\end{aligned}$$

But then $h(\nabla\psi_j) \to |F|$ in L^1, contradicting (7.4). Thus, $Qh(0) > 0$.

For $p \in (1, \infty)$, we may apply the Mihlin Multiplier Theorem A.35 to get analogously that

$$\|\nabla\psi_j\|_{\mathrm{L}^p} \le C\|\mathbf{M}\|_{\mathrm{C}^{\lfloor d/2\rfloor+1}} \|\mathbf{P}(\nabla\psi_j)\|_{\mathrm{L}^p} \to 0,$$

which is again at odds with (7.4).

It remains to show (7.7). Notice that $\mathbf{P}(a\otimes\xi) \neq 0$ for any $a \in \mathbb{C}^m \setminus \{0\}$, $\xi \in \mathbb{R}^d \setminus \{0\}$ by the assumption that rank $F \ge 2$ (whereby L does not contain a rank-one line). Thus, for some constant $C > 0$ we have the *ellipticity* estimate

$$|a \otimes \xi| \le C|\mathbf{P}(a \otimes \xi)| \qquad \text{for all } a \in \mathbb{C}^m, \xi \in \mathbb{R}^d.$$

The (complexified) projection $\mathbf{P} \colon \mathbb{C}^{m\times d} \to \mathbb{C}^{m\times d}$ has kernel $L^{\mathbb{C}} := \operatorname{span}_{\mathbb{C}} L$ (the complex span of L), which in the following we also denote just by L. Hence, $\mathbf{P}$ descends to the quotient

$$[\mathbf{P}] \colon \mathbb{C}^{m\times d}/L \to \operatorname{ran}\mathbf{P},$$

and $[\mathbf{P}]$ is an invertible linear map. For $\xi \in \mathbb{R}^d \setminus \{0\}$ let

$$\bigl\{F, \mathrm{e}_1 \otimes \xi, \ldots, \mathrm{e}_d \otimes \xi, G_{d+1}(\xi), \ldots, G_{md-1}(\xi)\bigr\}$$

be a $\mathbb{C}$-basis of $\mathbb{C}^{m\times d}$ with the property that the matrices $G_{d+1}(\xi), \ldots, G_{md-1}(\xi)$ depend smoothly on ξ and are positively 1-homogeneous in ξ, that is, $G_{d+1}(\alpha\xi) = \alpha G_{d+1}(\xi)$ for all $\alpha \ge 0$. Furthermore, for $\xi \in \mathbb{R}^d \setminus \{0\}$ denote by $\mathbf{Q}(\xi) \colon \mathbb{C}^{m\times d} \to \mathbb{C}^{m\times d}$ the (non-orthogonal) projection with

$$\ker \mathbf{Q}(\xi) = L,$$
$$\operatorname{ran} \mathbf{Q}(\xi) = \operatorname{span}\big\{\mathrm{e}_1 \otimes \xi, \dots, \mathrm{e}_d \otimes \xi, G_{d+1}(\xi), \dots, G_{md-k}(\xi)\big\}.$$

If we interpret $\mathrm{e}_1 \otimes \xi, \dots, \mathrm{e}_d \otimes \xi, G_{d+1}(\xi), \dots, G_{md-1}(\xi)$ as vectors in $\mathbb{R}^{md}$ and collect them into the columns of the matrix $X(\xi) \in \mathbb{R}^{md\times(md-1)}$, and if we further let $Y \in \mathbb{R}^{md\times(md-1)}$ be a matrix whose columns comprise an orthonormal basis of $L^\perp$, then, up to a change in sign for one of the G_l's, there exists a constant $c > 0$ such that

$$\det(Y^T X(\xi)) \geq c > 0, \qquad \text{for all } \xi \in \mathbb{S}^{d-1}.$$

Indeed, if $\det(Y^T X(\xi))$ was not uniformly bounded away from zero for all $\xi \in \mathbb{S}^{d-1}$, then by compactness there would exist a $\xi_0 \in \mathbb{S}^{d-1}$ with $\det(Y^T X(\xi_0)) = 0$, a contradiction. We can then write $\mathbf{Q}(\xi)$ explicitly as

$$\mathbf{Q}(\xi) = X(\xi)(Y^T X(\xi))^{-1} Y^T.$$

This implies that $\mathbf{Q}(\xi)$ depends positively 0-homogeneously and smoothly on $\xi \in \mathbb{R}^d \setminus \{0\}$. Also $\mathbf{Q}(\xi)$ descends to the quotient

$$[\mathbf{Q}(\xi)]\colon \mathbb{C}^{m\times d}/L \to \operatorname{ran} \mathbf{Q}(\xi),$$

which is now invertible. It is not difficult to see that $\xi \mapsto [\mathbf{Q}(\xi)]$ is still positively 0-homogeneous and smooth in $\xi \neq 0$ (by utilizing the basis given above).

Since $\widehat{w}(\xi) \otimes \xi \in \operatorname{ran} \mathbf{Q}(\xi)$, we have

$$[\mathbf{Q}(\xi)]^{-1}(\widehat{w}(\xi) \otimes \xi) = [\widehat{w}(\xi) \otimes \xi],$$

where $[\widehat{w}(\xi) \otimes \xi]$ designates the equivalence class of $\widehat{w}(\xi) \otimes \xi$ in $\mathbb{C}^{m\times d}/L$. This fact in conjunction with $\widehat{\nabla w}(\xi) = (2\pi \mathrm{i})\, \widehat{w}(\xi) \otimes \xi$ allows us to rewrite (7.6) in the form

$$(2\pi \mathrm{i})\, [\mathbf{P}][\mathbf{Q}(\xi)]^{-1}(\widehat{w}(\xi) \otimes \xi) = \widehat{R}(\xi),$$

or equivalently as

$$\widehat{\nabla w}(\xi) = (2\pi \mathrm{i})\, \widehat{w}(\xi) \otimes \xi = [\mathbf{Q}(\xi)][\mathbf{P}]^{-1} \widehat{R}(\xi).$$

The multiplier $\mathbf{M}(\xi)\colon \mathbb{R}^{m\times d} \to \mathbb{R}^{m\times d}$ for $\xi \in \mathbb{R}^d \setminus \{0\}$ is thus given by

$$\mathbf{M}(\xi) := [\mathbf{Q}(\xi)][\mathbf{P}]^{-1},$$

which is smooth and positively 0-homogeneous in ξ. Consequently, we have shown the multiplier equation (7.7).

For the last assertion in the statement of the lemma, it suffices to notice that h has p-growth and is non-negative. Hence, $Qh \leq h$ also has p-growth and is non-negative. □

7.2 Relaxation of Integral Functionals

We first consider the abstract principles of relaxation before moving on to more concrete integral functionals. Consider a functional $\mathscr{F}: X \to \mathbb{R}$, where X is a reflexive Banach space. Its **(weak) relaxation** $\mathscr{F}_*: X \to \mathbb{R} \cup \{-\infty\}$ is defined to be

$$\mathscr{F}_*[u] := \sup \left\{ \mathscr{H}[u] \;:\; \mathscr{H} \leq \mathscr{F} \text{ and } \mathscr{H} \text{ is weakly lower semicontinuous} \right\},$$

where $u \in X$; see Problem 7.5 for an alternative definition (also cf. Proposition 2.28 in the convex case).

Theorem 7.5. *Let X be a reflexive Banach space and let $\mathscr{F}: X \to \mathbb{R}$ be a functional. Assume furthermore:*

(WH1) ***Weak coercivity:*** *For all $\Lambda > 0$ the sublevel set*

$$\left\{ u \in X \;:\; \mathscr{F}[u] \leq \Lambda \right\} \qquad \text{is sequentially weakly precompact.}$$

Then, the relaxation $\mathscr{F}_*$ of $\mathscr{F}$ is weakly lower semicontinuous and

$$\min_X \mathscr{F}_* = \inf_X \mathscr{F}.$$

Proof. The functional $\mathscr{F}_*$ is weakly lower semicontinuous as the supremum of weakly lower semicontinuous functionals. Indeed, if $u_j \rightharpoonup u$ in X, then for all weakly lower semicontinuous $\mathscr{H}: X \to \mathbb{R}$ with $\mathscr{H} \leq \mathscr{F}$,

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

Taking the supremum over all such $\mathscr{H}$, we see that $\mathscr{F}_*[u] \leq \liminf_{j\to\infty} \mathscr{F}_*[u_j]$.

By the Direct Method, see Theorem 2.3, $\mathscr{F}_*$ attains its minimum. Since

$$\inf_X \mathscr{F} \leq \mathscr{F}_* \leq \mathscr{F},$$

the minimum of $\mathscr{F}_*$ must agree with the infimum of $\mathscr{F}$ over X. □

As usual, we are most interested in the concrete case of an integral functional

$$\mathscr{F}[u] := \int_\Omega f(x, \nabla u(x)) \,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m),$$

where $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain, $p \in (1, \infty)$, and $f : \Omega \times \mathbb{R}^{m\times d} \to \mathbb{R}$ is a Carathéodory integrand satisfying the p-growth and coercivity assumption

$$\mu|A|^p \le f(x, A) \le M(1 + |A|^p), \qquad (x, A) \in \Omega \times \mathbb{R}^{m\times d}, \tag{7.8}$$

for some $\mu, M > 0$. The main result of this section is the following relaxation theorem:

Theorem 7.6. *Let $\mathscr{F}$ be as above and assume furthermore that there exists a modulus of continuity ω (i.e., $\omega\colon [0, \infty) \to [0, \infty)$ continuous, increasing, and $\omega(0) = 0$) such that*

$$|f(x, A) - f(y, A)| \le \omega(|x - y|)(1 + |A|^p), \qquad x, y \in \Omega,\ A \in \mathbb{R}^{m\times d}. \tag{7.9}$$

Then, the relaxation $\mathscr{F}_*$ of $\mathscr{F}$ is

$$\mathscr{F}_*[u] = \int_\Omega Qf(x, \nabla u(x))\,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m),$$

where $Qf(x, \cdot)$ denotes the quasiconvex envelope of $f(x, \cdot)$ for $x \in \Omega$. The same conclusion holds if we prescribe fixed boundary values.

Proof. We define

$$\mathscr{G}[u] := \int_\Omega Qf(x, \nabla u(x))\,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m).$$

As $Qf(x, \cdot)$ is quasiconvex for all $x \in \Omega$ by Lemma 7.1, it is continuous by Lemma 5.6. We will moreover see below that $Qf(\cdot, A)$ is continuous for all fixed $A \in \mathbb{R}^{m\times d}$, hence in particular Qf is Carathéodory and $\mathscr{G}$ is well-defined.

We will show in the following that (a) $\mathscr{G} \le \mathscr{F}_*$ and (b) $\mathscr{G} \ge \mathscr{F}_*$.

To see (a), it suffices to observe that $\mathscr{G}$ is weakly lower semicontinuous by Morrey's Theorem 5.16 and that $\mathscr{G} \le \mathscr{F}$. Thus, from the definition of $\mathscr{F}_*$ we immediately get $\mathscr{G} \le \mathscr{F}_*$.

We will prove (b) in several steps.

Step 1. Let $\varepsilon > 0$ and fix $A \in \mathbb{R}^{m\times d}$. For $x \in \Omega$ let $\psi_x \in \mathrm{W}^{1,p}_0(B(0, 1); \mathbb{R}^m)$ be such that

$$\fint_{B(0,1)} f(x, A + \nabla\psi_x(z))\,\mathrm{d}z \le Qf(x, A) + \varepsilon.$$

Then we use (7.8) to observe

$$\mu \fint_{B(0,1)} |A + \nabla\psi_x(z)|^p\,\mathrm{d}z \le Qf(x, A) + \varepsilon \le M(1 + |A|^p) + \varepsilon.$$

Let now $x, y \in \Omega$ and estimate using (7.9) and the definition of $Qf(x, \cdot)$,

$$\begin{aligned} Qf(y,A) - Qf(x,A) &\le \fint_{B(0,1)} \bigl| f(y, A + \nabla\psi_x(z)) - f(x, A + \nabla\psi_x(z)) \bigr| \,\mathrm{d}z + \varepsilon \\ &\le \omega(|x-y|) \fint_{B(0,1)} 1 + |A + \nabla\psi_x(z)|^p \,\mathrm{d}z + \varepsilon \\ &\le C\omega(|x-y|)(1+|A|^p) + \varepsilon, \end{aligned}$$

where $C = C(\mu, M)$ is a constant. Letting $\varepsilon \downarrow 0$ and also exchanging the roles of x and y, we see that

$$\bigl| Qf(x,A) - Qf(y,A) \bigr| \le C\omega(|x-y|)(1+|A|^p) \tag{7.10}$$

for all $x, y \in \Omega$ and $A \in \mathbb{R}^{m\times d}$. In particular, $Qf(\,\bullet\,, A)$ is continuous.

Step 2. Next, we will show that it suffices to prove the claim (b) for countably piecewise affine u. If $u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$, then there exists a sequence $(v_j) \subset \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ of countably piecewise affine functions such that $v_j \to u$ in $\mathrm{W}^{1,p}$ and we may also require that $v|_{\partial\Omega} = u|_{\partial\Omega}$ (see Theorem A.29). Since Qf is Carathéodory and has p-growth (as $0 \le Qf \le f$), Theorem 2.13 shows that $\mathscr{G}[v_j] \to \mathscr{G}[u]$. Thus, if (b) holds for all countably piecewise affine u, we get using the lower semicontinuity of $\mathscr{F}_*$ that

$$\mathscr{G}[u] = \lim_{j\to\infty} \mathscr{G}[v_j] \ge \liminf_{j\to\infty} \mathscr{F}_*[v_j] \ge \mathscr{F}_*[u].$$

This proves (b) on all of $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$.

Step 3. Fix $\varepsilon > 0$ and let $u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ be countably piecewise affine, say $u(x) = v_k + A_k x$ (where $v_k \in \mathbb{R}^m$, $A_k \in \mathbb{R}^{m\times d}$) on the set D_k from a disjoint collection of open sets $D_k \subset \Omega$ ($k \in \mathbb{N}$) such that $|\Omega \setminus \bigcup_k D_k| = 0$. For any $k \in \mathbb{N}$ we may cover D_k up to a negligible set with countably many disjoint balls $B_l^{(k)} := B(x_l^{(k)}, r_l^{(k)}) \subset D_k$, where $x_l^{(k)} \in D_k$, $0 < r_l^{(k)} < \varepsilon$ ($l \in \mathbb{N}$) such that

$$\left| \fint_{B_l^{(k)}} Qf(x, A_k) \,\mathrm{d}x - Qf(x_l^{(k)}, A_k) \right| \le C\omega(\varepsilon)(1+|A_k|^p). \tag{7.11}$$

This covering exists by the Vitali Covering Theorem A.15 in conjunction with (7.10).

From the definition of Qf and the remarks following it, in each ball $B_l^{(k)}$, we can find a map $\psi_l^{(k)} \in \mathrm{W}_0^{1,\infty}(B_l^{(k)};\mathbb{R}^m)$ with $\|\psi_l^{(k)}\|_{\mathrm{L}^\infty} \le \varepsilon$ and

$$\left| Qf(x_l^{(k)}, A_k) - \fint_{B_l^{(k)}} f\bigl(x_l^{(k)}, A_k + \nabla\psi_l^{(k)}(z)\bigr) \,\mathrm{d}z \right| \le \varepsilon. \tag{7.12}$$

Set

$$v_\varepsilon(x) := u(x) + \psi_l^{(k)}(x) \qquad \text{if } x \in B_l^{(k)} \ (k, l \in \mathbb{N}).$$

In a similar way to Step 1 we can show that

$$\mu \fint_{B_l^{(k)}} |A_k + \nabla \psi_l^{(k)}(z)|^p \,\mathrm{d}z \le M(1 + |A_k|^p) + \varepsilon.$$

Thus,

$$\begin{aligned}\int_\Omega |\nabla v_\varepsilon|^p \,\mathrm{d}x &= \sum_{k,l} \int_{B_l^{(k)}} |A_k + \nabla \psi_l^{(k)}(x)|^p \,\mathrm{d}x \\ &\le \frac{M}{\mu} \int_\Omega 1 + |\nabla u(x)|^p \,\mathrm{d}x + \varepsilon \frac{|\Omega|}{\mu} \\ &< \infty.\end{aligned}$$

So, by the Poincaré inequality, $(v_\varepsilon)_{\varepsilon>0} \subset \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ is uniformly norm-bounded.

From the continuity assumption (7.9) we infer that

$$\left| \fint_{B_l^{(k)}} f(x_l^{(k)}, \nabla v_\varepsilon(x)) \,\mathrm{d}x - \fint_{B_l^{(k)}} f(x, \nabla v_\varepsilon(x)) \,\mathrm{d}x \right| \le \omega(\varepsilon) \fint_{B_l^{(k)}} 1 + |\nabla v_\varepsilon(x)|^p \,\mathrm{d}x.$$

We can combine this with (7.11), (7.12) to get

$$\begin{aligned}&\left| \fint_{B_l^{(k)}} Qf(x, \nabla u(x)) \,\mathrm{d}x - \fint_{B_l^{(k)}} f(x, \nabla v_\varepsilon(x)) \,\mathrm{d}x \right| \\ &\qquad \le C\omega(\varepsilon) \fint_{B_l^{(k)}} 2 + |A_k|^p + |\nabla v_\varepsilon(x)|^p \,\mathrm{d}x + \varepsilon.\end{aligned}$$

Multiplying both sides by $|B_l^{(k)}|$ and summing over all k, l, we arrive at

$$\big|\mathscr{G}[u] - \mathscr{F}[v_\varepsilon]\big| \le C\omega(\varepsilon) \int_\Omega 2 + |\nabla u(x)|^p + |\nabla v_\varepsilon(x)|^p \,\mathrm{d}x + \varepsilon|\Omega|,$$

and this vanishes as $\varepsilon \downarrow 0$. For $u_j := v_{1/j}$ we have $u_j \rightharpoonup u$ in $\mathrm{W}^{1,p}$ since (u_j) is weakly precompact in $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ and $\|u_j - u\|_{\mathrm{L}^\infty} \le 1/j \to 0$. Hence, by the weak lower semicontinuity of $\mathscr{F}_*$,

$$\mathscr{G}[u] = \lim_{j\to\infty} \mathscr{F}[u_j] \ge \liminf_{j\to\infty} \mathscr{F}_*[u_j] \ge \mathscr{F}_*[u],$$

which is (b) for countably piecewise affine u. □

7.3 Generalized Convexity Notions and Envelopes

The four major convexity conditions that play a role in the modern calculus of variations satisfy the following implications:

$$\text{convexity} \implies \text{polyconvexity} \implies \text{quasiconvexity} \implies \text{rank-one convexity},$$

where the second and third implications were established in Propositions 6.1 and 5.3, respectively.

Define for continuous $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ the **polyconvex envelope** $Ph\colon \mathbb{R}^{m\times d} \to \mathbb{R}\cup\{-\infty\}$ and the **rank-one convex envelope** $Rh\colon \mathbb{R}^{m\times d} \to \mathbb{R}\cup\{-\infty\}$ of h as

$$\begin{aligned} Ph(A) &:= \sup\big\{\, g(A) \;:\; g \text{ polyconvex and } g \le h \,\big\}, \\ Rh(A) &:= \sup\big\{\, g(A) \;:\; g \text{ rank-one convex and } g \le h \,\big\}, \end{aligned} \qquad A \in \mathbb{R}^{m\times d}.$$

In this context also recall that if $h \ge 0$ (or bounded below) we showed an analogous formula for the quasiconvex envelope in Lemma 7.2. As a consequence of the above implications between the convexity notions, we have that

$$h^{**} \le Ph \le Qh \le Rh \le h,$$

where we recall that h^{**} denotes the convex envelope of h. In general, Qh is difficult to compute, so Ph and Rh can give useful lower and upper bounds on Qh.

While in the scalar case ($d = 1$ or $m = 1$) all four generalized notions of convexity are equivalent, this no longer holds in higher dimensions. Clearly, the determinant function is polyconvex but not convex. We next exhibit an example of a quasiconvex, but not polyconvex function (also see Problem 6.5):

Example 7.7. Let $F \in \mathbb{R}^{3\times 3}$ with rank $F \ge 2$ and let $p \in (1, 2)$. Define

$$h(A) := \operatorname{dist}(A, \{-F, F\})^p, \qquad A \in \mathbb{R}^{3\times 3}.$$

Then, the quasiconvex envelope $Qh\colon \mathbb{R}^{3\times 3} \to [0,\infty)$ is quasiconvex by Lemma 7.1 and not convex (at zero) by Lemma 7.3. Since Qh has p-growth and we chose $p < 2$, it can be shown without too much effort that Qh cannot be polyconvex (using the fact that non-constant convex functions have at least linear growth in at least one direction); see, for instance, Corollary 5.9 (i) in [76].

Example 7.8 (*Alibert–Dacorogna–Marcellini 1988 [7, 78]*). From Example 5.4 we recall the Alibert–Dacorogna–Marcellini function

$$h_\gamma(A) := |A|^2\big(|A|^2 - 2\gamma \det A\big), \qquad A \in \mathbb{R}^{2\times 2}.$$

It can be shown that h_γ is polyconvex if and only if $|\gamma| \le 1$. Since it is quasiconvex if and only if $|\gamma| \le \gamma_{\mathrm{QC}}$, where $\gamma_{\mathrm{QC}} > 1$, this provides a further example of a quasiconvex function that is not polyconvex. See again Section 5.3.8 in [76] for the details.

When he introduced quasiconvexity, Morrey conjectured that rank-one convexity was a strictly weaker notion. This was one of the major open problems in the field for a long time:

Conjecture 7.9 (*Morrey 1952 [195]*). Rank-one convexity does not imply quasiconvexity.

A verification of this conjecture proved elusive until Švérak's 1992 counterexample, which proved the conjecture, at least for $d \geq 2$, $m \geq 3$, see below. The case $d = m = 2$ is still a major unsolved problem, also because of its connection to other branches of mathematics, see, for instance, [18]. A partial result for (2×2)-diagonal matrices is in [201].

Example 7.10 (*Švérak 1992 [252]*). We will show the non-equivalence of quasiconvexity and rank-one convexity for $d = 2$, $m = 3$ only. Higher dimensions can be treated using an embedding of $\mathbb{R}^{3\times 2}$ into $\mathbb{R}^{m\times d}$. We will construct a function $h \colon \mathbb{R}^{3\times 2} \to \mathbb{R}$ that is rank-one convex but not quasiconvex.

Define a linear subspace L of $\mathbb{R}^{3\times 2}$ as

$$L := \left\{ \begin{pmatrix} x & 0 \\ 0 & y \\ z & z \end{pmatrix} \ \colon x, y, z \in \mathbb{R} \right\}$$

and denote by $\mathbf{P} \colon \mathbb{R}^{3\times 2} \to L$ the linear projection onto L given as

$$\mathbf{P}(A) := \begin{pmatrix} a & 0 \\ 0 & d \\ (e+f)/2 & (e+f)/2 \end{pmatrix} \qquad \text{for} \qquad A = \begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} \in \mathbb{R}^{3\times 2}.$$

Also let $g \colon L \to \mathbb{R}$ be defined by

$$g \begin{pmatrix} x & 0 \\ 0 & y \\ z & z \end{pmatrix} := -xyz.$$

For $\alpha, \beta > 0$ we set $h_{\alpha,\beta} \colon \mathbb{R}^{3\times 2} \to \mathbb{R}$ to be

$$h_{\alpha,\beta}(A) := g(\mathbf{P}(A)) + \alpha\big(|A|^2 + |A|^4\big) + \beta|A - \mathbf{P}(A)|^2.$$

Below we will prove the following two properties of $h_{\alpha,\beta}$:

(i) For every $\alpha > 0$ sufficiently small and all $\beta > 0$ the function $h_{\alpha,\beta}$ is not quasiconvex.
(ii) For every $\alpha > 0$ there exists a $\beta = \beta(\alpha) > 0$ such that $h_{\alpha,\beta}$ is rank-one convex.

This implies the claim for suitable α, β.

Ad (i): For the periodic map $\phi \in \mathrm{W}^{1,\infty}_{\mathrm{per}}((0,1)^2; \mathbb{R}^3)$ given as

$$\phi(x_1, x_2) := \frac{1}{2\pi} \begin{pmatrix} \sin(2\pi x_1) \\ \sin(2\pi x_2) \\ \sin(2\pi(x_1 + x_2)) \end{pmatrix}, \qquad x = (x_1, x_2) \in (0,1)^2,$$

we have $\nabla\phi \in L$ and hence $\mathbf{P}(\nabla\phi) = \nabla\phi$. Thus, we may compute in an elementary way

$$\int_{(0,1)^2} g(\nabla\phi)\,\mathrm{d}x = -\int_0^1\int_0^1 (\cos 2\pi x_1)^2(\cos 2\pi x_2)^2\,\mathrm{d}x_1\,\mathrm{d}x_2 = -\frac{1}{4} < 0.$$

Then, for $\alpha > 0$ sufficiently small and all $\beta > 0$,

$$\int_{(0,1)^2} h_{\alpha,\beta}(\nabla\phi)\,\mathrm{d}x < 0 = h_{\alpha,\beta}(0). \tag{7.13}$$

It turns out that in the definition of quasiconvexity we may alternatively test with functions that have periodic boundary values, see Problem 5.2. Hence, (i) follows.

Ad (ii): By Problem 7.6, the rank-one convexity of $h_{\alpha,\beta}$ is equivalent to the **Legendre–Hadamard condition**

$$\mathrm{D}^2 h_{\alpha,\beta}(A)[B,B] := \frac{\mathrm{d}^2}{\mathrm{d}t^2} h_{\alpha,\beta}(A+tB)\bigg|_{t=0} \geq 0 \tag{7.14}$$

for all $A, B \in \mathbb{R}^{3\times 2}$ with rank $B \leq 1$.

The function g is a homogeneous polynomial of degree 3, whereby we can find $c > 0$ such that for all $A, B \in \mathbb{R}^{3\times 2}$ with rank $B \leq 1$,

$$G(A,B) := \mathrm{D}^2(g\circ\mathbf{P})(A)[B,B] := \frac{\mathrm{d}^2}{\mathrm{d}t^2} g(\mathbf{P}(A+tB))\bigg|_{t=0} \geq -c|A||B|^2.$$

A computation then shows that

$$\begin{aligned}\mathrm{D}^2 h_{\alpha,\beta}(A)[B,B] &= G(A,B) + 2\alpha|B|^2 + 4\alpha|A|^2|B|^2 + 8\alpha(A:B)^2 \\ &\quad + 2\beta|B-\mathbf{P}(B)|^2 \\ &\geq (-c+4\alpha|A|)|A||B|^2\end{aligned}$$

for some $c > 0$. Thus, for $|A| \geq c/(4\alpha)$, the Legendre–Hadamard condition (and hence the rank-one convexity) holds.

We still need to prove the Legendre–Hadamard condition for $|A| < c/(4\alpha)$. Since $\mathrm{D}^2 h_{\alpha,\beta}(A)[B,B]$ is homogeneous of degree 2 in B, we only need to consider A, B from the compact set

$$K := \left\{ (A,B) \in \mathbb{R}^{3\times 2}\times\mathbb{R}^{3\times 2} : |A| \leq \frac{c}{4\alpha},\ |B| = 1,\ \operatorname{rank} B = 1 \right\}.$$

By the estimate

$$\mathrm{D}^2 h_{\alpha,\beta}(A)[B,B] \geq G(A,B) + 2\alpha|B|^2 + 2\beta|B-\mathbf{P}(B)|^2 =: \kappa(A,B,\beta)$$

it suffices to show that there exists a $\beta = \beta(\alpha) > 0$ such that $\kappa(A, B, \beta) \geq 0$ for all $(A, B) \in K$. Assume that this is not the case. Then there is a sequence $\beta_j \to \infty$ and $(A_j, B_j) \in K$ with

$$0 > \kappa(A_j, B_j, \beta_j) = G(A_j, B_j) + 2\alpha + 2\beta_j |B_j - \mathbf{P}(B_j)|^2.$$

As K is compact, we may assume $(A_j, B_j) \to (A, B) \in K$, for which

$$G(A, B) + 2\alpha \leq 0, \qquad \mathbf{P}(B) = B, \qquad \text{and} \qquad \operatorname{rank} B = 1.$$

However, since rank $B = 1$, we have $g(\mathbf{P}(A+tB)) = 0$ for all $t \in \mathbb{R}$. This implies in particular $G(A, B) = \mathrm{D}^2(g \circ \mathbf{P})(A)[B, B] = 0$, yielding $2\alpha \leq 0$, which contradicts our assumption $\alpha > 0$.

Thus, in conclusion, we have shown that for a suitable choice of $\alpha, \beta > 0$ it indeed holds that $h_{\alpha,\beta}$ is rank-one convex.

Based on Švérak's example, it has been shown that quasiconvexity is not a local condition, meaning that there is no pointwise condition involving only the function and a finite number of its derivatives that is both necessary and sufficient for the function to be quasiconvex (verifying a conjecture by Morrey [195]). More precisely, let $\mathscr{Q}\colon \mathrm{C}^\infty(\mathbb{R}^{m\times d}) \to \mathbf{X}^{m\times d}$ be a nonlinear operator, where we denote by $\mathbf{X}^{m\times d}$ the space of functions from $\mathbb{R}^{m\times d}$ to $[-\infty, +\infty]$. Call $\mathscr{Q}$ **local** if $h = g$ in a neighborhood of $A \in \mathbb{R}^{m\times d}$ implies that also $\mathscr{Q}[h] = \mathscr{Q}[g]$ in a neighborhood of A.

Theorem 7.11 (**Kristensen 1999 [165]**). *Let $d \geq 2$ and $m \geq 3$. Then, there exists no local (nonlinear) operator $\mathscr{Q}\colon \mathrm{C}^\infty(\mathbb{R}^{m\times d}) \to \mathbf{X}^{m\times d}$ such that*

$$\mathscr{Q}[h] = 0 \quad \Longleftrightarrow \quad h \text{ is quasiconvex}$$

for all $h \in \mathrm{C}^\infty(\mathbb{R}^{m\times d})$.

This is in contrast to rank-one convexity, which is characterized by the local operator

$$\mathscr{R}[h](A) := \inf \left\{ \mathrm{D}^2 h(A)[a \otimes b, a \otimes b] \;:\; a \in \mathbb{R}^m,\; b \in \mathbb{R}^d \right\},$$

where $h \in \mathrm{C}^\infty(\mathbb{R}^{m\times d})$ and $A \in \mathbb{R}^{m\times d}$.

7.4 Young Measure Relaxation

As discussed at the beginning of this chapter, the relaxation strategy of Section 7.2 has one serious drawback: While it allows us to find the infimal *value*, the relaxed functional potentially says only very little about the "shape" of minimizing sequences (e.g. their oscillations). Often, however, this information is decisive. For example, in material science, oscillations in minimizing sequences correspond to crystalline microstructure, which greatly influences the material properties, see Section 1.8 and also Chapter 9. Therefore, in this section we implement the second strategy outlined at the beginning of the chapter, that is, we extend the minimization problem to a larger space and look for solutions there.

Let us first, in an abstract fashion, collect a few properties that our extension should satisfy. Assume we are given a metric space X with a convergence "$\to$" and a functional $\mathscr{F}\colon X \to \mathbb{R} \cup \{+\infty\}$. Then, we extend X to a complete metric space $\overline{X}$ with convergence "$\rightsquigarrow$". For this, we assume that there exists a (usually not continuous) map

$$\iota\colon X \to \overline{X}.$$

We then seek to extend $\mathscr{F}$ to a functional $\overline{\mathscr{F}}\colon \overline{X} \to \mathbb{R} \cup \{+\infty\}$, which we call the **extension–relaxation** of $\mathscr{F}$, such that the following conditions are satisfied:

(i) **Extension property:** $\overline{\mathscr{F}} \circ \iota = \mathscr{F}$.
(ii) **Lower bound:** If $(u_j) \subset X$ is precompact, then, up to selecting a subsequence, there exists a $\nu \in \overline{X}$ such that $\iota(u_j) \rightsquigarrow \nu$ in $\overline{X}$ and

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

(iii) **Recovery sequence:** For all $\nu \in \overline{X}$ there exists a **recovery sequence** $(u_j) \subset X$ with $\iota(u_j) \rightsquigarrow \nu$ in $\overline{X}$ and such that

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

Intuitively, these conditions entail that we can solve our minimization problem for $\mathscr{F}$ by passing to the extended space $\overline{X}$. In particular, if (u_j) is a minimizing and precompact sequence in X, then (ii) tells us that any limit $\nu \in \overline{X}$ as in (ii) should be considered a *generalized minimizer*. On the other hand, (iii) ensures that the minimization problems for $\mathscr{F}$ and for $\overline{\mathscr{F}}$ are sufficiently related. In particular, it is easy to see that the infima of $\mathscr{F}$ and $\overline{\mathscr{F}}$ agree and, if we additionally assume some coercivity, then $\overline{\mathscr{F}}$ attains its minimum, so

$$\min_{\overline{X}} \overline{\mathscr{F}} = \inf_X \mathscr{F}.$$

Let us specialize this abstract approach to our prototypical integral functional

$$\mathscr{F}[u] := \int_\Omega f(x, \nabla u(x)) \,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m),$$

with $\Omega \subset \mathbb{R}^d$ a bounded Lipschitz domain, $p \in (1, \infty)$, and $f\colon \Omega \times \mathbb{R}^{m\times d} \to [0, \infty)$ a Carathéodory integrand satisfying the p-growth and coercivity assumption

$$\mu|A|^p \le f(x, A) \le M(1 + |A|^p), \qquad (x, A) \in \Omega \times \mathbb{R}^{m \times d}, \tag{7.15}$$

for some $\mu, M > 0$. Then, X is a norm-bounded subset of $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ with the weak topology and we will use a space of gradient Young measures for $\overline{X}$. For $\nu = (\nu_x)_x \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m \times d})$, we define the **extension–relaxation**

$$\overline{\mathscr{F}} \colon \mathbf{GY}^p(\Omega; \mathbb{R}^{m \times d}) \to \mathbb{R}$$

as

$$\overline{\mathscr{F}}[\nu] := \langle\!\langle f, \nu \rangle\!\rangle = \int_\Omega \int f(x, A) \,\mathrm{d}\nu_x(A) \,\mathrm{d}x.$$

Then, our relaxation result takes the following form.

Theorem 7.12. *Let $\mathscr{F}, \overline{\mathscr{F}}$ be as above.*

(i) ***Extension property:*** *For every $u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ the elementary Young measure $\delta[\nabla u] = (\delta_{\nabla u(x)})_x \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m \times d})$ satisfies*

$$\mathscr{F}[u] = \overline{\mathscr{F}}[\delta[\nabla u]].$$

(ii) ***Lower bound:*** *If $(u_j) \subset \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ is weakly precompact, then, up to selecting a subsequence, there exists a Young measure $\nu \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m \times d})$ such that $\nabla u_j \overset{\mathbf{Y}}{\to} \nu$ with $[\nu] = \nabla u$ and*

$$\overline{\mathscr{F}}[\nu] \le \liminf_{j \to \infty} \mathscr{F}[u_j]. \tag{7.16}$$

(iii) ***Recovery sequence:*** *For all $\nu \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m \times d})$ there exists a recovery sequence $(u_j) \subset \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ with $\nabla u_j \overset{\mathbf{Y}}{\to} \nu$ and*

$$\lim_{j \to \infty} \mathscr{F}[u_j] = \overline{\mathscr{F}}[\nu]. \tag{7.17}$$

(iv) ***Equality of minima:*** *$\overline{\mathscr{F}}$ attains its minimum and*

$$\min_{\mathbf{GY}^p(\Omega; \mathbb{R}^{m \times d})} \overline{\mathscr{F}} = \inf_{\mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)} \mathscr{F}.$$

Furthermore, all these statements remain true if we prescribe boundary values. For a Young measure this refers to the underlying deformation (which is only determined up to a translation, of course).

Proof. Ad (i). This follows directly from the definition of $\overline{\mathscr{F}}$.

Ad (ii). The existence of ν is a consequence of the Fundamental Theorem of Young measure theory, Theorem 4.1. The lower bound (7.16) follows from Proposition 4.6.

Ad (iii). Via Lemma 4.13 we may construct a sequence $(u_j) \subset \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ with

(a) $u_j|_{\partial\Omega} = u|_{\partial\Omega}$, where $u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ is an underlying deformation of ν;
(b) the family $\{\nabla u_j\}_j$ is L^p-equiintegrable;
(c) $\nabla u_j \overset{\mathbf{Y}}{\to} \nu$.

For this special generating sequence we can now use the statement about representation of limits of integral functionals from the Fundamental Theorem 4.1 (here we need the p-equiintegrability) to get

$$\mathscr{F}[u_j] \to \langle\!\langle f, \nu \rangle\!\rangle = \overline{\mathscr{F}}[\nu],$$

which is nothing else than (7.17).

Ad (iv). This is not hard to see using (b), (c) and the coercivity assumption, that is, the lower bound in (7.15). □

Example 7.13. In our sailing example from Section 1.6, we were tasked with solving the optimal beating problem

$$\begin{cases} \text{Minimize} \quad \mathscr{F}[r] := \displaystyle\int_0^T v_{\max} \cdot \frac{\cos(4\arctan r'(t)) - 1}{2} + v_{\text{flow}}\left(\frac{r(t)^2}{R^2} - 1\right) \,\mathrm{d}t \\ \text{subject to} \quad r(0) = r(T) = 0, \;\; |r(t)| \le R. \end{cases}$$

Here, because of the additional constraints we can work in any L^p-space, even L^∞. Clearly, the integrand

$$f(r,a) := v_{\max} \cdot \frac{\cos(4\arctan a) - 1}{2} + v_{\text{flow}}\left(\frac{r^2}{R^2} - 1\right), \qquad (r,a) \in [-R,R] \times \mathbb{R},$$

is not convex in a, see Figure 1.5 (on p. 12). Technically, the preceding theorem is not applicable since f depends on r and a and $p = \infty$, but it is obvious that we can simply extend it to consider Young measures

$$(\delta_{r(t)} \otimes \nu_t)_{t\in(0,T)} \in \mathbf{Y}^\infty((0,T);\mathbb{R}\times\mathbb{R}) \quad \text{with} \quad \nu \in \mathbf{GY}^\infty((0,T)), \;\; [\nu] = r',$$

in a similar way to the strategy in Section 5.6. We collect all such product Young measures in the set $\overline{X}$, which we equip with the weak* convergence for Young measures, see (4.9).

Then, the extended–relaxed variational problem is

$$\begin{cases} \text{Minimize} \quad \overline{\mathscr{F}}[\delta_r \otimes \nu] := \langle\!\langle f, \delta_r \otimes \nu \rangle\!\rangle = \displaystyle\int_0^T \int f(r(t),a) \,\mathrm{d}\nu_t(a) \,\mathrm{d}t, \\ \text{over all} \qquad \delta_r \otimes \nu = (\delta_{r(t)} \otimes \nu_t)_t \in \overline{X} \text{ with } r(0) = r(T) = 0 \text{ and } |r(t)| \le R. \end{cases}$$

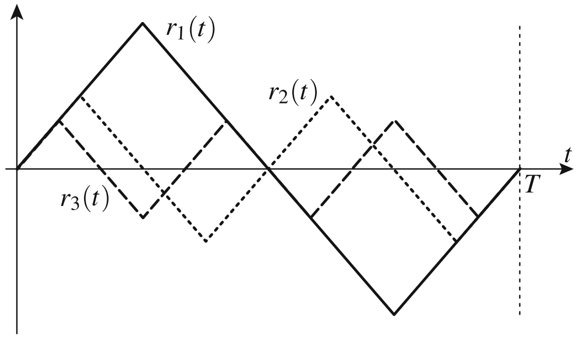


Fig. 7.2 The first few maps in the minimizing sequence for $\mathscr{F}$

Let us also construct a sequence of approximate solutions that generates the optimal Young measure solution. The first part of the integrand f, namely $v_{\max} \cdot (\cos(4 \arctan a) - 1)/2$, has two minima with value $-v_{\max}$ at $a = \pm 1$, see Figure 1.5 (on p. 12). The second part $v_{\text{flow}}(r^2/R^2 - 1)$ attains its minimum $-v_{\text{flow}}$ for $r = 0$. Thus,

$$f \geq -(v_{\max} + v_{\text{flow}}) =: f_{\min}.$$

We let

$$h(s) := \begin{cases} s & \text{if } s \in [0, 1], \\ 2 - s & \text{if } s \in (1, 3], \\ s - 4 & \text{if } s \in (3, 4], \end{cases}$$

and consider h to be extended to all $s \in \mathbb{R}$ by periodicity. Then set

$$r_j(t) := \frac{T}{4j} h\left(\frac{4j}{T} t\right), \qquad t \in [0, T],$$

see Figure 7.2. It is easy to see that $r_j' \in \{-1, 1\}$ and $r_j \to 0$ uniformly. Thus,

$$\mathscr{F}[r_j] \to T \cdot f_{\min} = \inf \mathscr{F}.$$

By the Fundamental Theorem 4.1 on Young measures, we deduce that we may select a subsequence of j's (not explicitly labeled) such that (see Lemma 5.19)

$$(r_j, r_j') \overset{\mathbf{Y}}{\to} \delta_r \otimes \nu = (\delta_{r(t)} \otimes \nu_t)_t \in \overline{X}.$$

From the construction of r_j we see that

$$\delta_r \otimes \nu = \delta_0 \otimes \left(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1}\right).$$

Clearly, this $\delta_r \otimes \nu$ is minimizing for $\overline{\mathscr{F}}$. Since above we have constructed a $\mathrm{W}^{1,\infty}$-bounded generating sequence for this optimal $\delta_r \otimes \nu$, we have also a posteriori justified our choice to work with L^∞-Young measures.

We close this section by showing how the two relaxation approaches are related.

Proposition 7.14. *Let $p \in (1,\infty)$ and let $h\colon \mathbb{R}^{m\times d} \to [0,\infty)$ be continuous and satisfy the p-growth and coercivity assumption*

$$\mu|A|^p \le h(A) \le M(1+|A|^p), \qquad A \in \mathbb{R}^{m\times d},$$

for some $\mu, M > 0$. Then, for all $F \in \mathbb{R}^{m\times d}$ there exists a homogeneous gradient Young measure $\nu^F \in \mathbf{GY}^p(B(0,1);\mathbb{R}^{m\times d})$ with $[\nu^F] = F$ and

$$Qh(F) = \int h \,\mathrm{d}\nu^F.$$

Proof. According to (7.1) and the remarks following it,

$$Qh(F) = \inf\left\{ \fint_{B(0,1)} h(F + \nabla\psi(z)) \,\mathrm{d}z \;\mathbf{:}\; \psi \in \mathrm{W}^{1,p}_0(B(0,1);\mathbb{R}^m) \right\}.$$

Let now $(\psi_j) \subset \mathrm{W}^{1,p}_0(B(0,1);\mathbb{R}^m)$ be a minimizing sequence in the above formula. This minimization problem is admissible in Theorem 7.12, which yields a Young measure minimizer $\nu \in \mathbf{GY}^p(B(0,1);\mathbb{R}^m)$ such that

$$Qh(F) = \fint_{B(0,1)} \int h \,\mathrm{d}\nu_x \,\mathrm{d}x.$$

It only remains to show that we can replace $(\nu_x)_x$ by a *homogeneous* Young measure ν^F. This, however, follows directly from the averaging principle for Young measures, Lemma 4.14. □

7.5 Characterization of Gradient Young Measures

In the previous section we replaced a minimization problem over a Sobolev space by its extension–relaxation, defined on the space of gradient Young measures, which is a strict subset of all Young measures, as we saw in Section 5.4. This subset of the space of Young measures is so far specified only *extrinsically*, i.e., through the existence of a generating sequence of gradients. The question arises whether there is also an *intrinsic* characterization of gradient Young measures. Intuitively, trying to understand gradient Young measures amounts to understanding the (asymptotic) oscillations that can occur in sequences of gradients, and it should be clear by now that this is a useful endeavor.

Recall that in Lemma 5.11 we showed that a homogeneous gradient Young measure $\nu \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m\times d})$ satisfies the Jensen-type inequality

$$h([\nu]) \le \int h \, d\nu$$

for all quasiconvex functions $h \colon \mathbb{R}^{m\times d} \to \mathbb{R}$ with p-growth. There, we interpreted this as an expression of the (generalized) convexity of h, in analogy with the classical Jensen inequality.

However, we may also switch to a dual point of view and consider the validity of the above Jensen-type inequality for quasiconvex functions as a property of gradient Young measures. The following result shows that this dual point of view is indeed valid and that the Jensen-type inequalities (essentially) characterize gradient Young measures.

Theorem 7.15 (Kinderlehrer–Pedregal 1991/1994 [157, 158]). *Assume that $\nu \in \mathbf{Y}^p(\Omega; \mathbb{R}^{m\times d})$, $p \in (1, \infty]$, is a Young measure with $[\nu] = \nabla u$ for some underlying deformation $u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$. Then, $\nu \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m\times d})$ if and only if for almost every $x \in \Omega$ the Jensen-type inequality*

$$h(\nabla u(x)) \le \int h \, d\nu_x \tag{7.18}$$

holds for all quasiconvex $h \colon \mathbb{R}^{m\times d} \to \mathbb{R}$ with p-growth if $p \in (1, \infty)$ (no growth condition if $p = \infty$).

Remark 7.16. It will follow from the proof that we only need to verify (7.18) for all quasiconvex $h \colon \mathbb{R}^{m\times d} \to \mathbb{R}$ such that

$$\lim_{|A|\to\infty} \frac{h(A)}{1 + |A|^p} \quad \text{exists (in } \mathbb{R}) \tag{7.19}$$

if $p \in (1, \infty)$. Note that by Lemma 5.11, the condition (7.19) is not needed for the Jensen-type inequality (7.18) to hold. For another strengthening of the Kinderlehrer–Pedregal theorem, see Problem 7.8.

The idea of the proof is to reduce to the case of *homogeneous* Young measures ν and to show that the set of homogeneous gradient Young measures is convex and weakly* closed in the set of homogeneous Young measures. Then, the Jensen-type inequalities express that ν cannot be separated from this set by an abstract "hyperplane" represented by a suitable integrand. Thus, the geometric Hahn–Banach theorem implies that ν actually lies in this set and hence must be a gradient Young measure. Note that this is a *non-constructive* argument, which does not produce a generating sequence.

For the functional analytic setup we define for $p \in (1, \infty)$ the following class of integrands:

$$\mathbf{I}^p(\mathbb{R}^{m\times d}) := \left\{ h \in \mathrm{C}(\mathbb{R}^{m\times d}) \;:\; \lim_{|A|\to\infty} \frac{h(A)}{1+|A|^p} \text{ exists (in } \mathbb{R}) \right\}, \tag{7.20}$$

which is a separable Banach space when equipped with the norm

$$\|h\|_{\mathbf{I}^p} := \left\| \frac{h}{1+|\bullet|^p} \right\|_\infty, \qquad h \in \mathbf{I}^p(\mathbb{R}^{m\times d}).$$

For the separability, see Problem 7.7. The set of homogeneous $\mathrm{W}^{1,p}$-gradient Young measures with barycenter $F \in \mathbb{R}^{m\times d}$ is defined as

$$\mathbf{GY}^p_{\mathrm{hom}}(F) := \left\{ \mu \in \mathscr{M}^1(\mathbb{R}^{m\times d}) \;:\; \mu \in \mathbf{GY}^p(B(0,1);\mathbb{R}^{m\times d}),\ [\mu] = F \right\},$$

which can be considered a subset of the dual space $\mathbf{I}^p(\mathbb{R}^{m\times d})^*$.

Lemma 7.17. *For any $F \in \mathbb{R}^{m\times d}$ the set $\mathbf{GY}^p_{\mathrm{hom}}(F)$ is convex and weakly* closed in $\mathbf{I}^p(\mathbb{R}^{m\times d})^*$.*

Proof. Step 1: Convexity. Recall that homogeneous Young measures have generating sequences on any bounded Lipschitz domain, see Lemma 4.14; we will use this fact several times in the sequel. Let $\mu_1, \mu_2 \in \mathbf{GY}^p_{\mathrm{hom}}(F)$ and $\theta \in (0,1)$. Choose a Lipschitz subdomain $D_1 \subset B(0,1)$ with

$$|D_1| = \theta\omega_d,$$

where we recall that $\omega_d := |B(0,1)|$. We assume that μ_1 is a Young measure on D_1 and μ_2 is a Young measure on $D_2 := B(0,1) \setminus D_1$ (see Lemma 4.14). Let $(u_j) \subset \mathrm{W}^{1,p}_{Fx}(D_1;\mathbb{R}^m)$, $(v_j) \subset \mathrm{W}^{1,p}_{Fx}(D_2;\mathbb{R}^m)$ with $\nabla u_j \xrightarrow{\mathbf{Y}} \mu_1$, $\nabla v_j \xrightarrow{\mathbf{Y}} \mu_2$ and $\{\nabla u_j\}_j, \{\nabla v_j\}_j$ L^p-equiintegrable, which can be constructed using Lemma 4.13. We define $(w_j) \subset \mathrm{W}^{1,p}_{Fx}(B(0,1);\mathbb{R}^m)$ through

$$w_j(x) := \begin{cases} u_j(x) & \text{if } x \in D_1, \\ v_j(x) & \text{if } x \in D_2, \end{cases}$$

which is a norm-bounded sequence. Thus, up to selecting a subsequence, we may assume that $\nabla w_j \xrightarrow{\mathbf{Y}} \nu \in \mathbf{GY}^p(B(0,1);\mathbb{R}^m)$ and that $\{\nabla w_j\}_j$ is L^p-equiintegrable. For all continuous $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ with p-growth we find

$$\begin{aligned} \fint_{B(0,1)} \int h \,\mathrm{d}\nu_x \,\mathrm{d}x &= \frac{1}{\omega_d} \cdot \lim_{j\to\infty} \left[\int_{D_1} h(\nabla u_j(x)) \,\mathrm{d}x + \int_{D_2} h(\nabla v_j(y)) \,\mathrm{d}y \right] \\ &= \theta \int h \,\mathrm{d}\mu_1 + (1-\theta) \int h \,\mathrm{d}\mu_2. \end{aligned}$$

Then use the averaging principle from Lemma 4.14 to get $\overline{\nu} \in \mathbf{GY}^p_{\mathrm{hom}}(F)$ with

$$\int h \, \mathrm{d}\overline{\nu} = \fint_{B(0,1)} \int h \, \mathrm{d}\nu_x \, \mathrm{d}x = \theta \int h \, \mathrm{d}\mu_1 + (1-\theta) \int h \, \mathrm{d}\mu_2$$

and the convexity of $\mathbf{GY}^p_{\mathrm{hom}}(F)$ follows.

Step 2: Weak-closedness.* Note that we need to show weak* *topological* closedness, not just sequential closedness.

If $\mu \in \mathbf{I}^p(\mathbb{R}^{m\times d})^*$ lies in the topological closure of $\mathbf{GY}^p_{\mathrm{hom}}(F)$, then we observe first that μ must be a probability measure. Indeed, the weak* topology on $\mathrm{C}_0(\mathbb{R}^{m\times d})^*$ is weaker than the weak* topology on $\mathbf{I}^p(\mathbb{R}^{m\times d})^*$, so μ is a positive measure. Moreover, $\mathbb{1} \in \mathbf{I}^p(\mathbb{R}^{m\times d})$, whereby $\mu \in \mathscr{M}^1(\mathbb{R}^{m\times d})$.

Take a countable collection $\{h_k\}_k$ that is dense in the separable Banach space $\mathbf{I}^p(\mathbb{R}^{m\times d})$. Then, by the topological definition of closure for the weak* (locally convex) topology on $\mathbf{I}^p(\mathbb{R}^{m\times d})^*$, for all $j \in \mathbb{N}$ there exists a $\mu_j \in \mathbf{GY}^p_{\mathrm{hom}}(F)$ with

$$\left| \int h_k \, \mathrm{d}(\mu_j - \mu) \right| \le \frac{1}{j} \qquad \text{for all } k \le j.$$

For every μ_j we find $u_j \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ with $\int_\Omega u_j \, \mathrm{d}x = 0$ such that

$$\left| \int_{B(0,1)} h_k(\nabla u_j) \, \mathrm{d}x - \int_{B(0,1)} \int h_k \, \mathrm{d}\mu_j \, \mathrm{d}x \right| \le \frac{1}{j} \qquad \text{for all } k \le j.$$

Additionally, we may assume, adding $h_0(A) := |A|^p$ to our collection $\{h_k\}_k$ and using the Poincaré inequality, that the sequence (u_j) is uniformly $\mathrm{W}^{1,p}$-bounded (see the proof of Proposition 2.5 for a more precise argument). Hence, up to a subsequence, $\nabla u_j \overset{\mathbf{Y}}{\to} \nu \in \mathbf{GY}^p(B(0,1);\mathbb{R}^{m\times d})$.

By the averaging principle, Lemma 4.14, we may furthermore assume that $\nu = \overline{\nu}$ is homogeneous and that (u_j) is the averaged generating sequence from the proof of the said lemma. Since the integrands h_k are independent of x, this does not change any of the above assertions. We have

$$\left| \int_{B(0,1)} h_k(\nabla u_j) \, \mathrm{d}x - \int_{B(0,1)} \int h_k \, \mathrm{d}\mu \, \mathrm{d}x \right| \le \frac{2}{j} \qquad \text{for all } k \le j,$$

and so, letting $j \to \infty$,

$$\int h_k \, \mathrm{d}\nu = \int h_k \, \mathrm{d}\mu \qquad \text{for all } k \in \mathbb{N}.$$

Thus, by the density of $\{h_k\}_k$, we conclude that $\mu = \nu$ is a homogeneous gradient Young measure. □

Proof of Theorem 7.15. By Lemma 5.11 only the sufficiency of (7.18) remains to be proved.

Step 1. We first prove the result for homogeneous Young measures and $p \in (1, \infty)$, so let $\mu \in \mathbf{Y}^p(B(0,1); \mathbb{R}^{m\times d}) \subset \mathscr{M}^1(\mathbb{R}^{m\times d})$ be homogeneous with

$$h(F) \le \int h \, \mathrm{d}\mu, \qquad F := [\mu], \tag{7.21}$$

for all quasiconvex $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ with p-growth. Now, for any $g \in \mathbf{I}^p(\mathbb{R}^{m\times d})$ set $g_\alpha := \max\{g, \alpha\}$, $\alpha \in \mathbb{R}$. By Lemma 7.1 (slightly generalized to h with $h \ge \alpha$, which is trivial) we know that Qg_α is quasiconvex. Then, (7.21) implies

$$Qg(F) \le Qg_\alpha(F) \le \int Qg_\alpha \, \mathrm{d}\mu \le \int g_\alpha \, \mathrm{d}\mu.$$

Hence, by the monotone convergence theorem,

$$Qg(F) \le \int g \, \mathrm{d}\mu \tag{7.22}$$

since $g_\alpha \downarrow g$ as $\alpha \downarrow -\infty$ (this also uses the p-growth of g).

Assume that μ is not a gradient Young measure. From the preceding lemma we know that $\mathbf{GY}^p_{\mathrm{hom}}(F)$ is convex and weakly* closed in $\mathbf{I}^p(\mathbb{R}^{m\times d})^*$. Then, applying the Hahn–Banach separation theorem in the version of Theorem A.1, there is a $g \in \mathbf{I}^p(\mathbb{R}^{m\times d})$ such that

$$\int g \, \mathrm{d}\mu < \inf_{\nu \in \mathbf{GY}^p_{\mathrm{hom}}(F)} \int g \, \mathrm{d}\nu.$$

In particular, we may test this with all $\nu := \overline{\delta_{F+\nabla\psi}}$ for $\psi \in \mathrm{W}^{1,\infty}_0(B(0,1); \mathbb{R}^m)$ (i.e. the homogeneous Young measures originating from the Riemann–Lebesgue Lemma 4.15) to see via (7.1) that

$$\int g \, \mathrm{d}\mu < Qg(F).$$

However, this contradicts (7.22).

Step 2. We now treat the inhomogeneous case for $p \in (1, \infty)$, but assuming that $u = 0$ almost everywhere. So let $\nu = (\nu_x)_x \in \mathbf{Y}^p(\Omega; \mathbb{R}^{m\times d})$ satisfy the Jensen-type inequality (7.18) for all quasiconvex $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ with p-growth. Take a countable collection $\{\phi_k \otimes h_k\}_{k\in\mathbb{N}}$ as in Lemma 4.7. By the first step we have that $\nu_x \in \mathbf{GY}^p(B(0,1); \mathbb{R}^{m\times d})$ for almost every $x \in \Omega$. Moreover, almost every $x \in \Omega$ is a Lebesgue point of the functions

$$x \mapsto \langle h_k, \nu_x \rangle \;\; (l \in \mathbb{N}) \qquad \text{and} \qquad x \mapsto \langle |\bullet|^p, \nu_x \rangle,$$

see Theorem A.20.

Fix $\varepsilon \in (0, 1)$ and cover Ω up to a negligible set with a countable Vitali collection of disjoint balls $B(a_n, r_n)$, that is,

$$\Omega = Z \cup \bigcup_{n=1}^{\infty} B(a_n, r_n), \qquad |Z| = 0,$$

where $a_n \in \Omega$, $r_n > 0$ are such that (7.18) holds at $x = a_n$, and

$$\left| \fint_{B(a_n,r_n)} \langle h_k, \nu_y \rangle \,\mathrm{d}y - \langle h_k, \nu_{a_n} \rangle \right| \le \varepsilon \qquad \text{for all } k \in \mathbb{N}, \tag{7.23}$$

as well as

$$\left| \fint_{B(a_n,r_n)} \langle |\bullet|^p, \nu_y \rangle \,\mathrm{d}y - \langle |\bullet|^p, \nu_{a_n} \rangle \right| \le \varepsilon. \tag{7.24}$$

These estimates can be achieved by the Lebesgue point property and the fact that in the Vitali cover we may choose every radius $r_n > 0$ to be arbitrarily small.

For each $n \in \mathbb{N}$ take a generating sequence $(v_j^{(n)}) \subset \mathrm{W}_0^{1,p}(B(0,1); \mathbb{R}^m)$ with $\nabla v_j^{(n)} \overset{\mathbf{Y}}{\to} \nu_{a_n}$, cf. Lemma 4.13. Define for $j \in \mathbb{N}$,

$$w_j^\varepsilon(x) := r_n v_j^{(n)}\left(\frac{x - a_n}{r_n}\right) \quad \text{if } x \in B(a_n, r_n) \ (n \in \mathbb{N}).$$

We then estimate for all $j \in \mathbb{N}$,

$$\begin{aligned}
\int_\Omega |\nabla w_j^\varepsilon|^p \,\mathrm{d}x &= \sum_{n=1}^{\infty} \int_{B(a_n,r_n)} \left| \nabla v_j^{(n)}\left(\frac{x - a_n}{r_n}\right) \right|^p \mathrm{d}x \\
&= \sum_{n=1}^{\infty} r_n^d \int_{B(0,1)} |\nabla v_j^{(n)}|^p \,\mathrm{d}y \\
&\le \sum_{n=1}^{\infty} |B(a_n, r_n)| \cdot \left(\langle |\bullet|^p, \nu_{a_n} \rangle + 1 \right) \\
&\le \sum_{n=1}^{\infty} \int_{B(a_n,r_n)} \langle |\bullet|^p, \nu_y \rangle \,\mathrm{d}y + (1 + \varepsilon)|\Omega| \\
&< \infty,
\end{aligned}$$

where we discarded some leading elements of $(v_j^{(n)})_j$ for every $n \in \mathbb{N}$ and also used (7.24). By the Poincaré inequality from Theorem A.26 (i) we thus get that $(w_j^\varepsilon)_j$ is uniformly norm-bounded in $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ and so, selecting a subsequence, we may assume that

$$\nabla w_j^\varepsilon \overset{\mathbf{Y}}{\to} \nu^\varepsilon \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m \times d}).$$

By a similar calculation as above, this time also using the uniform continuity of ϕ_k and (7.23), we get for every $k \in \mathbb{N}$,

$$\begin{aligned}
\langle\!\langle \phi_k \otimes h_k, \nu^\varepsilon \rangle\!\rangle &= \lim_{j\to\infty} \int_\Omega \phi_k(x) h_k(\nabla w_j^\varepsilon(x)) \,\mathrm{d}x \\
&= \lim_{j\to\infty} \sum_{n=1}^\infty r_n^d \biggl[\int_{B(0,1)} \phi_k(a_n) h_k(\nabla v_j^{(n)}(y)) \,\mathrm{d}y + E(\varepsilon) \biggr] \\
&= \sum_{n=1}^\infty |B(a_n, r_d)| \cdot \phi_k(a_n) \cdot \bigl\langle h_k, \nu_{a_n} \bigr\rangle + E(\varepsilon) \\
&= \sum_{n=1}^\infty \phi_k(a_n) \int_{B(a_n,r_n)} \bigl\langle h_k, \nu_y \bigr\rangle \,\mathrm{d}y + E(\varepsilon) \\
&= \int_\Omega \phi_k(y) \bigl\langle h_k, \nu_y \bigr\rangle \,\mathrm{d}y + E(\varepsilon) \\
&= \langle\!\langle \phi_k \otimes h_k, \nu \rangle\!\rangle + E(\varepsilon).
\end{aligned}$$

Here, $E(\varepsilon)$ is an error term that may change from line to line and vanishes as $\varepsilon \downarrow 0$. Thus, as $\varepsilon \downarrow 0$ we get that

$$\nu^\varepsilon \overset{*}{\rightharpoonup} \nu \quad \text{in } \mathbf{Y}^p(\Omega; \mathbb{R}^{m\times d}),$$

that is,

$$\langle\!\langle f, \nu^\varepsilon \rangle\!\rangle \to \langle\!\langle f, \nu \rangle\!\rangle \qquad \text{for all } f \in \mathrm{C}_0(\Omega \times \mathbb{R}^N),$$

see (4.9). As all the ν^ε are gradient Young measures with

$$\sup_{j\in\mathbb{N}} \langle\!\langle |\bullet|^p, \nu^\varepsilon \rangle\!\rangle < \infty,$$

a diagonal argument (similar to Step 2 in the proof of Lemma 7.17) yields that ν is also a gradient Young measure.

Step 3. Let now $p \in (1, \infty)$ and let u not be identically zero but, without loss of generality, finite everywhere. Then define the **shifted Young measure** $\widehat{\nu} = (\widehat{\nu}_x)_x \in \mathbf{Y}^p(\Omega; \mathbb{R}^{m\times d})$ by $\widehat{\nu}_x := \nu_x \star \delta_{-\nabla u(x)}$, that is,

$$\int h \,\mathrm{d}\widehat{\nu}_x = \int h(A - \nabla u(x)) \,\mathrm{d}\nu_x \qquad \text{for all } h \in \mathrm{C}_0(\mathbb{R}^{m\times d}),\ \text{a.e. } x \in \Omega.$$

We have $[\widehat{\nu}] = 0$ and the Jensen-type inequalities still hold for $\widehat{\nu}_x$ at almost every $x \in \Omega$: Let $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ be quasiconvex and have p-growth. Then, $\tilde{h}(A) := h(A - \nabla u(x))$ is also quasiconvex with p-growth, and so,

$$h([\widehat{\nu}_x]) = h([\nu_x] - \nabla u(x)) \leq \int h(A - \nabla u(x)) \, \mathrm{d}\nu_x(A) = \int h \, \mathrm{d}\widehat{\nu}_x.$$

Thus, the previous step applies to $\widehat{\nu}$. Consequently, $\widehat{\nu} \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m\times d})$ and there is a sequence $(w_j) \subset \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ such that $\nabla w_j \xrightarrow{\mathbf{Y}} \widehat{\nu}$. Then, for the inversely shifted maps

$$u_j(x) := w_j(x) + u(x)$$

we have $\nabla u_j \xrightarrow{\mathbf{Y}} \widehat{\nu} \star \delta_{\nabla u(x)} = \nu$ and thus $\nu \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m\times d})$. This finishes the proof for $p \in (1, \infty)$.

Step 4. For the sufficiency part of Theorem 7.15 in the case $p = \infty$, we simply apply the previous steps for one exponent $q \in (1, \infty)$. This yields that $\nu \in \mathbf{GY}^q(\Omega; \mathbb{R}^{m\times d})$. On the other hand, since $\nu \in \mathbf{Y}^\infty(\Omega; \mathbb{R}^{m\times d})$, there exists a compact set $K \subset \mathbb{R}^{m\times d}$ with $\operatorname{supp} \nu_x \subset K$ for almost every $x \in \Omega$, see the Fundamental Theorem 4.1. Zhang's lemma below then implies that $\nu \in \mathbf{GY}^\infty(\Omega; \mathbb{R}^{m\times d})$. □

For the case $p = \infty$, we still have to prove the following truncation result.

Lemma 7.18 (Zhang 1992 [284]). *Let $\nu \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m\times d})$ for some $p \in (1, \infty)$ and suppose that there exists a compact and convex convex set $K \subset \mathbb{R}^{m\times d}$ with*

$$\operatorname{supp} \nu_x \subset K \qquad \textit{for a.e. } x \in \Omega.$$

Then, $\nu \in \mathbf{GY}^\infty(\Omega; \mathbb{R}^{m\times d})$, that is, there exists a sequence $(u_j) \subset \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ with $\nabla u_j \xrightarrow{\mathbf{Y}} \nu$ and

$$\|\nabla u_j\|_{\mathrm{L}^\infty} \leq C|K|_\infty, \qquad \textit{where} \qquad |K|_\infty := \sup\big\{|A| : A \in K\big\},$$

and $C = C(d, m) > 0$ is a dimensional constant.

Remark 7.19. Zhang's result as stated here is far from being sharp. A refined version shows that in fact the constant C can be chosen arbitrarily close to 1 and for the sequence $(u_j) \subset \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ with $\nabla u_j \xrightarrow{\mathbf{Y}} \nu$ one can achieve $\operatorname{dist}(\nabla u_j, K) \to 0$ in L^∞. This is proved in [202].

Proof. Step 1. We suppose first that $[\nu] = 0$ and thus that there is a sequence $(v_j) \subset \mathrm{W}^{1,p}(\mathbb{R}^d; \mathbb{R}^m)$ with $v_j \rightharpoonup 0$ in $\mathrm{W}^{1,p}$ and $\nabla v_j \xrightarrow{\mathbf{Y}} \nu$. Define, using the maximal function from Appendix A.6,

$$V_j := M\big(|v_j| + |\nabla v_j|\big), \qquad j \in \mathbb{N},$$

and

$$G_j := \big\{ x \in \Omega : V_j(x) \leq 8|K|_\infty \big\}.$$

Theorem A.36 implies that v_j is Lipschitz continuous on G_j with Lipschitz constant $C|K|_\infty$ for some dimensional constant $C = C(d, m) > 0$. Hence, by the Kirszbraun Theorem A.34, we may extend $v_j|_{G_j}$ to a Lipschitz function $u_j \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ without changing the Lipschitz constant. Consequently, $\|\nabla u_j\|_{\mathrm{L}^\infty} \leq C|K|_\infty$.

By the weak-type estimate on the maximal function, see Theorem A.36, we furthermore get

$$\begin{aligned} |\Omega \setminus G_j| &\leq \bigl|\bigl\{ x \in \Omega \,:\, |Mv_j(x)| \geq 4|K|_\infty \bigr\}\bigr| \\ &\qquad + \bigl|\bigl\{ x \in \Omega \,:\, |M(\nabla v_j)(x)| \geq 4|K|_\infty \bigr\}\bigr| \\ &\leq \frac{C}{4|K|_\infty} \int_\Omega |v_j| \,\mathrm{d}x + \frac{C}{4|K|_\infty} \int_{\{|\nabla v_j| \geq 2|K|_\infty\}} |\nabla v_j| \,\mathrm{d}x \\ &\leq C \int_\Omega |v_j| \,\mathrm{d}x + C \int_\Omega g(|\nabla v_j|) \,\mathrm{d}x, \end{aligned} \tag{7.25}$$

where $g \colon [0, \infty) \to [0, \infty)$ is given as

$$g(s) := \begin{cases} 0 & \text{if } s < |K|_\infty, \\ 2(s - |K|_\infty) & \text{if } |K|_\infty \leq s < 2|K|_\infty, \\ s & \text{if } s \geq 2|K|_\infty. \end{cases}$$

The first term in (7.25) tends to zero since $v_j \to 0$ in L^1 by the compact embedding from $\mathrm{W}^{1,p}$ into L^1. By the Young measure representation (note that $(g(|\nabla v_j|))_j$ is uniformly L^p-bounded, hence equiintegrable), the second term converges to

$$\int_\Omega \int g(|A|) \,\mathrm{d}\nu_x(A) \,\mathrm{d}x = 0$$

since $\operatorname{supp} \nu_x \subset K$ for almost every $x \in \Omega$. Therefore, for all $\phi \in \mathrm{C}_0(\Omega)$ and $h \in \mathrm{C}_0(\mathbb{R}^m)$,

$$\int_\Omega |\phi h(\nabla u_j) - \phi h(\nabla v_j)| \,\mathrm{d}x \leq \|\phi\|_\infty \cdot \|h\|_\infty \cdot |\Omega \setminus G_j| \to 0,$$

and we may conclude that (∇u_j) generates ν, which therefore has been shown to lie in $\mathbf{GY}^\infty(\Omega; \mathbb{R}^{m \times d})$.

Step 2. If $[\nu] = \nabla u \neq 0$ for some $u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ (since $\nabla u \in K$ almost everywhere), we consider the shifted Young measure $\widehat{\nu} = (\widehat{\nu}_x)_x \in \mathbf{Y}^p(\Omega; \mathbb{R}^{m \times d})$ defined via $\widehat{\nu}_x := \nu_x \star \delta_{-\nabla u(x)}$, that is,

$$\int h \,\mathrm{d}\widehat{\nu}_x = \int h(A - \nabla u(x)) \,\mathrm{d}\nu_x \qquad \text{for all } h \in \mathrm{C}_0(\mathbb{R}^{m \times d}).$$

Then, since $\nabla u \in K$ almost everywhere (as we assumed that K is convex) we have

$$\operatorname{supp}\widehat{\nu}_x \subset K - K := \big\{\, A - B \,:\, A, B \in K \,\big\} \subset B(0, 2|K|_\infty).$$

Thus, the first step applies to $\widehat{\nu}$ and yields a sequence $(\widehat{u}_j) \subset \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ with $\nabla\widehat{u}_j \overset{\mathbf{Y}}{\to} \widehat{\nu} \in \mathbf{GY}^\infty(\Omega; \mathbb{R}^{m\times d})$ and $\|\nabla\widehat{u}_j\|_{\mathrm{L}^\infty} \leq 2C|K|_\infty$. Setting

$$u_j := \widehat{u}_j + u,$$

we get $\|\nabla\widehat{u}_j\|_{\mathrm{L}^\infty} \leq (2C+1)|K|_\infty$ and $\nabla u_j \overset{\mathbf{Y}}{\to} \nu \in \mathbf{GY}^\infty(\Omega; \mathbb{R}^{m\times d})$. This concludes the proof. □

Notes and Historical Remarks

Classically, one often defines the quasiconvex envelope through (7.3) and not as we have done through (7.1). In this case, (7.1) is usually called *Dacorogna's formula*, see Section 6.3 in [76] for further references.

The construction of Lemma 7.3 and Example 7.7 are from [250], but our proof also uses some ellipticity arguments similar to those in Lemma 2.7 of [203]. A different proof of Lemma 7.3 can be found in Section 5.3.9 of [76]. We refer to [33] for some regularity properties of quasiconvex envelopes.

Further relaxation formulas can be found in Chapter 11 of the textbook [19]; historically, Dacorogna's lecture notes [75] were also influential.

The conditions (i)–(iii) at the beginning of Section 7.4 are modeled on the concept of Γ-convergence (introduced by De Giorgi), see Chapter 13 for more on this topic.

The Kinderlehrer–Pedregal theorem is conceptually very important. In particular, it entails that if we could understand the class of quasiconvex functions, then we also could understand gradient Young measures and thus the asymptotic "shape" of gradients. Unfortunately, our knowledge of quasiconvex functions, and hence of gradient Young measures, is limited at present. There is some further discussion on this point throughout [222].

The truncation argument used in Zhang's Lemma 7.18 seems to be due originally to Acerbi–Fusco [1, 3]. The book [177] makes use of this technique in regularity theory and also contains several refinements.

Problems

7.1. Generalize Lemmas 7.1 and 7.2 to $h\colon \mathbb{R}^{m\times d} \to [0,\infty)$ with p-growth that are merely upper semicontinuous.

7.2. Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain and $F \in \mathbb{R}^{m\times d}$ with rank $F = 1$. Consider the functional

$$\mathscr{F}[u] := \int_\Omega \operatorname{dist}(\nabla u(x), \{-F, F\})^2 \, \mathrm{d}x, \qquad u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m).$$

Construct a sequence $(u_j) \subset \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ with $u_j \overset{*}{\rightharpoonup} 0$ in $\mathrm{W}^{1,\infty}$ and $\mathscr{F}[u_j] = 0$ for all $j \in \mathbb{N}$. Conclude that $\mathscr{F}$ is not lower semicontinuous with respect to weak* convergence in $\mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$.

7.3. Show that the integrand in the integral functional from the previous exercise is not rank-one convex.

7.4. Let L be a linear subspace of $\mathbb{R}^{m\times d}$ with $\operatorname{rank}(A - B) \geq 2$ for all $A, B \in L$, let $K \subset L$ be compact, and let $p \in (1, \infty)$. Show that for

$$h(A) := \operatorname{dist}(A, K)^p, \qquad A \in \mathbb{R}^{m\times d},$$

it holds that Qh is not convex (at zero). Conclude that for $p < 2$ and K not convex, Qh cannot be polyconvex.

7.5. Show that under the assumptions of Theorem 7.5 and assuming additionally that X is separable, it holds that

$$\mathscr{F}_*[u] = \inf \left\{ \liminf_{j\to\infty} \mathscr{F}[u_j] \,:\, u_j \rightharpoonup u \text{ in } X \right\}, \qquad u \in X.$$

7.6. Let $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ be twice continuously differentiable. Show that then h is rank-one convex if and only if h satisfies the Legendre–Hadamard condition, that is,

$$\begin{aligned} \mathrm{D}^2 h(A)[a\otimes b, a\otimes b] &= \frac{\mathrm{d}^2}{\mathrm{d}t^2} h(A + ta\otimes b)\bigg|_{t=0} \\ &= \sum_{i,k=1}^{m} \sum_{j,l=1}^{d} \frac{\partial^2 h(A)}{\partial A^i_j \partial A^k_l} a^i b^j a^k b^l \\ &\geq 0 \end{aligned} \tag{7.26}$$

for all $A \in \mathbb{R}^{m\times d}$ and all $a \in \mathbb{R}^m$, $b \in \mathbb{R}^d$.

7.7. Show that $\mathbf{I}^p(\mathbb{R}^{m\times d})$ (defined in (7.20)) is isomorphic to the (separable) space $\mathrm{C}(\alpha\mathbb{R}^{m\times d})$, where $\alpha\mathbb{R}^{m\times d}$ is the Alexandrov (one-point) compactification of $\mathbb{R}^{m\times d}$, or, equivalently, $\mathbf{I}^p(\mathbb{R}^{m\times d})$ is isometrically isomorphic to the set of all $\phi \in \mathrm{C}(\overline{\mathbb{B}^{m\times d}})$ with $\phi|_{\partial\mathbb{B}^{m\times d}}$ constant, where $\mathbb{B}^{m\times d}$ denotes the open unit ball in $\mathbb{R}^{m\times d}$ (with respect to the Frobenius norm). Conclude that $\mathbf{I}^p(\mathbb{R}^{m\times d})$ is separable.

7.8. Show that in the Kinderlehrer–Pedregal Theorem 7.15 in the case $p \in (1, \infty)$ it suffices to verify (7.18) for all quasiconvex $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ that are bounded from below and for which

$$\lim_{|A|\to\infty} \frac{h(A)}{1 + |A|} \quad \text{exists (in } \mathbb{R}).$$

7.9. Prove the claim of Remark 5.15.

7.10. Let $K \subset \mathbb{R}^{m \times d}$ be compact and non-empty. Also assume there is a norm-bounded sequence $(u_j) \subset \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$, $p \in (1, \infty)$ with $\operatorname{dist}(\nabla u_j, K) \to 0$ in measure. Show that then there exists a sequence $(v_j) \subset \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ such that also $\operatorname{dist}(\nabla v_j, K) \to 0$ in measure. *Hint:* Use Zhang's Lemma 7.18.

Part II
Advanced Topics

Chapter 8
Rigidity

We noted in several places that oscillations may develop in minimizing sequences. Now we will embark on a more detailed study of these oscillations. Inspired by (but not limited to) the example on crystalline microstructure in Section 1.8, our overarching philosophy is the following: Assume that we are trying to minimize the functional

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

where $f\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ ($d, m \geq 2$) is continuous, over a (Sobolev) class of functions $u\colon \Omega \to \mathbb{R}^m$ with prescribed boundary values. Here, as usual, we assume that $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain. We associate with $\mathscr{F}$ as above the *pointwise* **differential inclusion**

$$\nabla u(x) \in K := \bigl\{ A \in \mathbb{R}^{m\times d} \;:\; f(A) = \min f \bigr\}, \qquad x \in \Omega,$$

where $\min f$ denotes the pointwise minimum of f that we assume to exist in $\mathbb{R}$. Under a mild coercivity assumption on f we have that K is compact.

If $u\colon \Omega \to \mathbb{R}^m$ (with the prescribed boundary values) exists with $\nabla u \in K$ almost everywhere, then such a u clearly is a minimizer of $\mathscr{F}$. Of course, a minimizer u of $\mathscr{F}$ usually does not satisfy $\nabla u \in K$ almost everywhere, in particular if K is "small". However, the differential inclusion "$\nabla u \in K$" should hold at least in some approximate sense and any deviation of the gradient from K may be considered a perturbation. By analyzing solutions to the differential inclusion

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m), \\ \nabla u \in K \quad \text{in } \Omega \end{cases} \tag{8.1}$$

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8_8

for a non-empty compact set $K \subset \mathbb{R}^{m\times d}$, we can thus understand the "shape" of minimizers or approximate minimizers. Obviously, if $A \in K$ and $v_0 \in \mathbb{R}^m$, then $u(x) := v_0 + Ax$ solves (8.1) exactly. The question is whether these trivial solutions are the only solutions or whether there are other, non-affine ones. Here, we mostly focus on the $\mathrm{W}^{1,\infty}$-theory. By Zhang's Lemma 7.18 this is no restriction as long as K is compact, which is the most common case (see Problem 7.10).

There are two notions of solution for (8.1) that are relevant for our investigation:

- A map $u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ is an **exact solution** to (8.1) if the differential inclusion holds pointwise almost everywhere, that is,

$$\nabla u(x) \in K \qquad \text{for a.e. } x \in \Omega.$$

- A uniformly norm-bounded sequence $(u_j) \subset \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ is an **approximate solution** to (8.1) if

$$\mathrm{dist}(\nabla u_j, K) \to 0 \quad \text{in measure},$$

  that is, for every $\varepsilon > 0$,

$$\left|\left\{ x \in \Omega \;:\; \mathrm{dist}(\nabla u_j(x), K) > \varepsilon \right\}\right| \to 0 \qquad \text{as } j \to \infty.$$

There is no unified theory for (8.1) that is able to handle all possible sets K, but some techniques can be used repeatedly and we present a selection of them in this and the next chapter. Whereas this chapter focuses on situations where we have *rigidity*, that is, exact or approximate solutions to the differential inclusion under investigation are necessarily affine, the next chapter treats the complementary case where non-affine solutions do occur, which exhibit (usually very complex) microstructure.

After having considered a selection of differential inclusions as above, we also briefly touch on the related topic of *compensated compactness*.

8.1 Two-Gradient Inclusions

We start our investigation into differential inclusions with the smallest non-trivial set $K \subset \mathbb{R}^{m\times d}$ $(d, m \geq 2)$ and consider the **two-gradient inclusion**

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m), \quad A, B \in \mathbb{R}^{m\times d}, \;\; A \neq B, \\ \nabla u \in K := \{A, B\} \quad \text{in } \Omega. \end{cases} \tag{8.2}$$

By now it should come as no surprise that the behavior of (8.2) depends decisively on whether A, B are rank-one connected, that is, whether $\mathrm{rank}(A - B) = 1$.

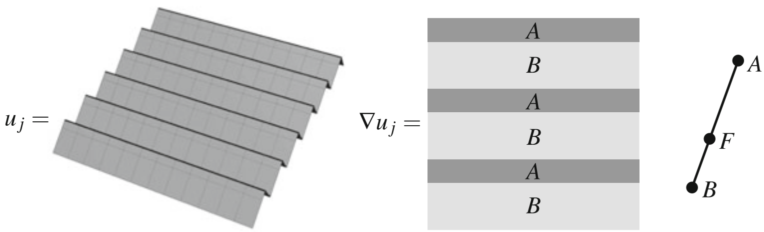


Fig. 8.1 The simple laminate u_j: map, gradient schematic, and rank-one diagram

Let us first assume that these matrices are not rank-one connected, i.e. $\operatorname{rank}(A - B) \geq 2$. Then, the Ball–James Rigidity Theorem 5.13 (i) (a) implies that any exact solution $u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ to (8.2) is in fact affine,

$$u(x) = v_0 + Fx \qquad \text{with } v_0 \in \mathbb{R}^m \text{ and } F = A \text{ or } F = B.$$

Turning to approximate solutions, let us assume that $(u_j) \subset \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ is uniformly norm-bounded and

$$\operatorname{dist}(\nabla u_j, \{A, B\}) \to 0 \quad \text{in measure.}$$

Then, assertion (ii) of the Ball–James rigidity theorem yields that, up to a subsequence,

$$\nabla u_j \to A \quad \text{in measure} \qquad \text{or} \qquad \nabla u_j \to B \quad \text{in measure.}$$

We now consider the case $\operatorname{rank}(A - B) \leq 1$, i.e., we assume that there exist $a \in \mathbb{R}^m$, $n \in \mathbb{S}^{d-1}$ such that

$$B - A = a \otimes n.$$

To see which solutions to (8.2) are now possible, recall the lamination construction from the proof of Proposition 5.3. There, on the oriented unit-volume cube Q_n with two faces orthogonal to n, we constructed an exact solution of (8.2). In fact, in the same way one can construct many solutions: For any $\theta \in (0, 1)$ let $F := \theta A + (1 - \theta)B$ and define $u_j \in \mathrm{W}^{1,\infty}(Q_n; \mathbb{R}^m)$, $j \in \mathbb{N}$, as the map

$$u_j(x) := Fx + \frac{1}{j}\psi(jx \cdot n)a \qquad x \in Q_n,$$

where

$$\psi(t) := \begin{cases} -(1-\theta)t & \text{if } t - \lfloor t \rfloor \in [0, \theta), \\ \theta t - \theta & \text{if } t - \lfloor t \rfloor \in (\theta, 1), \end{cases}$$

see Figure 8.1. For the gradients we compute

$$\nabla u_j(x) = \begin{cases} F - (1-\theta)a \otimes n = A & \text{if } jx \cdot n - \lfloor jx \cdot n \rfloor \in [0, \theta), \\ F - \theta a \otimes n = B & \text{if } jx \cdot n - \lfloor jx \cdot n \rfloor \in (\theta, 1). \end{cases}$$

Clearly, any such u_j solves (8.2) exactly.

We can further modify the sequence (u_j) to agree with Fx on ∂Q_n, for which it holds that $u_j \overset{*}{\rightharpoonup} Fx$ in $W^{1,\infty}$, but not $u_j \to Fx$ in measure. Finally, we may embed a rescaled copy of Q_n in the given Lipschitz domain Ω and extend the u_j by Fx to see that non-trivial approximate solutions also exist in other domains.

These observations inspire the following general definitions:

- The differential inclusion (8.1) is called **rigid for exact solutions** if all of its exact solutions are affine.
- The differential inclusion (8.1) is called **rigid for approximate solutions** if for all approximate solutions $(u_j) \subset W^{1,\infty}(\Omega; \mathbb{R}^m)$ with

  $$u_j \overset{*}{\rightharpoonup} u \in W^{1,\infty}(\Omega; \mathbb{R}^m) \quad \text{and} \quad u_j|_{\partial\Omega} = Fx \text{ for some } F \in \mathbb{R}^{m \times d},$$

  it holds that

  $$\nabla u_j \to \nabla u \quad \text{in measure} \quad \text{and} \quad \nabla u = \text{const} = F. \tag{8.3}$$

- The differential inclusion (8.1) is called **strongly rigid** if in the situation of the previous definition, (8.3) holds without any condition on the boundary values of the u_j.

Neither of the first two notions of rigidity implies the other, as we will see in Theorem 8.16, Proposition 8.17, and Problem 8.6. Of course, strong rigidity implies the other two rigidity notions.

With these definitions we can rephrase the Ball–James Rigidity Theorem 5.13 as follows:

Theorem 8.1 (Ball–James 1987 [30]).

(i) If $\operatorname{rank}(A - B) \geq 2$*, then the two-gradient inclusion (8.2) is strongly rigid; in particular, it is rigid for exact and for approximate solutions.*
(ii) If $\operatorname{rank}(A - B) = 1$*, then the two-gradient inclusion (8.2) is not rigid for exact and not rigid for approximate solutions.*

Note that in the definition of the rigidity for approximate solutions it is necessary to assume the a priori weak* convergence of (∇u_j); otherwise there are trivial counterexamples to the Ball–James rigidity theorem in the version above.

8.2 Linear Inclusions

After the basic two-gradient inclusions, we next consider **linear differential inclusions**, where the gradient of a $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$-map, $p \in [1,\infty]$, is restricted to lie in a linear subspace of $\mathbb{R}^{m\times d}$. Thus, in a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, we aim to solve

$$\begin{cases} u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m), \quad L \subset \mathbb{R}^{m\times d} \text{ a linear subspace,} \\ \nabla u \in L \quad \text{in } \Omega. \end{cases} \tag{8.4}$$

Since L is not compact, it is natural to look for solutions also in spaces of maps with unbounded gradients.

We first observe that if there is a nontrivial rank-one connection in L, that is, there are $A, B \in L$ with $\operatorname{rank}(A-B) = 1$, then we can construct a laminate whose gradient oscillates between A and B; this follows in the same fashion as in the previous section. Hence, in this case, the differential inclusion is not rigid for both exact and approximate solutions. We can even construct smooth solutions:

Example 8.2. Let $P_0 := a \otimes n$ with $a \in \mathbb{R}^m$, $n \in \mathbb{S}^{d-1}$. Then, for all $j \in \mathbb{N}$, the maps

$$u_j(x) := \frac{1}{j}\sin(jx\cdot n)a, \qquad x \in \mathbb{R}^d,$$

satisfy

$$\nabla u_j(x) = \cos(jx\cdot n)(a\otimes n) \in \operatorname{span}\{a\otimes n\} =: L.$$

We now prove a general result about (8.4). While necessarily falling short of rigidity in the sense of the previous section, it still expresses **ellipticity** of the differential inclusion, which can be interpreted as a weaker version of rigidity.

Theorem 8.3. *Let $L \subset \mathbb{R}^{m\times d}$ be a linear subspace that contains no rank-one line, that is, for all $A, B \in L$ with $A \neq B$ it holds that* $\operatorname{rank}(A-B) \geq 2$.

(i) If $u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$, $p \in [1,\infty]$, satisfies (8.4) exactly, then u is smooth.

(ii) If $u_j \rightharpoonup u$ in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ ($u_j \overset{*}{\rightharpoonup} u$ in $\mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^m)$ if $p = \infty$) and

$$\operatorname{dist}(\nabla u_j, L) \to 0 \quad \textit{in measure,}$$

then

$$\nabla u_j \to \nabla u \quad \textit{in measure} \qquad \textit{and} \qquad \nabla u \in L \quad \textit{a.e.}$$

Moreover, u is smooth.

Proof. The idea of the proof is that $\nabla u \in L$ can be written as the PDE

$$\mathbf{P}(\nabla u) = 0, \tag{8.5}$$

where we denote by $\mathbf{P}\colon \mathbb{R}^{m\times d} \to \mathbb{R}^{m\times d}$ the orthogonal projection onto the orthogonal complement $L^\perp$ of L. It turns out that under the assumptions of the theorem this PDE is *elliptic* and as such the usual higher-integrability estimates apply. However, since the argument is based on the Fourier-transform, which only operates on the whole space, we need to cut off suitably and this adds some complexity to the implementation of this strategy.

We will only prove the statement for $p \in (1, \infty)$. The case $p = \infty$ can obviously be reduced to the case $p = 2$ and for $p = 1$ the proof is the task of Problem 11.10.

Ad (i). For every smooth cut-off function $\rho \in \mathrm{C}_c^\infty(\Omega; [0, 1])$ with $\rho \equiv 1$ on an open set $U \Subset \Omega$, the function $w := \rho u$ satisfies

$$\nabla w = \rho \nabla u + u \otimes \nabla \rho.$$

Combining this with (8.5), we get

$$\mathbf{P}(\nabla w) = \mathbf{P}(u \otimes \nabla \rho) =: R \in \mathrm{L}^p(\mathbb{R}^d; \mathbb{R}^{m\times d}). \tag{8.6}$$

Hence, applying the Fourier transform to both sides of (8.6) and considering $\mathbf{P}$ to be identified with its complexification (that is, $\mathbf{P}(A + \mathrm{i}B) = \mathbf{P}(A) + \mathrm{i}\mathbf{P}(B)$ for $A, B \in \mathbb{R}^{m\times d}$), we arrive at

$$\mathbf{P}(\widehat{\nabla w}(\xi)) = (2\pi \mathrm{i})\, \mathbf{P}(\widehat{w}(\xi) \otimes \xi) = \widehat{R}(\xi). \tag{8.7}$$

Here we also used the fact that $\widehat{\nabla w}(\xi) = (2\pi \mathrm{i})\, \widehat{w}(\xi) \otimes \xi$ for $\xi \in \mathbb{R}^d$.

By a similar argument to the one in the proof of Lemma 7.3 (where now L is not necessarily one-dimensional, which, however, necessitates only minor modifications), we may rewrite (8.7) as the multiplier equation

$$\widehat{\nabla w}(\xi) = (2\pi \mathrm{i})\, \widehat{w}(\xi) \otimes \xi = \mathbf{M}(\xi) \widehat{R}(\xi), \tag{8.8}$$

where $\mathbf{M}(\xi)\colon \mathbb{R}^{m\times d} \to \mathbb{R}^{m\times d}$ is smooth and positively 0-homogeneous in $\xi \in \mathbb{R}^d \setminus \{0\}$. If $p = 2$ we may use Plancherel's identity (A.4) to estimate

$$\|\nabla w\|_{\mathrm{L}^2} = \|\widehat{\nabla w}\|_{\mathrm{L}^2} \le \|\mathbf{M}\|_\infty \|\widehat{R}\|_{\mathrm{L}^2} = \|\mathbf{M}\|_\infty \|R\|_{\mathrm{L}^2} \le C\|u\|_{\mathrm{L}^2}$$

for some constant $C > 0$ that depends on L though the operator norm of $\mathbf{P}$, and on the choice of ρ. If $p \in (1, \infty)$, we likewise infer from the Mihlin Multiplier Theorem A.35 that

$$\|\nabla w\|_{\mathrm{L}^p} \le C\|\mathbf{M}\|_{\mathrm{C}^{\lfloor d/2\rfloor+1}} \|R\|_{\mathrm{L}^p} \le C\|u\|_{\mathrm{L}^p}.$$

So, also using $\rho\nabla u = \nabla w - u \otimes \nabla\rho$, we get the estimate

$$\|\nabla u\|_{\mathrm{L}^p(U)} \leq \|\nabla w\|_{\mathrm{L}^p(\Omega)} + \|u \otimes \nabla\rho\|_{\mathrm{L}^p(\Omega)} \leq C\|u\|_{\mathrm{L}^p(\Omega)}$$

for some constant $C > 0$. Differentiating (8.5) in a weak sense (or using mollification), applying the above argument to $\partial_j u$ ($j = 1, \ldots, d$) and iterating (*bootstrapping* as in the argument for Corollary 3.13), we conclude that u is smooth.

Ad (ii). The assumptions imply by the dominated convergence theorem that

$$u_j \rightharpoonup u \quad \text{in } \mathrm{W}^{1,p} \qquad \text{and} \qquad \mathbf{P}(\nabla u_j) \to 0 \quad \text{in } \mathrm{L}^q \text{ for all } q \in (1, p).$$

Indeed, for the second assertion, we observe by Vitali's convergence theorem

$$\int_\Omega |\mathbf{P}(\nabla u_j)|^q \,\mathrm{d}x \leq \int_\Omega \mathrm{dist}(\nabla u_j, L)^q \,\mathrm{d}x \to 0$$

since the ∇u_j are L^p-uniformly bounded, and hence L^q-equiintegrable by Markov's inequality. Moreover, by the weak continuity of the linear operator $\mathbf{P}$, we get that

$$\mathbf{P}(\nabla u) = 0 \quad \text{a.e. in } \Omega. \tag{8.9}$$

Thus, for every smooth cut-off function $\rho \in \mathrm{C}_c^\infty(\Omega; [0, 1])$ we have

$$\mathbf{P}\big(\nabla(\rho(u_j - u))\big) = \rho\mathbf{P}(\nabla u_j) + \mathbf{P}((u_j - u) \otimes \nabla\rho) \to 0 \quad \text{in } \mathrm{L}^q.$$

Consequently, by a similar estimate as before, the ($\mathrm{L}^q \to \mathrm{L}^q$)-boundedness of the Fourier multiplier operator corresponding to the multiplier $\mathbf{M}(\xi)$ from (8.8) implies that also $\nabla(\rho(u_j - u)) \to 0$ in L^q. Together with (8.9) this yields the assertions of (ii). □

The preceding theorem has some immediate and interesting applications in the theory of elliptic PDEs:

Example 8.4. Let

$$L := \mathbb{R}^{d\times d}_{\mathrm{sym,dev}} := \big\{ A \in \mathbb{R}^{d\times d} \;:\; A^T = A,\ \mathrm{tr}\, A = 0 \big\}.$$

This linear space does not contain any rank-one connections: Assume that rank $(A - B) \leq 1$ for $A, B \in L$. Then, $A - B = a \otimes b$ for some $a, b \in \mathbb{R}^d$ and since $a \cdot b = \mathrm{tr}(a \otimes b) = \mathrm{tr}(A - B) = 0$, we must have that a is orthogonal to b. On the other hand, $b \otimes a = (A - B)^T = A - B = a \otimes b$, so a is also parallel to b, yielding $A = B$.

Any $v \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^d)$, $p \in [1, \infty]$, that satisfies

$$\nabla v \in L \quad \text{a.e. in } \Omega$$

is the gradient of a map $u \in \mathrm{W}^{2,p}(\Omega)$ satisfying

$$\Delta u = 0 \quad \text{a.e. in } \Omega,$$

i.e., u is harmonic, and vice versa. Indeed, if $\nabla v \in L$ almost everywhere, the symmetry in the definition of L implies that ∇v is in fact a Hessian. Thus, a $u \in \mathrm{W}^{2,p}(\Omega)$ exists with $\nabla u = v$. Then, $\Delta u = \operatorname{div} v = \operatorname{tr} \nabla v = 0$ almost everywhere. The other direction is obvious. Thus, Theorem 8.3 (i) yields a well-known result, namely that all (weakly) harmonic maps must be smooth; we already proved this in Example 3.15. The conclusion (ii) of Theorem 8.3 is perhaps less well-known for harmonic maps. Clearly, we can find examples of non-affine harmonic maps in $\mathrm{W}^{1,\infty}(\Omega)$, so Theorem 8.3 cannot in general be strengthened to yield the full rigidity for exact solutions.

Another related example of a linear differential inclusion is the topic of Problem 8.2.

A special case of the linear inclusion (8.4), which is often interesting, is the case when L is one-dimensional, i.e., the **polar differential inclusion**

$$\begin{cases} u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m), \quad \mathrm{P}_0 \in \mathbb{R}^{m \times d} \text{ a fixed matrix,} \\ \nabla u \in \operatorname{span}\{P_0\} = \{\, \lambda P_0 \,:\, \lambda \in \mathbb{R} \,\} \quad \text{in } \Omega. \end{cases} \tag{8.10}$$

Here, as before, $p \in [1, \infty]$. Equivalently to (8.10), we could require that u satisfies

$$\nabla u(x) = P_0 g(x), \qquad x \in \Omega,$$

for some *scalar* function $g \colon \Omega \to \mathbb{R}$, which then is an additional unknown in the problem. If we suppose (without loss of generality) that $|P_0| = 1$, this formulation explains the name "polar inclusion". An example of (8.10) was already exhibited in Example 8.2.

The reason why we are interested in such differential inclusions is that they occur naturally in a wide variety of variational problems as soon as we blow-up ("magnify") around a point, see the proof of Proposition 5.14 for an example of this technique and also Lemma 10.4.

For the polar inclusion, stronger results than for the general linear inclusion are available, summarized in the following theorem (with the various notions of rigidity suitably extended if $p < \infty$).

Theorem 8.5. *Let $\Omega \subset \mathbb{R}^d$ be open, bounded, and connected.*

(i) If $\operatorname{rank} P_0 \geq 2$, *then (8.10) is strongly rigid.*
(ii) If $\operatorname{rank} P_0 = 1$, *then (8.10) is not rigid for exact solutions and not rigid for approximate solutions. Moreover, if $P_0 = a \otimes n$ ($a \in \mathbb{R}^m$, $n \in \mathbb{S}^{d-1}$), then u is one-directional in direction n, i.e., there exist $h \in \mathrm{W}^{1,p}(\mathbb{R})$, $v_0 \in \mathbb{R}^m$ such that $u(x) = v_0 + h(x \cdot n)a$, $x \in \Omega$.*

The proof of this theorem is the task of Problem 8.3.

8.3 Relaxation and Quasiconvex Hulls of Sets

Before we can investigate more complicated differential inclusions, we need to develop a more sophisticated approach to rigidity, which is built on the theory of Young measures.

To motivate the basic strategy, we associate with our compact set $K \subset \mathbb{R}^{m\times d}$ the functional $\mathscr{F}^K : \mathrm{W}^{1,2}(\Omega;\mathbb{R}^m) \to \mathbb{R}$,

$$\mathscr{F}^K[u] := \int_\Omega \operatorname{dist}(\nabla u(x), K)^2 \,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,2}(\Omega;\mathbb{R}^m).$$

Then, just like in Chapter 7, we can ask about the relaxation $\mathscr{F}^K_*$ of $\mathscr{F}^K$. Theorem 7.6 (trivially extended to the weaker coercivity assumption $f(A) \ge \mu|A| - \mu^{-1}$ for a $\mu > 0$) tells us that

$$\mathscr{F}^K_*[u] = \int_\Omega Q \operatorname{dist}(\nabla u(x), K)^2 \,\mathrm{d}x,$$

where $Q \operatorname{dist}(A, K)^2 := Q[\operatorname{dist}(\,\cdot\,, K)^2](A)$ is the quasiconvex envelope of the integrand $\operatorname{dist}(\,\cdot\,, K)^2$ evaluated at $A \in \mathbb{R}^{m\times d}$, see (7.1). If we apply this chapter's overarching philosophy that the pointwise minimizer set for the integrand determines the admissible microstructure, we can now define a *relaxation* $\overline{K}$ of K via

$$\overline{K} := \bigl\{ F \in \mathbb{R}^{m\times d} \;:\; Q \operatorname{dist}(F, K)^2 = 0 \bigr\} \supset K.$$

According to Proposition 7.14, for all $F \in \mathbb{R}^{m\times d}$ there exists a homogeneous gradient Young measure $\mu^F \in \mathbf{GY}^2(B(0,1);\mathbb{R}^{m\times d})$ with $[\mu^F] = F$ and

$$Q \operatorname{dist}(F, K)^2 = \int \operatorname{dist}(\,\cdot\,, K)^2 \,\mathrm{d}\mu^F.$$

If $Q \operatorname{dist}(F, K)^2 = 0$, then μ^F is supported on K, that is, $\operatorname{supp}\mu^F \subset K$. By Zhang's Lemma 7.18, we may also assume that $\mu^F \in \mathbf{GY}^\infty(B(0,1);\mathbb{R}^{m\times d})$.

On the other hand, for any $\mu \in \mathbf{GY}^2(B(0,1);\mathbb{R}^{m\times d})$ with $[\mu] = F \in \mathbb{R}^{m\times d}$ and $\operatorname{supp}\mu \subset K$, there exists a sequence $(\psi_j) \subset \mathrm{W}^{1,\infty}_0(B(0,1);\mathbb{R}^m)$ with $\|\nabla\psi_j\|_{\mathrm{L}^\infty}$ uniformly bounded and $F + \nabla\psi_j \overset{\mathbf{Y}}{\to} \mu$, see Lemmas 4.13, 7.18 (observe that the truncation construction in Zhang's Lemma 7.18 can be performed so that the boundary values remain intact). By (7.1) we get

$$Q \operatorname{dist}(F, K)^2 \le \lim_{j\to\infty} \fint_{B(0,1)} \operatorname{dist}(F + \nabla\psi_j(z), K)^2 \,\mathrm{d}z = \int \operatorname{dist}(\,\cdot\,, K)^2 \,\mathrm{d}\mu = 0.$$

Thus, we have shown that $\overline{K}$ consists precisely of all barycenters of homogeneous $\mathrm{W}^{1,\infty}$-gradient Young measures that are supported on K. This set $\overline{K}$ is called the **quasiconvex hull** K^{qc} of K,

$$K^{\mathrm{qc}} := \{ [\mu] : \mu \in \mathscr{M}^{\mathrm{qc}}(K) \},$$

where

$$\mathscr{M}^{\mathrm{qc}}(K) := \{ \mu \in \mathbf{GY}^{\infty}(B(0,1);\mathbb{R}^{m\times d}) : \mu \text{ homogeneous, supp } \mu \subset K \}$$

is the set of homogeneous $\mathrm{W}^{1,\infty}$-gradient Young measures supported on K, which is a subset of $\mathscr{M}^1(K)$, the set of all probability measures supported on K. Note that by the Kinderlehrer–Pedregal Theorem 7.15,

$$\mathscr{M}^{\mathrm{qc}}(K) = \left\{ \mu \in \mathscr{M}^1(K) : h([\mu]) \le \int h \,\mathrm{d}\mu \text{ for all quasiconvex } h \in \mathrm{C}(\mathbb{R}^{m\times d}) \right\},$$

which explains the notation "$\mathscr{M}^{\mathrm{qc}}(K)$". It is the task of Problem 8.7 to show that for a compact set $K \subset \mathbb{R}^{m\times d}$, K^{qc} is also compact.

Example 8.6. Let $K = \{A, B\}$ with $\mathrm{rank}(A - B) \ge 2$. Then, by Theorem 8.1, $K^{\mathrm{qc}} = K$. Thus, in general $K^{\mathrm{qc}} \ne K^{**}$, the convexification of K.

The above discussion leads to the "Young measure approach" to (approximate) rigidity of the differential inclusion

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^m), \quad K \subset \mathbb{R}^{m\times d} \text{ compact and non-empty,} \\ \nabla u \in K \quad \text{in } \Omega, \end{cases} \tag{8.11}$$

as expressed in the following lemmas.

Lemma 8.7. *Assume that every measure in $\mathscr{M}^{\mathrm{qc}}(K)$ is a Dirac mass. Then:*

(i) $K^{\mathrm{qc}} = K$.
(ii) If $u \in \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^m)$ *satisfies*

$$\nabla u \in K \quad \textit{a.e.} \qquad \textit{and} \qquad u|_{\partial\Omega} = Fx$$

for some $F \in \mathbb{R}^{m\times d}$, *then* $\nabla u = F$ *almost everywhere.*

Note that (ii) is not the same as rigidity for exact solutions since here we additionally assume linear boundary values.

Proof. Ad (i). For $F \in K^{\mathrm{qc}}$ by definition there exists a $\mu \in \mathscr{M}^{\mathrm{qc}}(K)$ with $[\mu] = F$. By assumption, μ is the Dirac mass δ_F and, since $\mathrm{supp}\,\mu \subset K$, necessarily $F \in K$. Thus, $K^{\mathrm{qc}} \subset K$; the opposite inclusion is trivial.

Ad (ii). Thanks to the linear boundary values of u, we may apply the Riemann–Lebesgue lemma for gradient Young measures, Lemma 4.15. This yields $\overline{\nu} := \overline{\delta[\nabla u]} \in \mathscr{M}^{\mathrm{qc}}(K)$ with $[\overline{\nu}] = F$. By assumption, $\overline{\nu} = \delta_F$. Hence, ∇u must be constant (see (4.17)). □

Lemma 8.8. *The following statements are true:*

(i) The differential inclusion (8.11) is rigid for approximate solutions if and only if every measure in $\mathscr{M}^{\mathrm{qc}}(K)$ is a Dirac mass.
(ii) The differential inclusion (8.11) is strongly rigid if and only if it is rigid for exact solutions and every measure in $\mathscr{M}^{\mathrm{qc}}(K)$ is a Dirac mass.

Proof. Ad (i). Assume first that (8.11) is rigid for approximate solutions. To see the assertion concerning Young measures, let $\nu \in \mathscr{M}^{\mathrm{qc}}(K)$ with $[\nu] = F \in \mathbb{R}^{m\times d}$. By Lemmas 4.13, 7.18 (the truncation in Zhang's Lemma 7.18 can be performed so that the boundary values remain intact) there exists a sequence $(\psi_j) \subset \mathrm{W}^{1,\infty}_0(B(0,1);\mathbb{R}^m)$ such that the norms $\|\nabla\psi_j\|_{\mathrm{L}^\infty}$ are uniformly bounded and $F + \nabla\psi_j \overset{\mathbf{Y}}{\to} \nu$. Moreover,

$$\psi_j \overset{*}{\rightharpoonup} 0 \quad \text{in } \mathrm{W}^{1,\infty} \qquad \text{and} \qquad \operatorname{dist}(F + \nabla\psi_j, K) \to 0 \quad \text{in measure}$$

by Lemma 4.12. Thus, the rigidity for approximate solutions implies that

$$F + \nabla\psi_j \to F \quad \text{in measure.}$$

Another application of Lemma 4.12 then implies $\nu = \delta_F$.

For the converse implication, assume that every measure in $\mathscr{M}^{\mathrm{qc}}(K)$ is a Dirac mass and let $(u_j) \subset \mathrm{W}^{1,\infty}_{Fx}(\Omega;\mathbb{R}^m)$ for some $F \in \mathbb{R}^{m\times d}$ such that

$$u_j \overset{*}{\rightharpoonup} u \in \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^m) \qquad \text{and} \qquad \operatorname{dist}(\nabla u_j, K) \to 0 \quad \text{in measure.}$$

Up to selecting a subsequence, we may further assume $\nabla u_j \overset{\mathbf{Y}}{\to} \nu \in \mathbf{GY}^\infty(\Omega;\mathbb{R}^{m\times d})$ with $[\nu_x] = \nabla u(x)$ and $\operatorname{supp}\nu_x \subset K$ for almost every $x \in \Omega$. Almost every ν_x is a homogeneous gradient Young measure in it own right, $\nu_x \in \mathscr{M}^{\mathrm{qc}}(K)$, by the blow-up principle in Proposition 5.14. Then, by assumption, $\nu_x = \delta_{\nabla u(x)}$ and so, Lemma 4.12 implies that in fact $\nabla u_j \to \nabla u$ in measure. Since $u|_{\partial\Omega} = Fx$, part (ii) of the preceding Lemma 8.7 immediately yields that ∇u is constant almost everywhere. Thus, we have established that (8.11) is rigid for approximate solutions.

Ad (ii). The proof of the first direction is identical since rigidity for exact and approximate solutions clearly follows from strong rigidity. For the other direction we can no longer assume that $u_j|_{\partial\Omega} = Fx$ for some $F \in \mathbb{R}^{m\times d}$. However, the only place where we used this was when we proved that ∇u is a constant (via Lemma 8.7) and this now follows directly from the assumed rigidity for exact solutions. □

In assertion (ii) of the preceding lemma, the assumption of exact rigidity cannot be dispensed with, see Problem 8.6.

Corollary 8.9. *The differential inclusion (8.11) is strongly rigid if and only if it is rigid for exact solutions and rigid for approximate solutions.*

It is sometimes possible, by "testing" with carefully selected quasiconvex functions, to show that any Young measure supported on a set K necessarily must be a Dirac mass. Together with exact rigidity, which is often elementary, this then implies strong rigidity and $K^{\mathrm{qc}} = K$. We will see examples of this strategy in the following sections.

Example 8.10. In the physical context of crystalline microstructure as described in Section 1.8, the quasiconvex hull K^{qc} of K contains the deformations with almost zero energy. Indeed, from the above reasoning we get that precisely for $F \in K^{\mathrm{qc}}$ there exists a sequence $(\psi_j) \subset \mathrm{W}_0^{1,\infty}(B(0,1); \mathbb{R}^m)$ with $\|\nabla\psi_j\|_{\mathrm{L}^\infty}$ uniformly bounded, and such that

$$\int_{B(0,1)} \operatorname{dist}(F + \nabla\psi_j(z), K)^2 \, \mathrm{d}z \to 0 \qquad \text{as } j \to \infty.$$

Consequently, if the potential elastic energy is measured by the above integral, at least close to a pointwise minimizer, then the material can realize the linear boundary values Fx with almost no expenditure of energy, but the internal structure of the material may be very complicated. Thus, if we want to understand the macroscopic behavior of a crystal that develops microstructure, then the quasiconvex hull of the pointwise minimizer set of the integrand gives valuable information. For instance, if K^{qc} contains an open set, then we have very soft, fluid-like behavior since deformation gradients in this set cost almost no energy.

8.4 Multi-point Inclusions

We now investigate multi-point differential inclusions, where K has finitely many elements. So, we are trying to solve

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m), \quad A_1, \ldots, A_N \in \mathbb{R}^{m\times d}, \\ \nabla u \in K := \{A_1, \ldots, A_N\} \quad \text{in } \Omega \end{cases} \tag{8.12}$$

for (small) $N \in \mathbb{N}$. This is called the **N-gradient problem**. As discussed for the two-gradient problem in Section 8.1, when there is a rank-one connection in K, i.e., if $A_i - A_j = a \otimes n$ for some $i, j \in \{1, \ldots, N\}$, $a \in \mathbb{R}^m \setminus \{0\}$, $n \in \mathbb{S}^{d-1}$, then we can construct laminates whose gradients oscillate between A_i and A_j. In this case, the differential inclusion is always not rigid for exact and not rigid for approximate solutions. Thus, in the following we focus on the situation where there are no rank-one connections in K, that is, the following **incompatibility relation** holds:

$$\operatorname{rank}(A_i - A_j) \geq 2 \qquad \text{for all } i, j \in \{1, \ldots, N\} \text{ with } i \neq j. \tag{8.13}$$

For the three-gradient problem the situation is comparable to the two-gradient problem, but the proof is much more involved.

Theorem 8.11 **(Švérak 1991 [249]).** *Assume that* $K := \{A_1, A_2, A_3\} \subset \mathbb{R}^{m\times d}$ *contains no rank-one connection. Then, the inclusion (8.12) is strongly rigid and* $K^{\mathrm{qc}} = K$.

We will prove the rigidity for exact solutions and the rigidity for approximate solutions separately, which suffices for strong rigidity by Corollary 8.9. We start with the rigidity for exact solutions, for which we first prove a dimension-reduction lemma.

Lemma 8.12. *Let* $K = \{A_1, \ldots, A_N\} \subset \mathbb{R}^{m\times d}$ *contain no rank-one connections. If* $u \in \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^m)$ *is a non-affine map with* $\nabla u \in K$ *almost everywhere, then there exists a set* $\tilde{K} = \{\tilde{A}_1, \ldots, \tilde{A}_N\} \subset \mathbb{R}^{2\times 2}$ *without rank-one connections and a non-affine map* $\tilde{u} \in \mathrm{W}^{1,\infty}((0,1)^2;\mathbb{R}^2)$ *with* $\nabla\tilde{u} \in \tilde{K}$ *almost everywhere.*

Proof. As u is not affine, we may find $y_1, y_2 \in B(x_0, \varepsilon) \subset B(x_0, 3\varepsilon) \subset \Omega$ for a sufficiently small $\varepsilon > 0$ and $x_0 \in \Omega$ such that

$$u(y_1 + y_2 - x_0) + u(x_0) \neq u(y_1) + u(y_2).$$

Now pick $P_0 \in \mathbb{R}^{d\times 2}$, $Q_0 \in \mathbb{R}^{2\times m}$ with

$$P_0\mathrm{e}_1 = y_1 - x_0, \qquad P_0\mathrm{e}_2 = y_2 - x_0,$$

and

$$Q_0\bigl[u(y_1 + y_2 - x_0) + u(x_0) - u(y_1) - u(y_2)\bigr] \neq 0.$$

Since rank-two (invertible) matrices are dense in $\mathbb{R}^{2\times 2}$ and $K \times K$ is a finite set, we may find $P \in \mathbb{R}^{d\times 2}$, $Q \in \mathbb{R}^{2\times m}$ close to P_0, Q_0, respectively, such that, possibly slightly lowering $\varepsilon > 0$, the following conditions hold:

(a) $\mathrm{rank}(Q(A_i - A_j)P) = 2$ for all $i \neq j$;
(b) $Pz + x \in \Omega$ for all $z \in (0,1)^2$, $x \in B(x_0, \varepsilon)$;
(c) the non-affinity condition

$$Q\bigl[u(P\mathrm{e}_1 + P\mathrm{e}_2 + x) + u(x) - u(P\mathrm{e}_1 + x) - u(P\mathrm{e}_2 + x)\bigr] \neq 0 \qquad (8.14)$$

holds for all $x \in B(x_0, \varepsilon)$.

For almost every $x_1 \in \mathbb{R}^d$ and almost every $z \in \mathbb{R}^2$ (the exceptional negligible set may depend on x_1) with $Pz + x_1 \in \Omega$ we have

$$\nabla_z[Qu(Pz + x_1)] \in \tilde{K} := \{QA_1P, \ldots, QA_NP\} \subset \mathbb{R}^{2\times 2}$$

since otherwise $\nabla u \in K$ would be violated on a set of non-zero measure. Pick $x_1 \in B(x_0, \varepsilon)$ with this property and observe that the map $\tilde{u} \colon (0,1)^2 \to \mathbb{R}^2$ given as

$$\tilde{u}(z) := Qu(Pz + x_1)$$

is not affine by (8.14). □

Proof of Theorem 8.11: *Rigidity for exact solutions*. Assume that there is a non-affine $u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ with $\nabla u \in K$ almost everywhere. By Lemma 8.12 we can assume that $d = m = 2$ and $\Omega = (0,1)^2$.

We may suppose without loss of generality that $A_1 = 0$ (this transforms u into $\tilde{u}(x) := u(x) - A_1 x$), whereby $\det A_2, \det A_3 \neq 0$ by the incompatibility relation (8.13), and that $A_2 = \mathrm{Id}$ (this transforms $\tilde{u}$ into $\hat{\mathrm{u}}(x) := \tilde{u}(A_2^{-1}x)$). Finally, by a change of variables and utilizing the Jordan normal form, see Appendix A.1, we may assume that A_3 has one of the following two forms:

$$A_3 = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \qquad \text{with} \qquad a \neq 0, \ c \notin \{0, 1\},$$

or

$$A_3 = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \qquad \text{with} \qquad a^2 + b^2 \neq 0.$$

Here, the additional conditions on the coefficients follow from the incompatibility relation (8.13).

In the first case, for $u = (u^1, u^2)$ we have $\partial_1 u^2 = 0$ since all matrices A_1, A_2, A_3 have a zero as their $(2,1)$-element. Hence, with a slight abuse of notation, $u^2(x) = u^2(x_2)$. Clearly, since $c \notin \{0,1\}$, the value of $\partial_2 u^2(x) = \partial_2 u^2(x_2)$ determines which of the matrices A_1, A_2, A_3 our gradient $\nabla u(x)$ takes at x. Thus, $\nabla u(x)$ depends only on x_2 and therefore

$$\partial_1^2 u = \partial_2 \partial_1 u = \partial_1 \partial_2 u = 0 \qquad \text{in the sense of distributions.}$$

It follows that

$$\nabla u(x) = \begin{pmatrix} a & \partial_2 u^1(x_2) \\ 0 & \partial_2 u^2(x_2) \end{pmatrix} = (a, 0) \otimes \mathrm{e}_1 + \partial_2 u(x_2) \otimes \mathrm{e}_2$$

for some $a \in \mathbb{R}$. Hence, $\operatorname{rank}(\nabla u(x) - \nabla u(y)) \leq 1$ for all $x, y \in \Omega$. By our incompatibility assumption, this is only possible if ∇u is constant.

In the second case, we see that $\nabla u \in L$, where

$$L := \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} : a, b \in \mathbb{R} \right\} \subset \mathbb{R}^{2 \times 2}$$

once we observe that $A_1, A_2, A_3 \in L$. As it can be easily shown that L does not contain rank-one connections, Theorem 8.3 implies that u is smooth. As K is disconnected, however, ∇u must be constant. □

The proof of rigidity for approximate solutions is more involved. We will establish this result by a dimension-reduction to (2×2)-matrices and then employ the *separation method*, which entails testing with special quasiconvex functions that vanish on K. Only points where this test function is less than or equal to zero can potentially lie in K^{qc}, yielding an upper bound on K^{qc}.

We start by proving a useful rigidity lemma.

Lemma 8.13. *Let $K \subset \mathbb{R}^{2\times 2}$ be compact and non-empty such that*

$$\det(A - B) > 0 \qquad \textit{for all } A, B \in K \textit{ with } A \neq B.$$

Then, K is rigid for approximate solutions.

Proof. By Lemma 8.8 (i) it suffices to check that all $\nu \in \mathscr{M}^{\mathrm{qc}}(K)$ are Dirac masses. We will use the elementary formula

$$\det(A + B) = \det A + \operatorname{cof} A : B + \det B, \qquad A, B \in \mathbb{R}^{2\times 2}. \tag{8.15}$$

From the fact that $0 \le \det(A - B)$ for all $A, B \in \operatorname{supp} \nu$ in conjunction with (8.15) and Corollary 5.12, we see that

$$\begin{aligned}
0 &\le \int \int \det(A - B) \, d\nu(A) \, d\nu(B) \\
&= \int \int \det A - \operatorname{cof} A : B + \det B \;\; d\nu(A) \, d\nu(B) \\
&= \int \det [\nu] - \operatorname{cof} [\nu] : B + \det B \;\; d\nu(B) \\
&= \det [\nu] - \operatorname{cof} [\nu] : [\nu] + \det [\nu] \\
&= \det([\nu] - [\nu]) \\
&= 0.
\end{aligned}$$

Thus, $\det(A - B) = 0$ for $(\nu \otimes \nu)$-almost every $(A, B) \in K \times K$. On the other hand, by assumption, $\det(A - B) > 0$ for all such A, B with $A \neq B$. Thus, ν must be a Dirac mass. □

We will also need a special quasiconvex function on symmetric matrices:

Lemma 8.14. *Define* $\det^{++} : \mathbb{R}^{2\times 2}_{\mathrm{sym}} \to [0, \infty)$ *by*

$$\det{}^{++}(A) := \begin{cases} \det A & \textit{if } A \textit{ is positive semidefinite}, \\ 0 & \textit{otherwise}, \end{cases} \qquad A \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}.$$

Then, $\det^{++}$ *is* ***quasiconvex on symmetric matrices****, that is,*

$$\det^{++}(A) \leq \fint_{B(0,1)} \det^{++}(A + \nabla^2\psi(z)) \,\mathrm{d}z \tag{8.16}$$

for all $A \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}$ *and all* $\psi \in \mathrm{W}^{2,2}_c(B(0,1))$.

Here and in the following we set

$$\mathrm{W}^{2,q}_c(\Omega;\mathbb{R}^m) := \left\{ u \in \mathrm{W}^{2,q}(\Omega;\mathbb{R}^m) \;:\; \operatorname{supp} u \Subset \Omega \right\}$$

for $\Omega \subset \mathbb{R}^d$ a bounded Lipschitz domain and $q \in [1,\infty]$.

Proof. Let $A \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}$ be positive definite (the assertion is trivial otherwise). We will only show (8.16) for $\psi \in \mathrm{C}^\infty_c(B(0,1))$; the general case follows by approximation. Since we are dealing with symmetric matrices, which can be orthogonally diagonalized, we may assume that $A = \mathrm{Id}$ via a coordinate transformation. Define

$$u(z) := \frac{1}{2}|z|^2 + \psi(z), \qquad z \in B(0,1),$$

and set

$$D := \left\{ z \in B(0,1) \;:\; \nabla^2 u(z) \text{ positive semidefinite} \right\}.$$

We claim that $B(0,1) \subset \nabla u(D)$. Indeed, let $x_0 \in B(0,1)$ be arbitrary and take a point $z_0 \in B(0,1)$ such that $z \mapsto u(z) - x_0 \cdot z$ attains its minimum at z_0. Note that such a minimizer z_0 exists in $B(0,1)$ since $z \mapsto \frac{1}{2}|z|^2 - x_0 \cdot z$ attains its minimum in $B(0,1)$ and ψ has compact support. Differentiating, we get $\nabla u(z_0) = x_0$. Moreover, $\nabla^2 u(z_0)$ is positive semidefinite. Thus, $B(0,1) \subset \nabla u(D)$ and then

$$\omega_d \leq |\nabla u(D)| \leq \int_D \det \nabla^2 u(z) \,\mathrm{d}z = \int_{B(0,1)} \det^{++}(\mathrm{Id} + \nabla^2\psi(z)) \,\mathrm{d}z.$$

Consequently, (8.16) holds for $A = \mathrm{Id}$. □

Next, we prove the analogue of the Jensen-type inequality from Lemma 5.11 for functions that are quasiconvex on symmetric matrices.

Lemma 8.15. *Let* $K \subset \mathbb{R}^{d\times d}_{\mathrm{sym}}$ *and* $\mu \in \mathscr{M}^{\mathrm{qc}}(K)$. *Then, for all* $h \colon \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}$ *that are quasiconvex on symmetric matrices and have* p*-growth for some* $p \in [1,\infty)$*, the Jensen-type inequality*

$$h([\mu]) \leq \int h \,\mathrm{d}\mu$$

holds.

Proof. We will show that for every $q \in (1,\infty)$ there exists a norm-bounded sequence $(\psi_j) \subset (\mathrm{W}^{1,2}_0 \cap \mathrm{W}^{2,q})(B(0,1))$ such that $F + \nabla^2\psi_j \overset{\mathbf{Y}}{\to} \mu$, where $F := [\mu]$. Then, if $q > p$,

$$h(F) \le \lim_{j\to\infty} \fint_{B(0,1)} h(F + \nabla^2\psi_j(z))\,\mathrm{d}z = \int h\,\mathrm{d}\mu,$$

where the inequality follows by a cut-off procedure as in Lemma 4.13 (Step 3 of the proof) and the quasiconvexity on symmetric matrices. This implies the claim.

Assume without loss of generality that $[\mu] = F = 0$ and let the sequence $(v_j) \subset \mathrm{C}_c^\infty(B(0,1);\mathbb{R}^d)$ be such that $F + \nabla v_j \xrightarrow{\mathbf{Y}} \mu$ (see Lemmas 4.13 and 7.18). Define $\psi_j \in (\mathrm{W}_0^{1,2} \cap \mathrm{W}^{2,2})(B(0,1))$ as the (unique) solution of

$$\begin{cases} \Delta\psi_j = \operatorname{div} v_j & \text{in } B(0,1), \\ \psi_j \ = 0 & \text{on } \partial B(0,1). \end{cases}$$

By standard results (we proved this in Examples 3.4 and 3.15, also taking into account Section 6.3.2 in [111] for the regularity up to the boundary), such a solution ψ_j exists in $(\mathrm{W}_0^{1,2} \cap \mathrm{W}^{2,2})(B(0,1))$ and $\|\psi_j\|_{\mathrm{W}^{2,2}} \le C\|v_j\|_{\mathrm{W}^{1,2}}$. It is a well-known fact that even

$$\|\psi_j\|_{\mathrm{W}^{2,q}} \le C_q\|v_j\|_{\mathrm{W}^{1,\infty}}$$

for any $q \in (1,\infty)$ and a constant $C_q > 0$. This can be seen by the bootstrapping procedure explained in Section 3.2 and the fact that $\mathrm{W}^{k,2}(B(0,1))$ embeds into $\mathrm{W}^{2,q}(B(0,1))$ for sufficiently large k, see Theorem A.27.

For $w_j := v_j - \nabla\psi_j$ we have $\operatorname{div} w_j = 0$ (this gives the Helmholtz decomposition of v_j). Thus, with the usual definition $\operatorname{curl} w_j := \nabla w_j - (\nabla w_j)^T$, we may calculate, using integration by parts,

$$\begin{aligned}
&\int_{B(0,1)} |\operatorname{curl} w_j|^2\,\mathrm{d}x \\
&\quad = \int_{B(0,1)} |\nabla w_j - (\nabla w_j)^T|^2 + 2(\operatorname{div} w_j)^2\,\mathrm{d}x \\
&\quad = \int_{B(0,1)} |\nabla w_j|^2 - 2\nabla w_j : (\nabla w_j)^T + |(\nabla w_j)^T|^2 + 2\nabla w_j : (\nabla w_j)^T\,\mathrm{d}x \\
&\quad = \int_{B(0,1)} 2|\nabla w_j|^2\,\mathrm{d}x.
\end{aligned}$$

On the other hand, $\operatorname{curl} w_j = \operatorname{curl} v_j$ and by Young measure representation,

$$\begin{aligned}
\int_{B(0,1)} |\operatorname{curl} w_j|^2\,\mathrm{d}x &= \int_{B(0,1)} |\nabla v_j - (\nabla v_j)^T|^2\,\mathrm{d}x \\
&\to \int_{B(0,1)} \int |A - A^T|^2\,\mathrm{d}\mu(A)\,\mathrm{d}x = 0
\end{aligned}$$

since μ is supported on symmetric matrices. Thus, $\nabla w_j \to 0$ in L^2 and $F + \nabla^2\psi_j = F + \nabla v_j - \nabla w_j$ generates μ, as desired. □

We can now complete the proof of Theorem 8.11.

Proof of Theorem 8.11: *Rigidity for approximate solutions.* We will show that every Young measure in $\mathscr{M}^{\mathrm{qc}}(K)$ is a Dirac mass, which implies the claim by Lemma 8.8 (i).

We may assume that $A_3 = 0$ (by an affine shift). Let L be the linear subspace of $\mathbb{R}^{m\times d}$ spanned by the matrices A_1, A_2. We distinguish three cases.

Case 1: L contains at most one rank-one direction. If there is a rank-one direction in L, then let $L_0 \subset L$ be the rank-one line through the origin; otherwise set $L_0 := \{0\}$. Denote by $\mathbf{P}\colon \mathbb{R}^{m\times d} \to L$ and $\mathbf{Q}\colon L \to L_0$ the orthogonal projections onto L and L_0, respectively. With

$$
\begin{aligned}
g(A) &:= |A - \mathbf{Q}(A)|^2, \qquad A \in L,\\
f_{\varepsilon,k}(A) &:= -g(\mathbf{P}(A)) + \varepsilon|\mathbf{QP}(A)|^2 + k|A - \mathbf{P}(A)|^2, \qquad A \in \mathbb{R}^{m\times d},
\end{aligned}
$$

we claim that for every $\varepsilon > 0$ there is a $k \in \mathbb{N}$ such that for all $\mu \in \mathscr{M}^{\mathrm{qc}}(K)$ with $F := [\mu]$ the Jensen-type inequality

$$
f_{\varepsilon,k}(F) \le \int f_{\varepsilon,k} \,\mathrm{d}\mu \tag{8.17}
$$

holds. Then, since μ is supported in L, whereby also $F \in L$, we may let $\varepsilon \downarrow 0$ to conclude the reverse Jensen-type inequality for g, namely

$$
g(F) \ge \int g \,\mathrm{d}\mu.
$$

Since g is strictly convex, this implies that μ is a Dirac mass.

It remains to show (8.17). For this, we first note that for rank-one matrices $A \in \mathbb{R}^{m\times d}$ with $|A| = 1$ and $-g(\mathbf{P}(A)) + \varepsilon|\mathbf{QP}(A)|^2 \le 0$, it holds that $|A - \mathbf{P}(A)|^2 \ge c$ for some constant $c > 0$ that does not depend on A. Indeed, $|\mathbf{P}(A) - \mathbf{QP}(A)|^2 = g(\mathbf{P}(A)) \ge \varepsilon|\mathbf{QP}(A)|^2$ implies that A has positive distance from L_0. On the other hand, the rank-one cone intersects L only in L_0, so that we obtain that A has positive distance from L, hence $|A-\mathbf{P}(A)|^2 > 0$. Then we can minimize over the compact set of the A as above to obtain that $|A - \mathbf{P}(A)|^2 \ge c > 0$ for all A as above. Thus, since every term in the definition of $f_{\varepsilon,k}$ is positively 2-homogeneous, for every $\varepsilon > 0$ we can find $k \in \mathbb{N}$ such that $f_{\varepsilon,k}$ is positive on the rank-one cone.

We now have two ways to conclude. As the first option, we may invoke Tartar's theorem from the theory of compensated compactness in the version of Corollary 8.31 in Section 8.8 below, to see that

$$
f_{\varepsilon,k}(F) \le \liminf_{j\to\infty} \fint_{B(0,1)} f_{\varepsilon,k}(\nabla u_j) \,\mathrm{d}x
$$

for a sequence $(u_j) \subset \mathrm{W}^{1,\infty}_{Fx}(B(0,1);\mathbb{R}^m)$ such that $\nabla u_j \overset{\mathbf{Y}}{\to} \mu$ (also use Lemma 4.13 and Zhang's Lemma 7.18, keeping the boundary values intact). This directly implies (8.17).

As the second option, the positivity of $f_{\varepsilon,k}$ on the rank-one cone implies that $f_{\varepsilon,k}$ is rank-one convex (one can check the Legendre–Hadamard condition (7.14)), which in turn implies that $f_{\varepsilon,k}$ is also quasiconvex via Problem 5.7. Then, (8.17) follows from the Jensen-type inequality for gradient Young measures proved in Lemma 5.11.

Case 2: $d = m = 2$ and L contains the rank-one directions $\mathrm{e}_1 \otimes \mathrm{e}_1$, $\mathrm{e}_2 \otimes \mathrm{e}_2$. We can assume that $\det A_1, \det A_2 > 0$ and $A_3 = 0$. Indeed, at least two of the three numbers $\det(A_i - A_j)$ for $i, j \in \{1,2,3\}$ and $i < j$ must have the same sign. Multiplying the first row of the matrices by -1 (which corresponds to flipping the sign of the first component of a generating sequence of μ) and exchanging indices, we may thus assume that $\det(A_1 - A_3) > 0$, $\det(A_2 - A_3) > 0$. Then we shift all matrices by $-A_3$ to obtain $\tilde{A}_1 = A_1 - A_3$, $\tilde{A}_2 = A_2 - A_3$, $\tilde{A}_3 = 0$, of which the first two have positive determinant. In the following we drop the tildes.

We can further multiply the first and second components of generating sequences for μ with scalars to reduce to the situation where $K = \{A_1, A_2, A_3\}$ with

$$A_1 = \mathrm{Id}, \qquad A_2 = \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}, \qquad A_3 = 0$$

for

$$\alpha > \beta > 0 \qquad \text{or} \qquad 0 > \alpha > \beta$$

since $\det A_2 > 0$. Note also that $\alpha = 1$ or $\beta = 1$ is impossible since $\operatorname{rank}(A_2 - A_1) = 2$.

If $\alpha < 1$, then also $\beta < 1$ and so all differences $A_i - A_j$ for $i, j \in \{1,2,3\}$ have positive determinant. Thus, Lemma 8.13 implies that any $\mu \in \mathscr{M}^{\mathrm{qc}}(K)$ is a Dirac mass. If $\alpha > 1$ and $\beta > 1$, a similar argument applies, again yielding that any $\mu \in \mathscr{M}^{\mathrm{qc}}(K)$ is a Dirac mass.

It remains to investigate the situation where $\alpha > 1$ and $0 < \beta < 1$. First, for any $\mu \in \mathscr{M}^{\mathrm{qc}}(K)$, Corollary 5.12 shows that $\det F = \langle \det, \mu \rangle$, where $F := [\mu]$. Writing

$$\mu = \theta_1 \delta_{A_1} + \theta_2 \delta_{A_2} + \theta_3 \delta_{A_3}$$

for $\theta_1, \theta_2, \theta_3 \in [0,1]$ with $\theta_1 + \theta_2 + \theta_3 = 1$, this implies the relation

$$(\theta_1 + \theta_2 \alpha)(\theta_1 + \theta_2 \beta) = \theta_1 + \theta_2 \alpha\beta. \tag{8.18}$$

Second, from Lemma 8.14 we infer that the function $h \colon \mathbb{R}^{2\times 2}_{\mathrm{sym}} \to \mathbb{R}$ given by

$$h(A) := \det{}^{++}(A - X), \qquad \text{where} \qquad X := \begin{pmatrix} 1 & \\ & \beta \end{pmatrix},$$

is quasiconvex on symmetric matrices. By the Jensen-type inequality proved in Lemma 8.15 (applied to h), we thus get

$$\det{}^{++}(F-X) \le \int \det{}^{++}(A-X)\,\mathrm{d}\mu(A) = 0,$$

where the last equality holds since the three matrices $A_1 - X$, $A_2 - X$, $A_3 - X$ are not positive definite. Hence, also

$$F - X = \begin{pmatrix} \theta_1 + \theta_2\alpha - 1 & \\ & \theta_1 + \theta_2\beta - \beta \end{pmatrix}$$

cannot be positive definite, so at least one entry in this matrix needs to be nonpositive. Combining this fact with (8.18) we get that either

$$\theta_1 + \theta_2\alpha\beta \le \theta_1 + \theta_2\beta \qquad \text{or} \qquad \theta_1 + \theta_2\alpha\beta \le \theta_1\beta + \theta_2\alpha\beta.$$

In the first case, $\alpha > 1$ implies that $\theta_2 = 0$, which, however, by (8.18) immediately gives $\theta_1^2 = \theta_1$, whereby μ must be a Dirac mass. In the second case, from $\beta < 1$ we get $\theta_1 = 0$ and hence (8.18) implies $\theta_2^2 = \theta_2$, which again yields that μ is a Dirac mass.

Case 3: $d \ge 2, m \ge 2$ and L contains at least two (different) rank-one directions. We will reduce this case to the previous one via a dimension-reduction argument. Let $a_1 \otimes b_1$ and $a_2 \otimes b_2$ be two rank-one matrices in L with a_1, a_2 and b_1, b_2 linearly independent. By a change of variables we can assume that a_1, b_1 are the first unit vector e_1 (in $\mathbb{R}^m$ and $\mathbb{R}^d$, respectively) and a_2, b_2 are the second unit vector e_2. Then, A_1 and A_2 are diagonal,

$$A_1 = \begin{pmatrix} \gamma_1 & \\ & \gamma_2 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} \eta_1 & \\ & \eta_2 \end{pmatrix}$$

with $\gamma_1, \gamma_2, \eta_1, \eta_2 \in \mathbb{R} \setminus \{0\}$.

Let $(u_j) \subset \mathrm{W}^{1,\infty}((0,1)^d;\mathbb{R}^m)$ be such that $\nabla u_j \overset{\mathbf{Y}}{\to} \mu$. Then, for $z = (z_3, \ldots, z_d) \in (0,1)^{d-2}$, we define $v_j^{(z)} \in \mathrm{W}^{1,\infty}((0,1)^2;\mathbb{R}^2)$ by

$$v_j^{(z)}(x_1, x_2) := \begin{pmatrix} u_j^1(x_1, x_2, z) \\ u_j^2(x_1, x_2, z) \end{pmatrix}, \qquad (x_1, x_2) \in (0,1)^2.$$

For $\mathscr{L}^{d-2}$-almost every $z \in (0,1)^{d-2}$ it holds that $\nabla v_j^{(z)} \overset{\mathbf{Y}}{\to} \nu^{(z)} \in \mathscr{M}^{\mathrm{qc}}(\widehat{K})$ with

$$\widehat{K} := \left\{ \begin{pmatrix} \gamma_1 & \\ & \gamma_2 \end{pmatrix}, \begin{pmatrix} \eta_1 & \\ & \eta_2 \end{pmatrix}, 0 \right\} \subset \mathbb{R}^{2\times 2}.$$

Then, Case 2 is applicable and we conclude that $\nu^{(z)}$ is a Dirac mass. Since $\mu = \mathscr{L}_z^{d-2} \llcorner (0,1)^{d-2} \otimes \nu^{(z)}$ and we know that μ is homogeneous, we get that μ is also a Dirac mass. $\square$

For the four-gradient problem, rigidity for exact solutions obtains as well:

Theorem 8.16 **(Chlebík–Kirchheim 2002 [63]).** *Assume that the four-element set $K := \{A_1, A_2, A_3, A_4\} \subset \mathbb{R}^{m\times d}$ contains no rank-one connection. Then, (8.12) is rigid for exact solutions.*

The proof is beyond the scope of this book. It uses a reduction to the Monge–Ampère equation (like the original proof for the three-gradient problem in [249]), where one has to distinguish the elliptic case $\det \nabla^2 u > 0$ and the hyperbolic case $\det \nabla^2 u < 0$, see [63, 160] for the details.

Despite this result on rigidity for exact solutions, approximate solutions of the four-gradient inclusion are *not* rigid:

Proposition 8.17 **(T_4-configuration).** *Let $K_{\mathrm{T4}} := \{A_1, A_2, A_3, A_4\} \subset \mathbb{R}^{2\times 2}$ for the diagonal matrices*

$$A_1 := \begin{pmatrix} 3 & \\ & 1 \end{pmatrix}, \qquad A_2 := \begin{pmatrix} -1 & \\ & 3 \end{pmatrix}, \qquad A_3 := -A_1, \qquad A_4 := -A_2,$$

which do not have rank-one connections. Then, (8.12) for $K := K_{\mathrm{T4}}$ is not rigid for approximate solutions, that is, there exists a weakly converging sequence $(u_j) \subset \mathrm{W}^{1,\infty}((0,1)^2;\mathbb{R}^2)$ such that*

$$\operatorname{dist}(\nabla u_j, K_{\mathrm{T4}}) \to 0 \qquad \textit{in measure},$$

but (∇u_j) does not converge in measure. Moreover, $K_{\mathrm{T4}}^{\mathrm{qc}} \supsetneq K_{\mathrm{T4}}$.

Proof We define the intermediate matrices

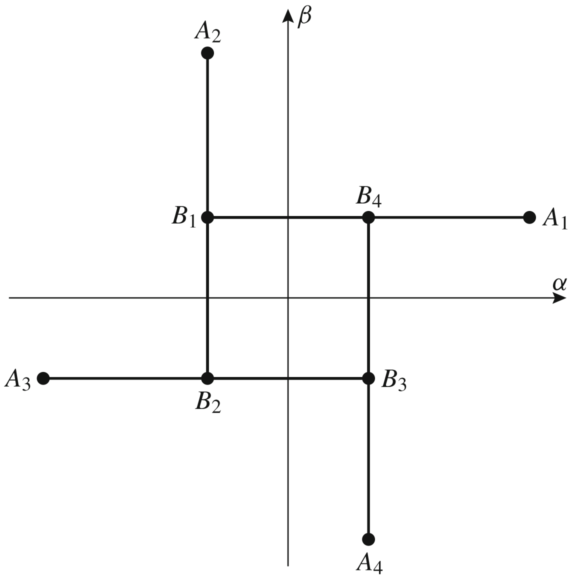
$$B_1 := \begin{pmatrix} -1 & \\ & 1 \end{pmatrix}, \qquad B_2 := \begin{pmatrix} -1 & \\ & -1 \end{pmatrix}, \qquad B_3 := -B_1, \qquad B_4 := -B_2,$$

see Figure 8.2.

The idea of the proof is to build a laminate of *infinite* order supported on K_{T4} and with barycenter B_4. First write B_4 as the rank-one convex combination between A_1 and B_1, which is possible since $B_4 = (A_1 + B_1)/2$ (note that $\operatorname{rank}(A_1 - B_1) = 1$). We have $A_1 \in K_{\mathrm{T4}}$ and we may write $B_1 = (A_2 + B_2)/2$. Again, $A_2 \in K_{\mathrm{T4}}$. We iterate this procedure and construct the corresponding laminates to obtain a sequence that satisfies the differential inclusion approximately.

To implement this strategy rigorously, we use the construction from the proof of Proposition 5.3 for $A, B \in \mathbb{R}^{2\times 2}$ with $A - B = a \otimes n$, where $a \in \mathbb{R}^2 \setminus \{0\}$, $n \in \mathbb{S}^{d-1}$, and $\theta \in (0,1)$. In this way, for any $\varepsilon > 0$ let $v \in \mathrm{W}^{1,\infty}_{Fx}(D_n;\mathbb{R}^2)$, where $D_n \subset \mathbb{R}^2$ is any rectangular parallelepiped with two faces orthogonal to n and $F := \theta A + (1-\theta)B$, be a map such that the following conditions hold:

Fig. 8.2 The rank-one diagram of the T_4-configuration in the space of diagonal matrices with entries α, β (solid lines signify rank-one connections)



(a) $\|v\|_{W^{1,\infty}} \le C(1 + |A| + |B|)$ for some $C > 0$;

(b) $\left| D_n \setminus \left(\bigcup_{i=1}^{N} E_i^A \cup \bigcup_{i=1}^{N} E_i^B \right) \right| \le \varepsilon$, where E_i^A, E_i^B are rectangular parallelepipeds with two faces orthogonal to n;

(c) $\nabla v = A$ almost everywhere in E_i^A, $\nabla v = B$ almost everywhere in E_i^B.

Fix $\delta > 0$. Apply this construction with $A := A_1$, $B := B_1$, $\theta := 1/2$ (hence $F = B_4$), and $\varepsilon := \delta/2$ in $D_n := (0, 1)^2$. This yields $u_1^{(\delta)}$ satisfying the above properties. Then apply the construction again in every $E_i^{B_1}$, $i = 1, \ldots, N_1$, with $A := A_2$, $B := B_2$, $\theta := 1/2$ (hence $F = B_1$), and $\varepsilon := \delta/(2^2 N_1)$ to get $u_2^{(\delta)}$ defined in $\bigcup_{i=1}^{N} E_i^{B_i}$. Note that we may just replace $u_1^{(\delta)}$ with $u_2^{(\delta)}$ in E^{B_1} since the boundary values agree. We proceed with this construction, successively eliminating more and more of the intermediate matrices B_1, B_2, B_3, B_4; in the k'th step we use $\varepsilon := \delta 2^{-k}/N_{k-1}$ on all the N_{k-1} pieces with an intermediate matrix in it.

Thus, we have constructed a sequence $(u_l^{(\delta)})_l \subset \mathrm{W}^{1,\infty}((0, 1)^2; \mathbb{R}^2)$ that is uniformly bounded, $u_l^{(\delta)}|_{\partial(0,1)^2} = B_4 x$, and

$$\left|\left\{ x \in (0, 1)^2 \;:\; \nabla u_l^{(\delta)}(x) \notin K_{\mathrm{T4}} \right\}\right| \le \sum_{k=1}^{l} \delta 2^{-k} + 2^{-l} \le \delta + 2^{-l}.$$

Since $(u_l^{(\delta)})_l$ is uniformly bounded in $\mathrm{W}^{1,\infty}$, we may pass to a weakly* converging subsequence with limit $u^\delta \in \mathrm{W}^{1,\infty}((0, 1)^2; \mathbb{R}^2)$. Then,

$$u^{(\delta)}|_{\partial\Omega} = B_4 x, \qquad \|u^{(\delta)}\|_{\mathrm{W}^{1,\infty}} \le C$$

with a δ-independent $C > 0$, and it also holds that

$$\left|\left\{ x \in (0,1)^2 \;:\; \nabla u^{(\delta)}(x) \notin K_{\mathrm{T4}} \right\}\right| \le \delta.$$

Now set $u_j := u^{(1/j)} \in \mathrm{W}^{1,\infty}((0,1)^2; \mathbb{R}^2)$, which is a uniformly bounded sequence with $u_j|_{\partial\Omega} = B_4 x$ and

$$\operatorname{dist}(\nabla u_j, K_{\mathrm{T4}}) \to 0 \quad \text{in measure.}$$

We can furthermore assume that (∇u_j) converges weakly* in $\mathrm{W}^{1,\infty}$ and generates a homogeneous gradient Young measure $\mu \in \mathscr{M}^{\mathrm{qc}}(K_{\mathrm{T4}})$ with $[\mu] = B_4 \notin K_{\mathrm{T4}}$. This in fact holds for the above construction, but one may also invoke Lemma 4.14 to ensure the homogeneity. Then, μ cannot be a Dirac mass, $K_{\mathrm{T4}}^{\mathrm{qc}} \supsetneq K_{\mathrm{T4}}$, and Lemma 8.8 (i) immediately yields that the inclusion $\nabla u \in K_{\mathrm{T4}}$ is not rigid for approximate solutions.

For the last step, one can also argue by observing that if, up to selecting a subsequence, we had $\nabla u_j \to \nabla u$ in measure as $j \to \infty$ for some $u \in \mathrm{W}^{1,\infty}((0,1)^2; \mathbb{R}^2)$, then it would hold that $\nabla u \in K_{\mathrm{T4}}$ almost everywhere and $u|_{\partial\Omega} = B_4 x$, which is impossible since $\nabla u \in K_{\mathrm{T4}}$ is rigid for exact solutions by Problem 8.4. □

Proposition 8.17 provides a *negative* answer to the following question of historical importance (as Tartar himself discovered in [271]).

Conjecture 8.18 (*Tartar 1982 [268]*). If a non-empty compact set $K \subset \mathbb{R}^{2\times 2}$ does not contain rank-one directions, then $\nabla u \in K$ is rigid for approximate solutions.

However, Tartar's conjecture turns out to be true for connected sets, see Problem 8.8 (i). It is a remarkable result that (generalized) rank-one connections and T_4-configuration are in fact the only obstructions to rigidity for approximate solutions, as was shown by Faraco and Székelyhidi Jr. in 2008 [114].

Finally, the five-gradient problem loses all rigidity:

Theorem 8.19 (**Kirchheim–Preiss 2003 [160]**). *There are matrices* $A_1, \ldots, A_5 \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}$ *without rank-one connections such that (8.12) is not rigid for exact solutions and not rigid for approximate solutions. In particular, there exists a non-affine* $u \in \mathrm{W}^{1,\infty}((0,1)^2; \mathbb{R}^2)$ *with*

$$\nabla u(x) \in K_{\mathrm{KP}} := \{A_1, \ldots, A_5\} \qquad \text{for a.e. } x \in (0,1)^2.$$

Moreover, $K_{\mathrm{KP}}^{\mathrm{qc}} \supsetneq K_{\mathrm{KP}}$.

It is remarkable that the matrices $A_1, \dots, A_5$ can be chosen to be *symmetric*. The proof is very geometric, see Section 4.3 in [160] and also [223], where the quasiconvex hull $K_{\mathrm{KP}}^{\mathrm{qc}}$ is identified explicitly.

8.5 The One-Well Inclusion

So far we have treated differential inclusions with linear and with discrete sets K. However, despite their importance, none of these situations are applicable to the problems of crystal microstructure described in Example 1.8, where we are interested in differential inclusions of the form

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^d), \\ \nabla u \in K := \mathrm{SO}(d)U_1 \cup \cdots \cup \mathrm{SO}(d)U_N \quad \text{in } \Omega \end{cases} \tag{8.19}$$

for distinct matrices $U_1, \dots, U_N \in \mathbb{R}^{d\times d}$. Every set $\mathrm{SO}(d)U$ for $U \in \mathbb{R}^{d\times d}$ is called a **well**. We will only consider the physically most relevant case where

$$\det U_1, \det U_2, \dots, \det U_N > 0.$$

In order to understand the microstructure that can form in such solids, we need to investigate the rigidity properties of the inclusion (8.19).

By the polar decomposition of matrices (see Appendix A.1) we may write $U_i = Q\tilde{U}_i$ with $Q \in \mathrm{SO}(d)$ and $\tilde{U}_i$ symmetric and positive definite. As we can absorb Q into the $\mathrm{SO}(d)$-factor in the well, we could assume in all of the following that the U_i are symmetric and positive definite to start with.

For the one-well problem, via a coordinate transformation we may in fact suppose that $U_1 = \mathrm{Id}$. We also observe that there are no rank-one connections in the well $\mathrm{SO}(d)$. Indeed, assume that there were $Q, R \in \mathrm{SO}(d)$ with $R = Q + a \otimes b$ for some $a, b \in \mathbb{R}^d \setminus \{0\}$ with $|a| = 1$ (without loss of generality). Then,

$$S := RQ^T = \mathrm{Id} + a \otimes (Qb) =: \mathrm{Id} + a \otimes \widehat{b} \in \mathrm{SO}(d).$$

Thus,

$$\mathrm{Id} + a \otimes \widehat{b} + \widehat{b} \otimes a + |\widehat{b}|^2 a \otimes a = SS^T = \mathrm{Id} = S^T S = \mathrm{Id} + a \otimes \widehat{b} + \widehat{b} \otimes a + |a|^2 \widehat{b} \otimes \widehat{b},$$

whereby $\widehat{b} = \lambda a$ for some $\lambda \neq 0$. Now take a rotation $P \in \mathrm{SO}(d)$ with $Pa = \mathrm{e}_1$. Then,

$$\det(PSP^T) = \det(\mathrm{Id} + \lambda \mathrm{e}_1 \otimes \mathrm{e}_1) = 1 + \lambda \neq 1,$$

contradicting the group property of $\mathrm{SO}(d)$. Hence, there are no rank-one connections in $\mathrm{SO}(d)$. Consequently, we may hope that the one-well inclusion is rigid. This is indeed true and in fact a classical result.

Theorem 8.20 (Reshetnyak 1967 [225]). *For $K := \mathrm{SO}(d)$ the one-well inclusion (8.19) is strongly rigid. In particular, $K^{\mathrm{qc}} = K$.*

Remark 8.21. The same result holds if $u \in \mathrm{W}^{1,\infty}$ is replaced with $u \in \mathrm{W}^{1,d}(\Omega; \mathbb{R}^d)$ in (8.19), as can be seen from the proof.

Proof. Step 1: Rigidity for exact solutions. Since $\operatorname{cof} Q = Q$ for all $Q \in \mathrm{SO}(d)$ (see Appendix A.1) we have for every $u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^d)$ with $\nabla u \in \mathrm{SO}(d)$ almost everywhere that

$$\operatorname{div} \operatorname{cof} \nabla u = \Delta u.$$

On the other hand, by the Piola identity (5.10), $\operatorname{div} \operatorname{cof} \nabla u = 0$. Thus, our u is harmonic and hence smooth (see Example 3.15 or Example 8.4). Recalling that $|\bullet|$ for us always means the Frobenius norm, $|\nabla u|^2 = \operatorname{tr}[(\nabla u)^T \nabla u] = d$ and so

$$2|\nabla^2 u|^2 = 2\nabla u \cdot \nabla \Delta u + 2|\nabla^2 u|^2 = \Delta\big[|\nabla u|^2\big] = 0.$$

Thus, ∇u is constant.

Step 2: Strong rigidity. As a preparation define

$$g(A) := |A|^d - d^{d/2} \det A, \qquad A \in \mathbb{R}^{d \times d},$$

which satisfies $g \geq 0$ and $g(A) = 0$ if and only if $A = \alpha Q$ for some $Q \in \mathrm{SO}(d)$ and $\alpha \geq 0$. To see the first claim we assume $\det A > 0$ and use the polar decomposition to write $A = QS$ with $Q \in \mathrm{SO}(d)$ and S symmetric and positive definite and then diagonalize S as $G^T S G = D = \operatorname{diag}(\lambda_1, \ldots, \lambda_d)$ with $G \in \mathbb{R}^{d \times d}$ orthogonal and $\lambda_1, \ldots, \lambda_d \geq 0$. We compute

$$\begin{aligned}
\det A &= \lambda_1 \cdots \lambda_d \\
&\leq \left(\frac{\lambda_1 + \cdots + \lambda_d}{d}\right)^d \\
&\leq d^{-d/2}\big(\lambda_1^2 + \cdots + \lambda_d^2\big)^{d/2} \\
&= d^{-d/2}|D|^d \\
&= d^{-d/2}|A|^d,
\end{aligned}$$

where we used the vector norm inequality $|\bullet|_1 \leq \sqrt{d}|\bullet|_2$ and the orthogonal invariance of the Frobenius norm. For the second claim, the above inequalities are in fact equalities, which is the case if and only if $\lambda_1 = \cdots = \lambda_d =: \alpha$, that is, $A = \alpha Q$.

Now suppose that $(u_j) \subset \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^d)$ is an approximate solution of (8.19), that is, $(u_j) \subset \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^d)$ with $u_j \overset{*}{\rightharpoonup} u$ for some map $u \in \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^d)$ and $\mathrm{dist}(\nabla u_j, \mathrm{SO}(d)) \to 0$ in measure. Then,

$$\begin{aligned}
0 &\le \int_\Omega g(\nabla u)\,\mathrm{d}x \\
&= \int_\Omega |\nabla u|^d - d^{d/2}\det\nabla u\,\mathrm{d}x \\
&\le \liminf_{j\to\infty}\int_\Omega |\nabla u_j|^d - d^{d/2}\det\nabla u_j\,\mathrm{d}x \\
&= \liminf_{j\to\infty}\int_\Omega g(\nabla u_j)\,\mathrm{d}x \\
&= 0,
\end{aligned}$$

where we used that the L^d-norm is weakly* lower semicontinuous and the determinant is weakly* continuous by Lemma 5.10. Hence, all inequalities are in fact equalities and we conclude that

$$g(\nabla u) = 0 \quad \text{a.e.} \qquad \text{and} \qquad \|\nabla u_j\|_{\mathrm{L}^d} \to \|\nabla u_j\|_{\mathrm{L}^d}.$$

This implies that $\nabla u_j \to \nabla u$ in L^d by the Radon–Riesz Theorem A.14 and consequently

$$\nabla u_j \to \nabla u \quad \text{in measure.}$$

Moreover, $\nabla u(x) = \alpha(x)Q(x)$ with $\alpha(x) \ge 0$ and $Q(x) \in \mathrm{SO}(d)$ for almost every $x \in \Omega$. On the other hand, $\lim_{j\to\infty} |\nabla u_j(x)|^2 = d$ for almost every $x \in \Omega$, whereby also $|\nabla u(x)|^2 = d$. Thus, $\alpha(x) = 1$ almost everywhere and so u is an exact solution to our differential inclusion $\nabla u(x) \in \mathrm{SO}(d)$. Then, we may conclude that ∇u is constant via the exact rigidity proved in Step 1. This establishes the strong rigidity. The fact that $K^{\mathrm{qc}} = K$ then follows from Lemmas 8.7 (i) and 8.8 (ii). □

We mention that there is also a *quantitative* one-well rigidity estimate:

Theorem 8.22 (Friesecke–James–Müller 2002 [127]). *Let* $u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^d)$*,* $p \in (1,\infty)$*. Then,*

$$\inf_{Q\in\mathrm{SO}(d)}\int_\Omega |\nabla u - Q|^p\,\mathrm{d}x \le C\int_\Omega \mathrm{dist}(\nabla u, \mathrm{SO}(d))^p\,\mathrm{d}x,$$

where $C = C(\Omega, p) > 0$ *is a constant that only depends on* Ω *and* p*.*

This result can be seen as a nonlinear analogue of the **Korn inequality**

$$\|\nabla u\|_{\mathrm{L}^2} \le C_K\big(\|u\|_{\mathrm{L}^2}^2 + \|\mathcal{E}u\|_{\mathrm{L}^2}^2\big)^{1/2}, \tag{8.20}$$

see Proposition I.1.1 in [273]. A proof of Theorem 8.22 can be found in the original work [127] for $p = 2$ and in [73] for $p \in (1, \infty)$.

8.6 Multi-well Inclusions in 2D

On the topic of multi-well problems, we first consider the situation in two dimensions. The reason that this case is easier than the corresponding problem in higher dimensions is that for $A, B \in \mathbb{R}^{2\times 2}$ we have the very useful equivalence

$$\det(A - B) \neq 0 \qquad \Longleftrightarrow \qquad \operatorname{rank}(A - B) = 2,$$

which is of course false in higher dimensions.

The two-dimensional N-well problem without rank-one connections (and positive determinants) can be completely analyzed as follows.

Theorem 8.23 **(Švérak 1993 [253, 254]).** *Let*

$$K := \mathrm{SO}(2)U_1 \cup \cdots \cup \mathrm{SO}(2)U_N$$

for $U_1, \ldots, U_N \in \mathbb{R}^{2\times 2}$ *with* $\det U_1, \det U_2, \ldots, \det U_N > 0$. *If* K *does not contain any rank-one connections, then the multi-well inclusion (8.19) is strongly rigid.*

For the proof we will need an elliptic regularity lemma.

Lemma 8.24. *Let* $K \subset \mathbb{R}^{2\times 2}$ *be compact and non-empty with the property that*

$$\det(A - B) \geq \beta |A - B|^2 \qquad \text{for all } A, B \in K \text{ and some } \beta > 0. \tag{8.21}$$

Then, for $u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^2)$ *satisfying* $\nabla u \in K$ *almost everywhere it holds that* $u \in \mathrm{W}^{2,2}_{\mathrm{loc}}(\Omega; \mathbb{R}^2)$.

Proof. Recall from Section 3.2 the definition of the k'th difference quotient ($k = 1, 2$) of u,

$$\mathrm{D}^h_k u(x) := \frac{u(x + h\mathrm{e}_k) - u(x)}{h}, \qquad h \in \mathbb{R} \setminus \{0\}.$$

Applying (8.21) to $\rho(x)(\nabla u(x + h\mathrm{e}_k) - \nabla u(x))$, where $\rho \in \mathrm{C}^\infty_c(\Omega; [0, 1])$ is a cut-off function, and integrating, we get

$$\beta \int_\Omega \rho^2 |\mathrm{D}^h_k \nabla u|^2 \,\mathrm{d}x \leq \int_\Omega \det[\rho \mathrm{D}^h_k \nabla u] \,\mathrm{d}x. \tag{8.22}$$

By the identity (8.15),

$$\begin{aligned}
&\int_\Omega \det[\rho \mathrm{D}_k^h \nabla u]\,\mathrm{d}x \\
&\quad= \int_\Omega \det\bigl[\nabla(\rho \mathrm{D}_k^h u) - \mathrm{D}_k^h u \otimes \nabla\rho\bigr]\,\mathrm{d}x \\
&\quad= \int_\Omega \det[\nabla(\rho \mathrm{D}_k^h u)] - \operatorname{cof}\,[\nabla(\rho \mathrm{D}_k^h u)] : (\mathrm{D}_k^h u \otimes \nabla\rho) + \det[\mathrm{D}_k^h u \otimes \nabla\rho]\,\mathrm{d}x \\
&\quad\le 0 + C\int_\Omega |\nabla(\rho \mathrm{D}_k^h u)| \cdot |\mathrm{D}_k^h u| \cdot |\nabla\rho| + |\mathrm{D}_k^h u|^2 \cdot |\nabla\rho|^2\,\mathrm{d}x \\
&\quad\le C\int_\Omega \rho|\mathrm{D}_k^h \nabla u| \cdot |\mathrm{D}_k^h u| \cdot |\nabla\rho|\,\mathrm{d}x + 2C\int_\Omega |\mathrm{D}_k^h u|^2 \cdot |\nabla\rho|^2\,\mathrm{d}x \\
&\quad\le \frac{\beta}{2}\int_\Omega \rho^2|\mathrm{D}_k^h \nabla u|^2\,\mathrm{d}x + \tilde{C}\int_\Omega |\mathrm{D}_k^h u|^2 \cdot |\nabla\rho|^2\,\mathrm{d}x,
\end{aligned}$$

where we also used that the determinant is a null-Lagrangian, see Lemma 5.8, and Young's inequality (with $\delta := \beta/C$). Combine this with (8.22) to get the estimate

$$\int_\Omega \rho^2|\mathrm{D}_k^h \nabla u|^2\,\mathrm{d}x \le \frac{2\tilde{C}}{\beta}\int_\Omega |\mathrm{D}_k^h u|^2 \cdot |\nabla\rho|^2\,\mathrm{d}x,$$

where the term on the right-hand side is bounded uniformly in h by Lemma 3.12 (i). By part (ii) of that lemma we then conclude that $u \in \mathrm{W}^{2,2}_{\mathrm{loc}}(\Omega;\mathbb{R}^2)$. □

Proof of Theorem 8.23. The idea is to reduce to the one-well inclusion.

Since there are no rank-one connections in K, we have for all $i, j \in \{1, \dots, N\}$, $i \neq j$, and all $Q, R \in \mathrm{SO}(2)$ that

$$\det(QU_i - RU_j) \neq 0.$$

It can be shown that this expression must be positive for all $Q, R \in \mathrm{SO}(d)$ or negative for all $Q, R \in \mathrm{SO}(d)$ since $\mathrm{SO}(2) \times \mathrm{SO}(2)$ is connected, see Problem 8.8 (i). In the following we only treat the case

$$\det(A - B) > 0 \qquad \text{for all } A, B \in K \text{ with } A \neq B; \tag{8.23}$$

the other case is analogous.

We first prove the quantitative estimate

$$\det(A - B) \ge \beta|A - B|^2 \qquad \text{for all } A, B \in K \text{ and some } \beta > 0. \tag{8.24}$$

Set $K_i := \mathrm{SO}(2)U_i$, $i = 1, \dots, N$, which are distinct, hence disjoint. If $A \in K_i$ and $B \in K_j$ for $i \neq j$, then $|A - B| > 0$. The continuous and bounded function

$$(A, B) \in K_i \times K_j \mapsto g(A, B) := \frac{\det(A - B)}{|A - B|^2}$$

is strictly positive by (8.23) on the compact and non-empty set $K_i \times K_j$, hence bounded from below. Thus, for this choice of A, B, (8.24) holds.

For $A, B \in K_i$ with a fixed i we assume $U_i = \mathrm{Id}$, that is, $A, B \in \mathrm{SO}(2)$. Observe that

$$\frac{\det(A-B)}{|A-B|^2} = \frac{1}{2} \cdot \frac{\det(\mathrm{Id} - A^T B)}{|\mathrm{Id} - A^T B|^2}.$$

It is well known that the tangent space of SO(2) at the identity matrix consists of all skew-symmetric matrices (this is really a restatement of the fact that the Lie algebra of the Lie group SO(2) is $\mathfrak{so}(2)$, the vector space of all skew-symmetric matrices). In particular, if $Q_j := A_j^T B_j \to \mathrm{Id}$ for $A_j, B_j \in \mathrm{SO}(2)$, then

$$\frac{\mathrm{Id} - Q_j}{|\mathrm{Id} - Q_j|} \to W, \qquad \text{where } W^T = -W, |W| = 1.$$

Thus,

$$\lim_{j\to\infty} \frac{\det(\mathrm{Id} - Q_j)}{|\mathrm{Id} - Q_j|^2} = \cdot \lim_{j\to\infty} \det\left(\frac{\mathrm{Id} - Q_j}{|\mathrm{Id} - Q_j|}\right) = \det W > 0.$$

Here we used that a non-zero skew-symmetric matrix $W \in \mathbb{R}^{2\times 2}$ has two non-zero eigenvalues of the form $\pm\mathrm{i}\alpha$, $\alpha \in \mathbb{R}$ and thus $\det W = \alpha^2 > 0$. This implies that the function g is continuous, strictly positive, and bounded on $\mathrm{SO}(2) \times \mathrm{SO}(2)$. Thus, (8.24) also holds in that case.

The rigidity for exact solutions $u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^2)$ with $\nabla u \in K$ almost everywhere now follows readily: By Lemma 8.24 we conclude that $\nabla u \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^{2\times 2})$. Then, $\nabla u \in K_i$ almost everywhere for one $i \in \{1, \ldots, N\}$; otherwise there would be a jump in ∇u. Thus, Theorem 8.20 becomes applicable and we conclude that ∇u is almost everywhere constant. This shows the rigidity for exact solutions.

With (8.23) already established, the rigidity for approximate solutions follows directly from Lemma 8.13. We can hence conclude strong rigidity by invoking Corollary 8.9. □

A special case of the preceding result concerns the two-well problem in two space dimensions:

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^2), \quad U_1, U_2 \in \mathbb{R}^{2\times 2}, \quad U_1 \neq U_2, \quad \det U_1, \det U_2 > 0, \\ \nabla u \in K := \mathrm{SO}(2)U_1 \cup \mathrm{SO}(2)U_2. \end{cases} \tag{8.25}$$

We can normalize U_1, U_2 to the case

$$U_1 = \mathrm{Id} = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \neq U_2 = \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}, \qquad 0 < \alpha \leq \beta, \qquad \alpha\beta \geq 1.$$

Indeed, assume that $\det U_2 \geq \det U_1 > 0$. Then, right-multiplying by U_1^{-1} (which transforms u correspondingly), we can reduce to the case when $U_1 = \mathrm{Id}$ and

$\det U_2 \geq 1$. Then, use the polar decomposition for matrices (see Appendix A.1) and a diagonalization (of the resulting symmetric and positive definite matrices) to further reduce to the case when $U_2 = \operatorname{diag}(\alpha, \beta)$ with $0 < \alpha \leq \beta$. Because $\det U_2 \geq 1$, it necessarily also follows that $\alpha\beta \geq 1$.

In this case, we can explicitly expose the rank-one connections in the set K:

Lemma 8.25. *Let* $0 < \alpha \leq \beta$ *with* $\alpha\beta \geq 1$ *and set*

$$K := \mathrm{SO}(2) \cup \mathrm{SO}(2) \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}.$$

Then:

(i) If $\alpha > 1$*, then there are no rank-one connections in* K.
(ii) If $\alpha = 1$*, then every matrix in* K *is rank-one connected to exactly one other matrix in* K.
(iii) If $\alpha < 1$*, then every matrix in* K *is rank-one connected to exactly two other matrices in* K.

Proof. Recall that for $A, B \in K$ with $A \neq B$ we have that $\operatorname{rank}(A - B) = 1$ if and only if $\det(A - B) = 0$. We showed toward the beginning of Section 8.5 that the only possible rank-one connections in K are *between* the wells. So, for $Q \in \mathrm{SO}(2)$, we are trying to find all $R \in \mathrm{SO}(2)$ such that

$$0 = \det\left(Q - R\begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}\right) = \det\left(\mathrm{Id} - Q^T R\begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}\right).$$

Via the identification $R \leftrightarrow Q^T R =: S \in \mathrm{SO}(2)$ we are thus trying to find all

$$S = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \qquad \text{with} \qquad a^2 + b^2 = 1$$

such that

$$0 = \det\left(\mathrm{Id} - S\begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}\right) = \det\begin{pmatrix} 1 - \alpha a & -\beta b \\ \alpha b & 1 - \beta a \end{pmatrix} = 1 - (\alpha + \beta)a + \alpha\beta. \tag{8.26}$$

In this case, $a = (1 + \alpha\beta)/(\alpha + \beta) \in (0, \infty)$.

Ad (i). For $\alpha > 1$ we have (using $\beta > 1$, which always follows from the above conditions) that $0 < (1 - \alpha)(1 - \beta) = 1 - \alpha - \beta + \alpha\beta$, whereby necessarily $a > 1$. This, however, contradicts the condition $a^2 + b^2 = 1$. Thus, there is no $S \in \mathrm{SO}(2)$ such that (8.26) holds and consequently, for all $Q, R \in \mathrm{SO}(2)$ we have that Q and $R\operatorname{diag}(\alpha, \beta)$ are not rank-one connected.

Ad (ii), (iii). By a similar argument as for (i), for $\alpha \leq 1$ we have $a \leq 1$. So, with $b := \pm\sqrt{1 - a^2}$ we find the solutions $S, S^T \in \mathrm{SO}(2)$ of (8.26). Thus, $R_1 := QS$, $R_2 := QS^T \in \mathrm{SO}(2)$ are rank-one connected to Q. If $\alpha = 1$, then $a = 1$, $b = 0$, $S = S^T$, and $R_1 = R_2$. Hence, (ii) follows. Finally, if $\alpha < 1$, we have $a < 1$, $S \neq S^T$, whereby also $R_1 \neq R_2$. Thus, (iii) is also established. □

It follows from this discussion that if we are interested in rigidity for (8.25), it suffices to consider the case

$$K = \mathrm{SO}(2) \cup \mathrm{SO}(2)\begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}, \qquad 1 < \alpha \leq \beta.$$

In this situation, the strong rigidity of the corresponding differential inclusion follows immediately from Theorem 8.23.

8.7 Two-Well Inclusions in 3D

The study of the two-well problem in three (and higher) dimensions is still incomplete and in particular the following is still open:

Conjecture 8.26 (*Kinderlehrer 1988, reported in [101, 182]*). If $K := \mathrm{SO}(3)U_1 \cup \mathrm{SO}(3)U_2$, where $U_1, U_2 \in \mathbb{R}^{3\times 3}$ with $\det U_1, \det U_2 > 0$, contains no rank-one connections, then the differential inclusion $\nabla u \in K$ is rigid for approximate solutions.

By similar arguments as in the last section, it can be shown that we may reduce to the case

$$K = \mathrm{SO}(3) \cup \mathrm{SO}(3)U$$

with

$$U = \begin{pmatrix} \alpha_1 & & \\ & \alpha_2 & \\ & & \alpha_3 \end{pmatrix}, \qquad \alpha_1 \geq \alpha_2 \geq \alpha_3 > 0. \tag{8.27}$$

Moreover, a calculation shows that there are no rank-one connections in K if and only if $\alpha_2 \neq 1$, see [153].

All currently known rigidity results for the three-dimensional two-well problem need some form of *strong incompatibility* between the wells. We will only survey the available results here.

Theorem 8.27 (**Matos 1992 [182]**). *Let K be as in (8.27) such that $\alpha_2 \neq 1$. If, with cyclic indices $i \in \{1, 2, 3\}$, it holds that*

$$(1 - \alpha_i)(1 - \alpha_{i-1}\alpha_{i+1}) \geq 0 \qquad \text{for at least one } i \in \{1, 2, 3\},$$

then the two-well inclusion (8.19) is strongly rigid.

Another result is the following:

Theorem 8.28 **(Dolzmann–Kirchheim–Müller–Šverák 2000 [101]).** *Let K be as in (8.27) with $\alpha_2 \neq 1$. If*

$$\alpha_1 \geq \alpha_2 > 1 > \alpha_3 \geq \frac{1}{3} \quad \textit{or} \quad 3 > \alpha_1 > 1 > \alpha_2 \geq \alpha_3 > 0,$$

then the two-well inclusion (8.19) is rigid for approximate solutions.

Finally, we remark that for the two-well problem there is a quantitative rigidity result, even for arbitrary dimensions. For this, call disjoint non-empty compact sets $K_1, K_2 \subset \mathbb{R}^{m\times d}$ **strongly incompatible** if for every gradient Young measure $\nu \in \mathbf{GY}^\infty(\Omega; \mathbb{R}^{m\times d})$ ($\Omega \subset \mathbb{R}^d$ any Lipschitz domain) it holds that if $\operatorname{supp} \nu_x \subset K_1 \cup K_2$ for almost every $x \in \Omega$, then either $\operatorname{supp} \nu_x \subset K_1$ for almost every $x \in \Omega$ or $\operatorname{supp} \nu_x \subset K_2$ for almost every $x \in \Omega$.

Theorem 8.29 **(Chaudhuri–Müller 2004 [61]).** *Let $K_1, K_2 \subset \mathbb{R}^{m\times d}$ be disjoint, non-empty, compact, and strongly incompatible sets and let $p \in [1, \infty)$. Then,*

$$\min\left(\int_\Omega \operatorname{dist}(\nabla u, K_1)^p \,\mathrm{d}x, \int_\Omega \operatorname{dist}(\nabla u, K_2)^p \,\mathrm{d}x\right) \leq C \int_\Omega \operatorname{dist}(\nabla u, K_1 \cup K_2)^p \,\mathrm{d}x,$$

where $C = C(\Omega, p) > 0$ is a constant.

Note that here, curiously, the case $p = 1$ is allowed in contrast to Theorem 8.22. A result for more than two wells can be found in [62].

8.8 Compensated Compactness

We finish this chapter with a more abstract look at rigidity. One can view the strong rigidity of a differential inclusion $\nabla u \in K$, where $K \subset \mathbb{R}^{m\times d}$ is compact, as the following "weak-to-strong" convergence principle:

$$\left.\begin{aligned} (V_j) \subset \mathrm{L}^\infty(\Omega; \mathbb{R}^{m\times d}) \\ V_j \overset{*}{\rightharpoonup} V \quad \text{in } \mathrm{L}^\infty \\ \operatorname{curl} V_j = 0 \quad \text{as distributions} \\ \operatorname{dist}(V_j, K) \to 0 \quad \text{in measure} \end{aligned}\right\} \quad \Longrightarrow \quad \begin{cases} h(V_j) \to h(V) \quad \text{in measure} \\ \text{for all bounded } h \in \mathrm{C}(\mathbb{R}^{m\times d}). \end{cases}$$

Here, $\operatorname{curl} V := (\partial_i V_j^k - \partial_j V_i^k)_{i,j,k}$ (in the sense of distributions). This formulation stresses that it is the interplay between the differential constraint $\operatorname{curl} V_j = 0$ and the nonlinear pointwise condition $\operatorname{dist}(V_j, K) \to 0$ that ensures convergence in measure, which is equivalent to the absence of oscillations. Exploring this interaction between pointwise and differential constraints systematically is the core idea of the theory of *compensated compactness*.

Even in the absence of a general weak-to-strong convergence principle as above, we have already seen situations where *certain* nonlinear expressions commute with the operation of taking weak(*) limits. Most prominently, we showed the weak continuity of minors in Lemma 5.10.

In order to set the stage for a general result, we let $\mathscr{A}$ be a homogeneous first-order linear PDE operator with constant coefficients,

$$\mathscr{A} := \sum_{l=1}^{d} A_l \partial_l, \tag{8.28}$$

where $A_l \in \mathbb{R}^{M\times N}$, $l = 1, \dots, d$ ($M, N \in \mathbb{N}$). The system of PDEs

$$\mathscr{A}V = 0, \qquad V \in \mathrm{L}^2(\Omega; \mathbb{R}^N),$$

in the following is to be interpreted in the $\mathrm{W}^{-1,2}$-sense, that is, we require

$$\langle V, \mathscr{A}^T w\rangle = \left\langle V, \sum_{l=1}^{d} A_l^T \partial_l w \right\rangle = 0 \qquad \text{for all } w \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^M).$$

We define the **wave cone** of $\mathscr{A}$ as

$$\Lambda_{\mathscr{A}} := \bigcup_{\xi\in\mathbb{S}^{d-1}} \ker \mathbb{A}(\xi), \qquad \text{where} \qquad \mathbb{A}(\xi) := (2\pi \mathrm{i}) \sum_{l=1}^{d} A_l \xi_l.$$

Here, $\mathbb{A}(\xi)$ is called the **symbol** of $\mathscr{A}$. The wave cone contains all amplitudes for which $\mathscr{A}$ is not elliptic. As such, it plays a fundamental role in the study of oscillations in sequences of functions $(V_j) \subset \mathrm{L}^2(\Omega; \mathbb{R}^N)$ with $V_j \rightharpoonup V$ in L^2 and $\mathscr{A}V_j \to 0$ in $\mathrm{W}^{-1,2}$. Let us illustrate this with a purely formal, yet instructive, argument: For the limit V we must have

$$\mathscr{A}V = 0.$$

Fourier transforming this, we get

$$\widehat{V}(\xi) \in \ker \mathbb{A}(\xi) \subset \Lambda_{\mathscr{A}} \qquad \text{for all } \xi \in \mathbb{R}^d \setminus \{0\}.$$

Thus, intuitively, all mass of $\widehat{V}_j(\xi)$ outside of $\ker \mathbb{A}(\xi)$ has to disappear in the limit and cannot contribute to oscillations in the sequence (V_j).

This fact is rigorously expressed in the following fundamental result.

Theorem 8.30 (**Tartar 1979 [267]**). *Assume that* $(V_j) \subset \mathrm{L}^2(\Omega; \mathbb{R}^N)$ *with* $V_j \rightharpoonup V$ *in* L^2 *and that* $(\mathscr{A} V_j)_j$ *is (strongly) precompact in* $\mathrm{W}^{-1,2}_{\mathrm{loc}}(\Omega; \mathbb{R}^M)$*. Suppose furthermore that* $q : \mathbb{R}^N \to \mathbb{R}$ *is a quadratic form with*

$$q(A) \geq 0 \qquad \textit{for all } A \in \Lambda_{\mathscr{A}}. \tag{8.29}$$

Then, if $q(V_j) \overset{*}{\rightharpoonup} \mu$ *in* $\mathscr{M}(\Omega)$*, it holds that*

$$q(V) \leq \mu \qquad \textit{as measures on } \Omega.$$

In particular, if $q(A) = 0$ *for all* $A \in \Lambda_{\mathscr{A}}$*, then* $q(V_j) \overset{*}{\rightharpoonup} q(V)$ *in* $\mathscr{M}(\Omega)$*.*

Here and in the following we identify $q(V_j)$ with the measure $q(V_j)\,\mathscr{L}^d \llcorner \Omega$. We also note that condition (8.29) is equivalent to the convexity of q in the directions in $\Lambda_{\mathscr{A}}$.

Proof Step 1. We first show that we may assume that $V \equiv 0$. Indeed, set

$$\tilde{V}_j := V_j - V.$$

Then, $\tilde{V}_j \rightharpoonup 0$ in L^2 and $(\mathscr{A}\tilde{V}_j)_j$ is (strongly) precompact in $\mathrm{W}^{-1,2}_{\mathrm{loc}}(\Omega; \mathbb{R}^m)$. If $q(A) = A^T \mathbf{S} A$ ($A \in \mathbb{R}^N$) with a symmetric matrix $\mathbf{S} \in \mathbb{R}^{N\times N}_{\mathrm{sym}}$, we compute

$$q(\tilde{V}_j) = q(V_j) - 2V_j^T \mathbf{S} V + q(V) \overset{*}{\rightharpoonup} \mu - q(V) =: \tilde{\mu} \quad \text{in } \mathscr{M}(\Omega).$$

Thus, it remains to show that $\tilde{\mu} \geq 0$. In the following we drop the tildes.

Step 2. Let $\phi \in \mathrm{C}^\infty_c(\Omega)$ and set $W_j := \phi V_j$. Then,

$$\mathscr{A} W_j = \phi \mathscr{A} V_j + \sum_{l=1}^{d} A_l V_j \partial_l \phi$$

and this sequence is (strongly) precompact in the space $\mathrm{W}^{-1,2}(\Omega; \mathbb{R}^M)$ since the second term is uniformly bounded in $\mathrm{L}^2(\Omega; \mathbb{R}^M)$, which is compactly embedded in $\mathrm{W}^{-1,2}(\Omega; \mathbb{R}^M)$. Thus, we may assume that

$$\mathscr{A} W_j \to 0 \quad \text{in } \mathrm{W}^{-1,2}. \tag{8.30}$$

Here, the convergence to zero follows since $\langle \mathscr{A} W_j, w\rangle = -\langle W_j, \mathscr{A}^T w\rangle \to 0$ for all $w \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^M)$ and $W_j \rightharpoonup 0$ in L^2. Moreover,

$$q(W_j) = \phi^2 q(V_j) \overset{*}{\rightharpoonup} \phi^2 \mu \quad \text{in } \mathscr{M}(\Omega).$$

Step 3. In the following we will show that

$$\int_\Omega \phi^2 \,\mathrm{d}\mu = \lim_{j\to\infty} \int_\Omega q(W_j)\,\mathrm{d}x \geq 0 \qquad \text{for all } \phi \in \mathrm{C}^\infty_c(\Omega), \tag{8.31}$$

which will prove the claim. We assume that q is extended to $\mathbb{C}^N$ as a Hermitian form, that is, for $q(A) = A^T \mathbf{S} A$ with $\mathbf{S} \in \mathbb{R}^{N\times N}$ set

$$q(Z) := Z^* \mathbf{S} Z \qquad Z \in \mathbb{C}^N.$$

If $Z = A + \mathrm{i}B \in \Lambda_{\mathscr{A}} + \mathrm{i}\Lambda_{\mathscr{A}}$, we have

$$\operatorname{Re} q(Z) = \operatorname{Re}\big[q(A) + q(B) + \mathrm{i}(A^* \mathbf{S} B - B^* \mathbf{S} A)\big] = q(A) + q(B) \geq 0. \quad (8.32)$$

By the (vector-valued) Parseval relation (A.5),

$$\int_\Omega q(W_j)\,\mathrm{d}x = \int_\Omega W_j^* \mathbf{S} W_j\,\mathrm{d}x = \int \widehat{W}_j^* \mathbf{S} \widehat{W}_j\,\mathrm{d}\xi = \int q(\widehat{W}_j)\,\mathrm{d}\xi$$
$$= \int \operatorname{Re} q(\widehat{W}_j)\,\mathrm{d}\xi,$$

where the last equality follows from the identity $\overline{\widehat{W}_j(\xi)} = \widehat{W}_j(-\xi)$ since W is real-valued. Hence, to prove our claim (8.31), we will in the following establish

$$\lim_{j\to\infty} \int \operatorname{Re} q(\widehat{W}_j)\,\mathrm{d}\xi \geq 0. \quad (8.33)$$

In order to see this, we split the integral's domain into $B(0,1)$ and $\mathbb{R}^d \setminus B(0,1)$ and prove (8.33) separately for these two parts.

On $B(0,1)$ this claim is straightforward: Since $W_j \rightharpoonup 0$ in L^2 and the W_j's have uniformly bounded supports, it holds that

$$\widehat{W}_j(\xi) = \int W_j(x) \mathrm{e}^{-2\pi \mathrm{i} x\cdot\xi}\,\mathrm{d}x \to 0 \qquad \text{for all } \xi \in \mathbb{R}^d.$$

Moreover, $|\widehat{W}_j(\xi)| \leq \|W_j\|_{\mathrm{L}^1} \leq C$ for all $\xi \in \mathbb{R}^d$ and a uniform constant $C > 0$. Thus,

$$\widehat{W}_j \to 0 \quad \text{in } \mathrm{L}^2_{\mathrm{loc}}.$$

Consequently,

$$\lim_{j\to\infty} \int_{B(0,1)} \operatorname{Re} q(\widehat{W}_j)\,\mathrm{d}\xi = 0. \quad (8.34)$$

Step 4. We next show that for all $\delta > 0$ there exists a constant $C_\delta > 0$ such that

$$\operatorname{Re} q(Z) \geq -\delta |Z|^2 - C_\delta |\mathbb{A}(\eta) Z|^2 \quad (8.35)$$

for all $Z \in \mathbb{C}^N$ and all $\eta \in \mathbb{S}^{d-1}$. Indeed, if this was not the case there would exist $\delta > 0$ and sequences $(Z_m) \subset \mathbb{C}^N$ with $|Z_m| = 1$ and $(\eta_m) \subset \mathbb{S}^{d-1}$ with the property

$$\operatorname{Re} q(Z_m) < -\delta |Z_m|^2 - m |\mathbb{A}(\eta_m) Z_m|^2.$$

Without loss of generality we may assume that $\eta_m \to \eta$ and $Z_m \to Z$ with $|Z| = 1$. Then, there is a constant $C > 0$ such that

$$|\mathbb{A}(\eta_m)Z_m|^2 \leq \frac{C}{m},$$

and thus $Z \in \ker_{\mathbb{C}} \mathbb{A}(\eta) \subset \Lambda_{\mathscr{A}} + \mathrm{i}\Lambda_{\mathscr{A}}$. By (8.32), $\operatorname{Re} q(Z) \geq 0$. On the other hand,

$$\operatorname{Re} q(Z) = \lim_{m\to\infty} \operatorname{Re} q(Z_m) \leq -\delta,$$

a contradiction. Hence, (8.35) must hold.

Step 5. From (8.30) we get that $(\mathrm{id} - \Delta)^{-1/2}[\mathscr{A} W_j] \to 0$ in L^2, whereby

$$\frac{1}{(1 + 4\pi^2|\xi|^2)^{1/2}} \mathbb{A}(\xi)\widehat{W}_j(\xi) \to 0 \quad \text{in } \mathrm{L}^2.$$

Thus, since $(1 + 4\pi^2|\xi|^2)^{1/2} \sim |\xi|$ for $|\xi| \geq 1$, we have also

$$\mathbb{A}\left(\frac{\xi}{|\xi|}\right)\widehat{W}_j(\xi) = \frac{1}{|\xi|}\mathbb{A}(\xi)\widehat{W}_j(\xi) \to 0 \quad \text{in } \mathrm{L}^2(\mathbb{R}^d \setminus B(0, 1)). \tag{8.36}$$

Applying (8.35) for $Z := \widehat{W}_j(\xi)$ and $\eta := \xi/|\xi|$, we get

$$\begin{aligned}&\int_{\mathbb{R}^d\setminus B(0,1)} \operatorname{Re} q(\widehat{W}_j(\xi)) \,\mathrm{d}\xi \\ &\qquad \geq -\int_{\mathbb{R}^d\setminus B(0,1)} \delta|\widehat{W}_j(\xi)|^2 \,\mathrm{d}\xi - C_\delta \int_{\mathbb{R}^d\setminus B(0,1)} \left|\mathbb{A}\left(\frac{\xi}{|\xi|}\right)\widehat{W}_j(\xi)\right|^2 \,\mathrm{d}\xi.\end{aligned}$$

Combine this with (8.36) to deduce that

$$\lim_{j\to\infty} \int_{\mathbb{R}^d\setminus B(0,1)} \operatorname{Re} q(\widehat{W}_j(\xi)) \,\mathrm{d}\xi \geq -C\delta$$

for a constant $C > 0$, where we also used that the $\widehat{W}_j$ are uniformly norm-bounded in $\mathrm{L}^2(\mathbb{R}^d)$. Thus, as $\delta > 0$ was arbitrary,

$$\lim_{j\to\infty} \int_{\mathbb{R}^d\setminus B(0,1)} \operatorname{Re} q(\widehat{W}_j(\xi)) \,\mathrm{d}\xi \geq 0. \tag{8.37}$$

From (8.34) and (8.37) we conclude (8.33), finishing the proof in the case $q \geq 0$ on $\Lambda_{\mathscr{A}}$.

If $q \equiv 0$ on $\Lambda_{\mathscr{A}}$, simply apply the first part of the theorem to q and to $-q$. □

Corollary 8.31. *Assume that* $(u_j) \subset \mathrm{W}^{1,2}(\Omega;\mathbb{R}^m)$ *with* $u_j \rightharpoonup u \in \mathrm{W}^{1,2}(\Omega;\mathbb{R}^m)$. *Suppose furthermore that* $q\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ *is a quadratic form with*

$$q(A) \geq 0 \qquad \textit{for all } A \in \mathbb{R}^{m\times d} \textit{ with } \operatorname{rank} A \leq 1.$$

Then,

$$\int_\Omega \phi q(\nabla u)\,\mathrm{d}x \leq \liminf_{j\to\infty} \int_\Omega \phi q(\nabla u_j)\,\mathrm{d}x$$

for all $\phi \in \mathrm{C}_0(\Omega)$. *Moreover, if* $\nabla u \equiv F \in \mathbb{R}^{m\times d}$ *is constant and* $u_j|_{\partial\Omega} = Fx$, *then*

$$q(F) \leq \liminf_{j\to\infty} \fint_\Omega q(\nabla u_j)\,\mathrm{d}x. \tag{8.38}$$

In particular, q *is quasiconvex.*

Proof. This is just a reformulation of Tartar's Theorem 8.30 for (with an obvious abuse of notation)

$$\mathscr{A} V := \operatorname{curl} V := \bigl(\partial_j V_i^k - \partial_i V_j^k\bigr)_{i,j=1,\ldots,d}^{k=1,\ldots,m}.$$

To explain our particular choice of "curl", we note that for any $u \in \mathrm{C}^2(\mathbb{R}^d;\mathbb{R}^m)$ it holds that

$$\partial_i [\nabla u]_j^k = \partial_i \partial_j u^k = \partial_j \partial_i u^k = \partial_j [\nabla u]_i^k \qquad \text{for all } i,j = 1,\ldots,d;\, k = 1,\ldots,m.$$

It is a classical result that these *integrability conditions* in fact characterize gradients on simply connected domains.

We can now easily compute that $A \in \ker \mathbb{A}(\xi)$ for $\xi \in \mathbb{S}^{d-1}$ if and only if

$$\xi_j A_i^k = \xi_i A_j^k \qquad \text{for all } i,j = 1,\ldots,d;\, k = 1,\ldots,m.$$

Thus, $A_i^k = \xi_i A_j^k / \xi_j$ for all j such that $\xi_j \neq 0$. This is only possible if $A = a \otimes \xi$ for some $a \in \mathbb{R}^m$. Hence,

$$\Lambda_{\mathrm{curl}} = \bigcup_{\xi \in \mathbb{S}^{d-1}} \ker \mathbb{A}(\xi) = \bigl\{\, a \otimes \xi \in \mathbb{R}^{m\times d} \;:\; a \in \mathbb{R}^m, \xi \in \mathbb{S}^{d-1} \,\bigr\}.$$

The additional assertion (8.38) from the statement of the corollary follows by extending all u_j to a larger domain $\Omega' \Supset \Omega$ and applying Tartar's theorem there. The quasiconvexity of q is then a consequence of a construction like the one in the proof of Proposition 5.18. □

As the most well-known compensated compactness result we have the following *div-curl lemma*:

Lemma 8.32 **(Murat–Tartar 1974 [209]).** *Assume that the sequences* $(u_j), (v_j) \subset \mathrm{L}^2(\Omega; \mathbb{R}^d)$ *are such that* $u_j \rightharpoonup u$, $v_j \rightharpoonup v$ *in* L^2 *and that*

$$(\operatorname{div} u_j)_j, (\operatorname{curl} v_j)_j \text{ are precompact in } \mathrm{W}^{-1,2}_{\mathrm{loc}}.$$

Then,

$$u_j \cdot v_j \rightharpoonup u \cdot v \quad \text{in } \mathrm{L}^1_{\mathrm{loc}}.$$

Proof. We set $V_j := (u_j, v_j)$ and $\mathscr{A} := (\operatorname{div}, \operatorname{curl})$. Then, we may compute (see Problem 8.9)

$$\Lambda_{\mathscr{A}} = \bigl\{ (a, b) \in \mathbb{R}^d \times \mathbb{R}^d \; \mathbf{:} \; a \perp b \bigr\}.$$

Thus, for the quadratic function $q(a, b) := a \cdot b$ we have that q vanishes on $\Lambda_{\mathscr{A}}$ and the conclusion thus follows from Tartar's theorem. □

Let us finally remark that one cannot generalize Tartar's Theorem 8.30 to non-quadratic functions h that are convex in the directions of $\Lambda_{\mathscr{A}}$.

Example 8.33. Take the rank-one convex but not quasiconvex function h from Šverák's Example 7.10 (which has 4-growth) and set $\mathscr{A} := \operatorname{curl}$ (defined as in the proof of Corollary 8.31 above). By Proposition 5.18, the functional

$$\mathscr{F}[u] := \int_{(0,1)^2} h(\nabla u(x)) \, \mathrm{d}x, \qquad u \in \mathrm{W}^{1,4}((0, 1)^2; \mathbb{R}^3),$$

is not weakly lower semicontinuous. Thus, there exists a sequence of maps $(u_j) \subset \mathrm{W}^{1,4}((0, 1)^2; \mathbb{R}^3)$ with $u_j \rightharpoonup u$ in $\mathrm{W}^{1,4}$ and such that

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

In particular, for this h (an L^4-version of) Tartar's Theorem 8.30 fails.

This can be explained as follows: In the case $\mathscr{A} = \operatorname{curl}$, the positivity of the quadratic form q on the wave cone $\Lambda_{\mathrm{curl}} = \{ a \otimes b \; \mathbf{:} \; a \in \mathbb{R}^m, \; b \in \mathbb{R}^d \}$ (see the proof of Corollary 8.31) implies the rank-one convexity of the quadratic form, as can be easily verified. This rank-one convexity, however, is equivalent to the quasiconvexity of the quadratic form by Problem 5.7. Thus, Tartar's theorem in this case just says that (a version of) lower semicontinuity holds, which is not surprising with all the theory that we have available now. As we have seen before, Šverák's (non-quadratic) example precisely distinguishes between these two notions of convexity and thus between the validity or non-validity of weak lower semicontinuity.

Notes and Historical Remarks

The origin of the rigidity theory as presented in this chapter lies in the Murat–Tartar Div-Curl Lemma 8.32, first published in [209] (but established four years before in 1974) and the Ball–James Rigidity Theorem 8.1 from [30]. The latter can further be traced back to **Hadamard's jump condition**: For a matrix-valued function $V : \mathbb{R}^d \to \mathbb{R}^{m \times d}$ of the form

$$V(x) = \begin{cases} A & \text{if } x \cdot n \leq 0, \\ B & \text{if } x \cdot n > 0, \end{cases}$$

where $A, B \in \mathbb{R}^{m \times d}$ and $n \in \mathbb{S}^{d-1}$, to be the gradient of a function $u : \mathbb{R}^d \to \mathbb{R}^m$, it is necessary and sufficient that

$$A - B = a \otimes n \qquad \text{for some } a \in \mathbb{R}^m.$$

Many related rigidity and compensated compactness theorems have been proved since, of which we could only present a selection.

The term "rigidity" itself is unfortunately used in different ways by different authors. In fact, *any* kind of restriction on the shape of a map is sometimes called "rigidity". Here, however, we reserve this term for the conclusion that a map is affine. Our definitions of rigidity for exact and for approximate solutions follow most closely those in Kirchheim's influential lecture notes [160]. In particular, we require linear boundary values along approximate solutions, but no boundary condition for exact solutions. This is explained as follows: Many differential inclusions with a discrete set K are trivially rigid when we impose affine boundary conditions, even if there are rank-one connections in K. On the other hand, rigidity for approximate solutions is most interesting when imposing linear (or affine) boundary conditions, see the discussion in Section 8.3.

Theorem 8.3 is originally contained in a more general result of Tartar [268]. It is called the "span restriction" in [43]. A special case of Theorem 8.5 (i) for $d = m = 2$ was shown in Lemma 1.4 of [90], also see the proof of Theorem 3.95 in [15]. The "blow-up technique" mentioned in connection with the polar inclusion (8.10) is systematically explained in great detail in [123, 124]. Theorem 8.11 was first established in full in [249], which was never published. The rigidity for exact solutions was known before, namely through more general results in [154] and an unpublished manuscript by Zhang. The presented proof of rigidity for exact solutions, however, is due to Kirchheim and reproduced in [203]. Our proof of approximate rigidity follows an idea from [251] and also uses Lemma 8.13, which is from [253]. Another proof is in [17] based on the theory of quasiregular mappings (which also mentions an unpublished similar argument by Ball and James).

Constructions similar to the fundamental T_4-configuration were probably first employed by Scheffer [239], but its importance in the present context was only realized after Tartar's work, see [271], which refers to his work of 1983. Similar examples to Tartar's were found in [21, 43, 215] (in particular, [43] adapted Tartar's original example to the present version with diagonal matrices).

The Young measure approach of Section 8.3 and the general compensated compactness philosophy discussed at the beginning of Section 8.8 is again mostly due to Tartar [267, 268, 270, 271]. This theory has also proved to be very fruitful in the study of hyperbolic conservation laws, see Chapter XVI in [81] for an overview and many references to the vast literature. Some recent investigations into various notions of "incompatibility" between several sets $K_1, \ldots, K_n$, which generalizes our notions of rigidity, can be found in [32] and the references cited therein.

The first part of Theorem 8.20 is (a version of) the classical Liouville theorem (also see Problem 8.2 (ii)); the extension to Sobolev functions as well as part (ii) is the work of Reshetnyak [225]. Our proof is due to Kinderlehrer [156].

The rigidity for the two-well problem in two dimensions and its extension to the N-well problem are from [253, 254]. More general N-well problems for $N \geq 3$ in three dimensions were investigated by Kirchheim [159, 160].

A direct proof of the div-curl lemma using elliptic regularity theory can be found in Theorem 16.2.1 of [81]. In the context of such compensated compactness problems, extensions of Young measure theory that allow one to pass to the limit in quadratic expressions have been developed by Tartar [269] under the name "H-measures" and, independently, by Gérard [130], who called them "micro-local defect measures", cf. the survey articles [126, 272].

Problems

8.1. Let $A, B \in \mathbb{R}^{m\times d}$ with $A - B = a \otimes n$ ($a \in \mathbb{R}^m \setminus \{0\}$, $n \in \mathbb{S}^{d-1}$) and let $\theta \in (0, 1)$. Set $F := \theta A + (1-\theta)B$ and prove that there exists a uniformly $\mathrm{W}^{1,\infty}$-bounded sequence $(u_j) \subset \mathrm{W}^{1,\infty}_{Fx}(\Omega; \mathbb{R}^m)$ that approximately solves the inclusion $\nabla u \in \{A, B\}$ and $\nabla u_j \overset{*}{\rightharpoonup} F$ in L^∞ but not $\nabla u_j \to F$ in measure.

8.2. Consider a complex function $f : \mathbb{C} \to \mathbb{C}$ as a mapping $f = (u, v)$ from $\mathbb{R}^2$ to itself (identify $z = x + \mathrm{i}y \in \mathbb{C}$ with $(x, y) \in \mathbb{R}^2$).

(i) Find a subspace $L \subset \mathbb{R}^{2\times 2}$ such that $\nabla f(x, y) \in L$ for all $(x, y) \in \mathbb{R}^2$ is equivalent to the (weak) *Cauchy–Riemann equations*:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

Thus prove that all (weakly) holomorphic functions are smooth.

(ii) Prove the following well-known theorem from complex analysis: Let $f : \mathbb{C} \to \mathbb{C}$ be holomorphic with $|f| = \text{const}$. Then, f is constant.

(iii) Prove *Montel's theorem*: A sequence of uniformly bounded holomorphic functions $f_j : \mathbb{D} \to \mathbb{C}$ on the unit disc $\mathbb{D} \subset \mathbb{C}$ (or any other bounded open domain) that converges to $f : \mathbb{D} \to \mathbb{C}$ pointwise or in measure converges in fact locally uniformly and the limit f is itself holomorphic.

8.3. Prove Theorem 8.5. *Hint:* Inspect the proof of the Ball–James Rigidity Theorem 5.13 and also use Theorem 8.3 or prove the rigidity for exact solutions for rank $P_0 \geq 2$ in the following alternative way: For $u \in \mathrm{C}^1(\Omega;\mathbb{R}^m)$ satisfying $\nabla u(x) = P_0 g(x)$ with $g \in \mathrm{C}(\Omega)$ first show the projection relation $\nabla g(x) = (\xi \cdot \nabla g(x)^T)\xi^T$ and conclude from there.

8.4. Even before Theorem 8.16 was established, it was already known that the inclusion $\nabla u \in K_{\mathrm{T4}}$ with K_{T4} from Proposition 8.17 is rigid for exact inclusions. Prove this in an elementary way. *Hint:* Inspect the proof of rigidity for exact solutions in Theorem 8.11 for inspiration.

8.5. Prove that the sequence $(u_j) \subset \mathrm{W}^{1,\infty}((0,1)^2;\mathbb{R}^m)$ that was constructed in Proposition 8.17 generates the homogeneous Young measure

$$\nu = \frac{8}{15}\delta_{A_1} + \frac{4}{15}\delta_{A_2} + \frac{2}{15}\delta_{A_3} + \frac{1}{15}\delta_{A_4}.$$

8.6. Find a non-empty compact set $K \subset \mathbb{R}^{m\times d}$ such that $\nabla u \in K$ is not rigid for exact solutions, but rigid for approximate solutions.

8.7. Show that if the set $K \subset \mathbb{R}^{m\times d}$ is compact, then K^{qc} is also compact.

8.8. Let $K \subset \mathbb{R}^{2\times 2}$ be a compact, non-empty and *connected* set without rank-one connections.

(i) Show that it either holds that $\det(A - B) > 0$ for all $A, B \in K$ with $A \neq B$ or $\det(A - B) < 0$ for all $A, B \in K$ with $A \neq B$. Conclude that K is rigid for approximate solutions and $K^{\mathrm{qc}} = K$. This proves Tartar's conjecture for sets K as above and was established by Šverák in 1993. *Hint:* Adapt the arguments from Section 8.6.

(ii) Show that if K is a closed connected smooth manifold and *elliptic* in the sense that for all $A \in K$ the tangent space of K at A does not contain rank-one directions, then either $\det(A-B) > c|A-B|^2$ for all $A, B \in K$ or $\det(A-B) < -c|A-B|^2$ for all $A, B \in K$. Conclude that every $u \in \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^2)$ with $\nabla u \in K$ is in $\mathrm{W}^{2,2}_{\mathrm{loc}}(\Omega;\mathbb{R}^2)$.

8.9. Let "$\mathscr{A} := (\mathrm{div}, \mathrm{curl})$", which needs to be defined properly in the sense of (8.28). Then compute that

$$\ker \Lambda_{\mathscr{A}} = \bigl\{ (a,b) \in \mathbb{R}^d \times \mathbb{R}^d \;:\; a \perp b \bigr\}.$$

8.10. Assume that $(u_j), (v_j) \subset \mathrm{L}^\infty(\Omega)$ with $u_j \overset{*}{\rightharpoonup} u$ and $v_j \overset{*}{\rightharpoonup} v$ in L^∞ and such that $(\partial_1 u_j), (\partial_2 v_j)$ exist in the weak sense and are uniformly bounded in L^∞. Show that then

$$u_j \cdot v_j \overset{*}{\rightharpoonup} u \cdot v \quad \text{in } \mathrm{L}^\infty.$$

Hint: Use Tartar's Theorem 8.30.

Chapter 9
Microstructure

Motivated by the example on crystal microstructure in Section 1.8 and the remarks in Section 8.3 about the connection of the quasiconvex hull to the relaxation of integral functionals, in this chapter we continue our analysis of the differential inclusion

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m), \quad u|_{\partial\Omega} = Fx, \\ \nabla u \in K \quad \text{in } \Omega. \end{cases} \tag{9.1}$$

Here, $K \subset \mathbb{R}^{m \times d}$ $(d, m \geq 2)$ is assumed to be compact and non-empty, and

$$F \in K^{\mathrm{qc}}.$$

Unlike in the previous chapter, however, now we consider the complementary case where $K^{\mathrm{qc}} \supsetneq K$ and there is no approximate rigidity (see Lemma 8.8 (i)).

In the first part of this chapter we concern ourselves with approximate solutions to (9.1) and various hulls of K. As we have seen before, the structure of the quasiconvex hull K^{qc} of K is intimately related to the rank-one connections in K. However, we already encountered sets without rank-one connections that are *not* rigid for approximate solutions, most strikingly the T_4-configuration K_{T4}, for which

$$K_{\mathrm{T4}}^{\mathrm{qc}} \supsetneq K_{\mathrm{T4}}$$

since at least the intermediate matrices B_1, B_2, B_3, B_4 have to lie in $K_{\mathrm{T4}}^{\mathrm{qc}}$. To compute the quasiconvex hull of K_{T4} and of other sets we will introduce lower and upper bounds on K^{qc} in the form of the lamination-convex hull, the rank-one convex hull and the polyconvex hull of K. If the upper and lower bounds agree, then they must be equal to the quasiconvex hull as well.

It is a surprising fact, which we will study in the second part of this chapter, that in many interesting applications (9.1) can also be solved *exactly*, but that non-trivial

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8_9

exact solutions necessarily display highly complex oscillations. The technique to construct these exact solutions is nowadays customarily called *convex integration* and is based on iterative addition of high-frequency, low-amplitude oscillations. Starting from ideas first employed by John F. Nash and Mikhail Gromov, this tool has now led to many astonishing results. Here, we only discuss the applications to microstructure, but similar techniques also have profound implications in geometry and fluid dynamics. There are several approaches to convex integration and we present two of them: the classical convex integration scheme based on in-approximations as well as the abstract and elegant Baire category method.

9.1 Laminates and Hulls of Sets

In this and the next section we will consider several ways to compute or at least estimate quasiconvex hulls.

For a non-empty compact set $K \subset \mathbb{R}^{m\times d}$ we define the set $\mathscr{M}^{\mathrm{lc}}(K)$ of **laminates of finite order** supported on K as

$$\mathscr{M}^{\mathrm{lc}}(K) := \bigcup_{i=0}^{\infty} \mathscr{M}^{\mathrm{lc},i}(K),$$

where

$$\begin{aligned}\mathscr{M}^{\mathrm{lc},0}(K) &:= \bigl\{\, \delta_A \;:\; A \in K \,\bigr\},\\ \mathscr{M}^{\mathrm{lc},i+1}(K) &:= \bigl\{\, \theta\mu_1 + (1-\theta)\mu_2 \;:\; \mu_1, \mu_2 \in \mathscr{M}^{\mathrm{lc},i}(K),\ \operatorname{rank}([\mu_1]-[\mu_2]) \le 1,\\ &\qquad\qquad \theta \in [0,1] \,\bigr\}.\end{aligned}$$

We call an element of $\mathscr{M}^{\mathrm{lc},i}(K) \setminus \bigcup_{j=0}^{i-1} \mathscr{M}^{\mathrm{lc},j}(K)$ a **laminate of order** i. Lemma 9.3 below will show that all elements of $\mathscr{M}^{\mathrm{lc}}(K)$ are homogeneous gradient Young measures. Problem 9.2 contains an alternative characterization of $\mathscr{M}^{\mathrm{lc}}(K)$. The **lamination-convex hull** of K is defined to be

$$K^{\mathrm{lc}} := \bigl\{\, [\mu] \;:\; \mu \in \mathscr{M}^{\mathrm{lc}}(K) \,\bigr\}.$$

Thus, K^{lc} is obtained from K by inductively adding rank-one lines. Indeed, we could also write

$$K^{\mathrm{lc}} = \bigcup_{i=0}^{\infty} K^{\mathrm{lc},i},$$

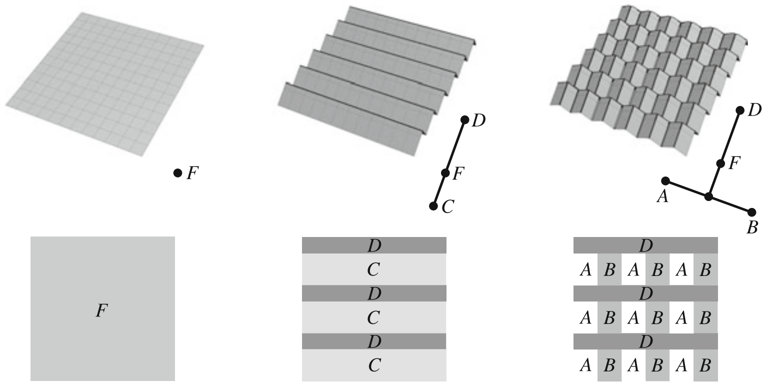


Fig. 9.1 Laminates of order 0, 1, and 2 (only one element of the generating sequence and its rank-one diagram as well as the gradient schematic are shown)

where

$$
\begin{aligned}
K^{\mathrm{lc},0} &:= K, \\
K^{\mathrm{lc},i+1} &:= \big\{ \theta A + (1-\theta)B \;:\; A, B \in K^{\mathrm{lc},i},\; \operatorname{rank}(A-B) \le 1,\; \theta \in [0,1] \big\}.
\end{aligned}
$$

We illustrate laminates of different orders in Figure 9.1.

Example 9.1. Set for some $a \in \mathbb{R}^m$, $b \in \mathbb{R}^d$,

$$K := \{0, a \otimes b\} \subset \mathbb{R}^{m \times d}.$$

Then,

$$K^{\mathrm{lc}} = K^{\mathrm{qc}} = K^{**} = \big\{ \theta a \otimes b \;:\; \theta \in [0,1] \big\},$$

as can be easily verified.

We will see in Lemma 9.3 below that the lamination-convex hull K^{lc} is a lower bound on the quasiconvex hull K^{qc} in the sense that $K^{\mathrm{lc}} \subset K^{\mathrm{qc}}$. Unfortunately, K^{lc} is often strictly smaller than K^{qc}, limiting its usefulness. For instance, since there are no rank-one connections in the T_4-configuration K_{T4}, we get that $K_{\mathrm{T4}}^{\mathrm{lc}} = K_{\mathrm{T4}}$, so we have not come nearer to bounding $K_{\mathrm{T4}}^{\mathrm{qc}}$ from below. The reason for this is that the procedure employed in the proof of Proposition 8.17 to construct non-trivial elements of $\mathscr{M}^{\mathrm{qc}}(K_{\mathrm{T4}})$ and thus of points in $K_{\mathrm{T4}}^{\mathrm{qc}} \setminus K_{\mathrm{T4}}$, needed *infinitely* many lamination steps. Consequently, we define the set of **laminates of infinite order** (also called **rank-one convex measures**) supported on a non-empty compact set $K \subset \mathbb{R}^{m \times d}$ as

$$\mathscr{M}^{\mathrm{rc}}(K) := \big\{ \mu \in \mathscr{M}^1(K) \;:\; \mu_j \overset{*}{\rightharpoonup} \mu \text{ for some } (\mu_j) \subset \mathscr{M}^{\mathrm{lc}}(K) \big\}.$$

The corresponding **rank-one-convex hull** of K is

$$K^{\mathrm{rc}} := \big\{ [\mu] \;:\; \mu \in \mathscr{M}^{\mathrm{rc}}(K) \big\}.$$

Theorem 9.2 (Pedregal 1993 [221]). *Let $\mu \in \mathscr{M}^1(K)$. Then, $\mu \in \mathscr{M}^{\mathrm{rc}}(K)$ if and only if*

$$h([\mu]) \le \int h \,\mathrm{d}\mu$$

for all rank-one convex functions $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$.

A proof of this fact can be found in [221]. Thus,

$$\mathscr{M}^{\mathrm{rc}}(K) = \biggl\{ \mu \in \mathscr{M}^1(K) \;:\; h([\mu]) \le \int h \,\mathrm{d}\mu \text{ for all rank-one convex functions } h \in \mathrm{C}(\mathbb{R}^{m\times d}) \biggr\}.$$

We also define the set of **polyconvex measures** supported on K as

$$\mathscr{M}^{\mathrm{pc}}(K) := \bigl\{ \mu \in \mathscr{M}^1(K) \;:\; M([\mu]) = \langle M, \mu\rangle \text{ for all minors } M \bigr\},$$

where the minors are defined as in Section 5.2. Recall that for $m = d = 3$, we can equivalently choose $M \in \{\mathrm{id}, \mathrm{cof}, \det\}$. Then, the **polyconvex hull** of a non-empty compact set $K \subset \mathbb{R}^{m\times d}$ is defined as

$$K^{\mathrm{pc}} := \bigl\{ [\mu] \;:\; \mu \in \mathscr{M}^{\mathrm{pc}}(K) \bigr\}.$$

Since the minors are precisely the quasiaffine functions, see Corollary 5.9, we get immediately that

$$\mathscr{M}^{\mathrm{pc}}(K) = \biggl\{ \mu \in \mathscr{M}^1(K) \;:\; h([\mu]) \le \int h \,\mathrm{d}\mu \text{ for all polyconvex } h \in \mathrm{C}(\mathbb{R}^{m\times d}) \biggr\}.$$

There are further equivalent ways to define the various hulls, see Problem 9.3 and Problem 9.4.

Lemma 9.3. *For a compact set $K \subset \mathbb{R}^{m\times d}$ it holds that*

$$\mathscr{M}^{\mathrm{lc}}(K) \subset \mathscr{M}^{\mathrm{rc}}(K) \subset \mathscr{M}^{\mathrm{qc}}(K) \subset \mathscr{M}^{\mathrm{pc}}(K)$$

and thus

$$K \subset K^{\mathrm{lc}} \subset K^{\mathrm{rc}} \subset K^{\mathrm{qc}} \subset K^{\mathrm{pc}} \subset K^{**}.$$

Proof. The inclusion $\mathscr{M}^{\mathrm{lc}}(K) \subset \mathscr{M}^{\mathrm{rc}}(K)$ is trivial. Moreover, it follows from Corollary 5.12 in conjunction with Corollary 5.9 that for any homogeneous gradient Young measure $\mu \in \mathscr{M}^{\mathrm{qc}}(K)$ and any minor M we have $M([\mu]) = \langle M, \mu\rangle$, whereby $\mu \in \mathscr{M}^{\mathrm{pc}}(K)$.

It remains to show $\mathscr{M}^{\mathrm{lc}}(K) \subset \mathscr{M}^{\mathrm{qc}}(K)$, from which it follows that also $\mathscr{M}^{\mathrm{rc}}(K) \subset \mathscr{M}^{\mathrm{qc}}(K)$ by the weak*-closedness of $\mathscr{M}^{\mathrm{qc}}(K)$, which can be seen by a diagonal argument similar to the one in the proof of Lemma 7.17. We proceed inductively. For $\mu \in \mathscr{M}^{\mathrm{lc},0}(K)$ there is nothing to show. For the inductive step assume that $\mu \in \mathscr{M}^{\mathrm{lc},i+1}(K)$ is of the form

$$\mu = \theta\mu_1 + (1-\theta)\mu_2, \qquad \theta \in [0,1],$$

where $\mu_1, \mu_2 \in \mathscr{M}^{\mathrm{lc},i}(K) \subset \mathscr{M}^{\mathrm{qc}}(K)$. Set $A_1 := [\mu_1]$, $A_2 := [\mu_2]$ and $F := \theta A_1 + (1-\theta)A_2$, such that $A_1 - A_2$ has rank at most one.

Apply the construction of the proof of Proposition 5.3 to get a sequence $(v_j) \subset \mathrm{W}^{1,\infty}_{Fx}(Q;\mathbb{R}^m)$, where $Q \subset \mathbb{R}^d$ is a rotated unit-volume cube, with

$$\nabla v_j \overset{\mathbf{Y}}{\to} \theta\delta_{A_1} + (1-\theta)\delta_{A_2}.$$

Now, on each solid piece of Q where ∇v_j is equal to A_1, say $R_1^{(j)}, \ldots, R_M^{(j)} \subset Q$, we re-define ∇v_j to be equal to the n'th element of a generating sequence for μ_1 (by Lemma 4.14 we may choose the domain of definition freely and by Lemma 4.13 we can ensure that the sequence has boundary values $A_1 x$). Similarly, if ∇v_j is equal to A_2 on the solid pieces $S_1^{(j)}, \ldots, S_N^{(j)} \subset Q$, then re-define ∇v_j to be equal to the n'th element of a generating sequence for μ_2. This yields a map $w_{j,n} \in \mathrm{W}^{1,\infty}_{Fx}(Q;\mathbb{R}^m)$ (n is the index of the "inner" sequences), which agrees with v_j outside the sets $R^{(j)} := \bigcup_m R_m^{(j)}$ and $S^{(j)} := \bigcup_m S_m^{(j)}$.

Let $\{\varphi_k \otimes h_k\}_{k\in\mathbb{N}} \subset \mathrm{C}_0(Q) \otimes \mathrm{C}_0(\mathbb{R}^{m\times d})$ be the countable set of integrands from Lemma 4.7. Then we we may choose $n = n(j)$ sufficiently large so that

$$\left| \int_{R^{(j)}\cup S^{(j)}} h_k(\nabla w_{j,n(j)}) \,\mathrm{d}x - |R^{(j)}| \cdot \int h_k \,\mathrm{d}\mu_1 - |S^{(j)}| \cdot \int h_k \,\mathrm{d}\mu_2 \,\mathrm{d}x \right| \leq \frac{1}{j}$$

for all $k \leq j$. Since we also have that $|R^{(j)}| \to \theta$ and $|S^{(j)}| \to 1-\theta$ as well as

$$\left| \int_{Q\setminus(R^{(j)}\cup S^{(j)})} \varphi_k h_k(\nabla v_j) \,\mathrm{d}x \right| \to 0 \qquad \text{as } j \to \infty,$$

we infer via Lemma 4.7 that $\nabla w_{j,n(j)} \overset{\mathbf{Y}}{\to} \nu \in \mathbf{GY}^\infty(Q;\mathbb{R}^{m\times d})$ with

$$\int_Q \int h_k \,\mathrm{d}\nu_x \,\mathrm{d}x = \theta \int h_k \,\mathrm{d}\mu_1 + (1-\theta) \int h_k \,\mathrm{d}\mu_2 \qquad \text{for all } k \in \mathbb{N}.$$

Finally, using the averaging principle of Lemma 4.14 (this additional averaging is in fact not really necessary, as one can check through a closer inspection of the proof of Proposition 5.3), we get a sequence $(u_j) \subset \mathrm{W}^{1,\infty}_{Fx}(Q;\mathbb{R}^m)$ such that

$$\nabla u_j \overset{\mathbf{Y}}{\to} \overline{\nu} = \theta\mu_1 + (1-\theta)\mu_2 = \mu.$$

Thus, $\mu \in \mathscr{M}^{\mathrm{qc}}(K)$ and the proof is finished. □

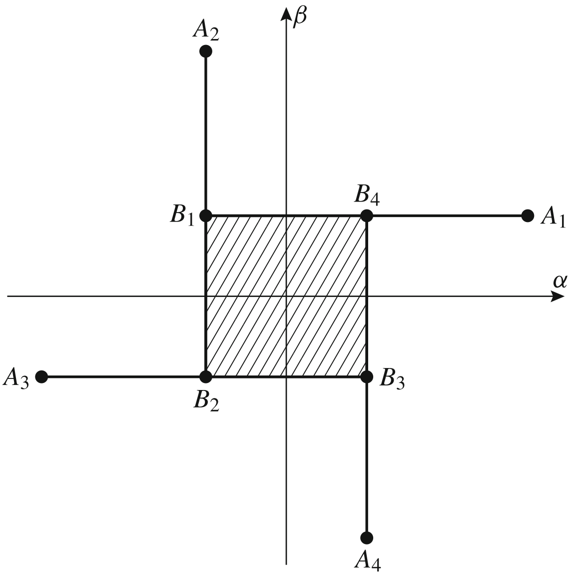


Fig. 9.2 The quasiconvex hull $K_{\mathrm{T4}}^{\mathrm{qc}}$ of the T_4-configuration (hatched area and solid lines) in the space of diagonal matrices with entries α, β

As a first application of the preceding lemma, we may now calculate the various hulls of the T_4-configuration.

Proposition 9.4. *For* $K_{\mathrm{T4}} = \{A_1, A_2, A_3, A_4\}$ *from Proposition 8.17 we have (see Figure 9.2)*

$$K_{\mathrm{T4}} = K_{\mathrm{T4}}^{\mathrm{lc}} \subsetneq K_{\mathrm{T4}}^{\mathrm{rc}} = K_{\mathrm{T4}}^{\mathrm{qc}} = K_{\mathrm{T4}}^{\mathrm{pc}} = \{B_1, B_2, B_3, B_4\}^{**} \cup \bigcup_{i=1}^{4} \{A_i, B_i\}^{**},$$

where

$$B_1 := \begin{pmatrix} -1 & \\ & 1 \end{pmatrix}, \quad B_2 := \begin{pmatrix} -1 & \\ & -1 \end{pmatrix}, \quad B_3 := -B_1, \quad B_4 := -B_2.$$

Here, $\{B_1, B_2, B_3, B_4\}^{**}$ and $\{A_i, B_i\}^{**}$ denote the closed square with vertices B_1, B_2, B_3, B_4 and the (closed) line segment from A_i to B_i, respectively.

Proof. The fact that $K_{\mathrm{T4}} = K_{\mathrm{T4}}^{\mathrm{lc}}$ is clear since there are no rank-one connections in K_{T4}.

Step 1. We first compute $K_{\mathrm{T4}}^{\mathrm{rc}}$ as a lower bound on $K_{\mathrm{T4}}^{\mathrm{qc}}$ and $K_{\mathrm{T4}}^{\mathrm{pc}}$; this technique is sometimes called the *lamination method*. By (the proof of) Proposition 8.17 we have that the four intermediate matrices B_1, B_2, B_3, B_4 lie in $K_{\mathrm{T4}}^{\mathrm{rc}}$. Hence also the rank-one lines between B_i and B_{i+1} lie in $K_{\mathrm{T4}}^{\mathrm{rc}}$, where $i \in \{1, 2, 3, 4\}$ is a cyclic index. Since its edges are rank-one lines, the closed square $\{B_1, B_2, B_3, B_4\}^{**}$ then lies in $K_{\mathrm{T4}}^{\mathrm{rc}}$ as well. Clearly, the additional rank-one lines $\{A_i, B_i\}^{**}$ also lie in that set. So, we have shown that

$$\hat{K} := \{B_1, B_2, B_3, B_4\}^{**} \cup \bigcup_{i=1}^{4} \{A_i, B_i\}^{**} \subset K_{\mathrm{T4}}^{\mathrm{rc}} \subset K_{\mathrm{T4}}^{\mathrm{qc}}.$$

Step 2. Next, we show that $K_{\mathrm{T4}}^{\mathrm{qc}} \subset \hat{K}$, for which we employ the *separation method*, which we already used in the proof of Theorem 8.11. We first observe that since $K^{\mathrm{qc}} \subset K^{**}$ and all matrices in K are diagonal, the same must hold for K^{qc}.

Let $F \in K^{\mathrm{qc}}$ and $\mu \in \mathscr{M}^{\mathrm{qc}}(K)$ with $F = [\mu]$. In Lemma 8.14 we saw that the function

$$\det{}^{++}(A) := \begin{cases} \det A & \text{if } A \text{ is positive semidefinite,} \\ 0 & \text{otherwise,} \end{cases} \qquad A \in \mathbb{R}^{2\times 2}_{\mathrm{sym}},$$

is quasiconvex on symmetric matrices, i.e.,

$$\det{}^{++}(A) \le \fint_{B(0,1)} \det{}^{++}(A + \nabla^2 \psi(z)) \,\mathrm{d}z$$

for all $A \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}$ and all $\psi \in \mathrm{W}^{2,2}_c(B(0,1))$. Then, $A \mapsto \det{}^{++}(A - B_1)$ is also quasiconvex on symmetric matrices. This corresponds to shifting the whole T_4-configuration so that $B_1 = 0$.

Via the Jensen-type inequality from Lemma 8.15, we thus get

$$\det{}^{++}(F - B_1) \le \int \det{}^{++}(A - B_1) \,\mathrm{d}\mu(A) = 0.$$

The last equality follows since $\det{}^{++}(A_i - B_1) = 0$ for $i = 1, 2, 3, 4$. Therefore, we also must have $\det{}^{++}(F - B_1) = 0$, that is, $F - B_1$ is not positive definite. Thus, F cannot lie in the hatched region in Figure 9.3.

Repeating this argument after rotating the T_4-configuration, we see that it must hold that $F \in \hat{K}$, whereby $K_{\mathrm{T4}}^{\mathrm{qc}} \subset \hat{K}$. Together with Step 1 we have thus shown that $K_{\mathrm{T4}}^{\mathrm{rc}} = K_{\mathrm{T4}}^{\mathrm{qc}} = \hat{K}$.

Step 3. We finally prove that $K^{\mathrm{pc}} \subset \hat{K}$. In fact, the proof of this is the same as the previous step since

$$\det{}^{++}\left(\begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}\right) = \alpha^+ \beta^+,$$

where $s^+ := \max\{s, 0\}$ for $s \in \mathbb{R}$. So, $\det^{++}$ equals a polyconvex function on the subspace of diagonal matrices. □

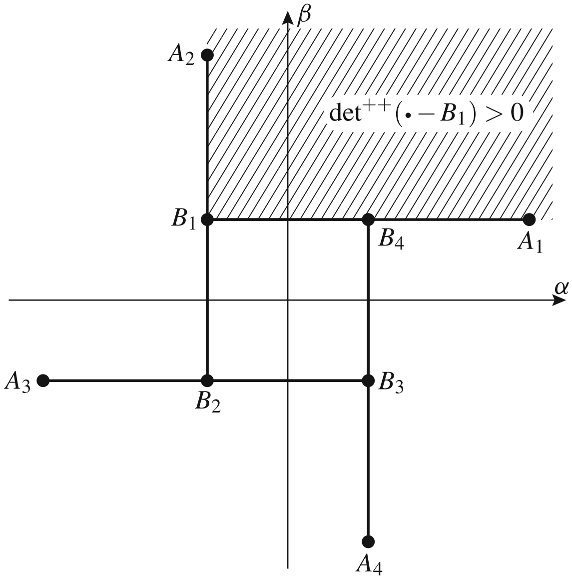


Fig. 9.3 One step in the determination of an upper bound for $K_{\mathrm{T4}}^{\mathrm{qc}}$

9.2 Multi-well Inclusions

Returning to the two-well problem in two dimensions, we now investigate the non-rigid case. Thus, consider (after normalization)

$$K := \mathrm{SO}(2) \cup \mathrm{SO}(2)U, \quad U := \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}, \quad 0 < \alpha \leq 1 \leq \beta, \quad \alpha\beta \geq 1. \tag{9.2}$$

By Lemma 8.25, for $\alpha = 1$ the two wells are simply rank-one connected and for $\alpha < 1$ they are doubly rank-one connected. In order to describe the hulls of K efficiently, it is convenient to introduce the so-called **conformal coordinates**: For any $A \in K^{**}$ we may write

$$A = \begin{pmatrix} y_1 & -y_2 \\ y_2 & y_1 \end{pmatrix} + \begin{pmatrix} z_1 & -z_2 \\ z_2 & z_1 \end{pmatrix} U =: (y, z),$$

where $y, z \in \mathbb{R}^2$ with

$$|y|^2 = y_1^2 + y_2^2 \leq 1 \qquad \text{and} \qquad |z|^2 = z_1^2 + z_2^2 \leq 1,$$

see Problem 9.5.

The following result then completely describes all the hulls of K:

Theorem 9.5 (**Šverák 1993 [254]**). *Let K be as in (9.2).*

(i) If $\alpha = 1$ (simple rank-one connection between wells), then

$$\begin{aligned} K^{\mathrm{lc}} &= K^{\mathrm{rc}} = K^{\mathrm{qc}} = K^{\mathrm{pc}} \\ &= \left\{ QU_s \in \mathbb{R}^{2\times 2} \;:\; Q \in \mathrm{SO}(2), U_s = \begin{pmatrix} 1 & \\ & s \end{pmatrix}, s \in [1, \beta] \right\}. \end{aligned}$$

(ii) If $\alpha < 1$ (double rank-one connection between wells) and $\det U = \alpha\beta = 1$, *then*

$$K^{\mathrm{lc}} = K^{\mathrm{rc}} = K^{\mathrm{qc}} = K^{\mathrm{pc}} = \left\{ A = (y, z) \in K^{**} : |y|+|z| \le 1 \text{ and } \det A = 1 \right\}.$$

(iii) If $\alpha < 1$ and $\det U = \alpha\beta > 1$, *then*

$$\begin{aligned} K^{\mathrm{lc}} &= K^{\mathrm{rc}} = K^{\mathrm{qc}} = K^{\mathrm{pc}} \\ &= \left\{ A = (y, z) \in K^{**} \;:\; |y| \le \tfrac{\det U - \det A}{\det U - 1} \text{ and } |z| \le \tfrac{\det A - 1}{\det U - 1} \right\}. \end{aligned}$$

Proof. Ad (i). Let $F \in K^{\mathrm{pc}}$ and take $\mu \in \mathscr{M}^{\mathrm{pc}}(K)$ with $[\mu] = F$. We will show that $\mu \in \mathscr{M}^{\mathrm{lc},1}(K)$, which will allow us to conclude the assertion via Lemma 9.3. By the structure of K, we may write

$$\mu = \theta\gamma + (1-\theta)R(U)_{\#}\eta, \qquad \text{where} \qquad \gamma, \eta \in \mathscr{M}^1(\mathrm{SO}(2)),\; \theta \in [0, 1].$$

Here, $R(U)_{\#}\eta(B) := \eta(B \cdot U^{-1})$ for any Borel set $B \subset \mathbb{R}^{2\times 2}$. Then, with $G := [\gamma]$, $H := [\eta]$, we have $F = \theta G + (1-\theta)HU$ and, since $G, H \in \mathrm{SO}(2)^{**}$, it holds that (see Appendix A.1)

$$G = \begin{pmatrix} g_1 & -g_2 \\ g_2 & g_1 \end{pmatrix}, \qquad H = \begin{pmatrix} h_1 & -h_2 \\ h_2 & h_1 \end{pmatrix}$$

with $g, h \in \mathbb{R}^2$ such that

$$\det G = |g|^2 = g_1^2 + g_2^2 \le 1 \qquad \text{and} \qquad \det H = |h|^2 = h_1^2 + h_2^2 \le 1.$$

Thus, we can compute, using Young's inequality,

$$\begin{aligned} G : \mathrm{cof}(HU) &= \begin{pmatrix} g_1 & -g_2 \\ g_2 & g_1 \end{pmatrix} : \begin{pmatrix} h_1 & -\beta h_2 \\ h_2 & \beta h_1 \end{pmatrix} \\ &= (1+\beta)(g_1 h_1 + g_2 h_2) \\ &\le \frac{1+\beta}{2}(|g|^2 + |h|^2) \\ &\le 1 + \beta \end{aligned}$$

and equality can only hold if $G = H$ and $\det G = \det H = 1$. From the definition of polyconvex measures we further have

$$\det F = \theta + (1-\theta)\beta.$$

Then, using (8.15) and the above considerations, we compute

$$\begin{aligned}\det F &= \det(\theta G + (1-\theta)HU)\\ &= \theta^2 \det G + \theta(1-\theta)G : \operatorname{cof}(HU) + (1-\theta)^2 \det H \det U\\ &\le \theta^2 + \theta(1-\theta)(1+\beta) + (1-\theta)^2\beta\\ &= \theta + (1-\theta)\beta\\ &= \det F.\end{aligned}$$

Hence, the above inequality is in fact an equality, from which we infer $G : \operatorname{cof}(HU) = 1 + \beta$, whereby $G = H$ and $\det G = \det H = 1$. This implies $G = H \in \mathrm{SO}(2)$ and

$$\mu = \theta\delta_G + (1-\theta)\delta_{GU} \in \mathscr{M}^{\mathrm{lc},1}(K) \tag{9.3}$$

since any probability measure $\sigma \in \mathscr{M}^1(\mathrm{SO}(2)^{**})$ with $[\sigma] \in \mathrm{SO}(2)$ is a Dirac mass. Indeed, for

$$A = \begin{pmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{pmatrix} \in \mathrm{SO}(2)^{**}$$

we have $\det A = a_1^2 + a_2^2 \le 1$ and so the determinant is a strictly convex function on $\mathrm{SO}(2)^{**}$. Thus, by Jensen's inequality,

$$1 = \det\,[\sigma] \le \int \det A \,\mathrm{d}\sigma(A) \le 1.$$

Consequently, σ is carried by SO(2), but it is elementary to see that any non-trivial convex combination of matrices in SO(2) has determinant strictly less than 1. Hence, $\operatorname{supp}\sigma$ must be a singleton.

Finally, from (9.3) it also follows that

$$[\mu] = G(\theta \mathrm{Id} + (1-\theta)U),$$

which immediately implies the claimed formula for $K^{\mathrm{lc}} = K^{\mathrm{rc}} = K^{\mathrm{qc}} = K^{\mathrm{pc}}$.

Ad (ii). From the definition of the polyconvex hull we have that $\mu \in \mathscr{M}^{\mathrm{pc}}(K)$ if and only if $\mu \in \mathscr{M}^1(\mathbb{R}^{2\times 2})$ and

$$H([\mu], \det[\mu]) \le \int H(A, \det A)\,\mathrm{d}\mu(A) \qquad \text{for all convex } H \in \mathrm{C}(\mathbb{R}^{2\times 2}\times\mathbb{R}).$$

By Jensen's inequality, we thus get

$$K^{\rm pc} = \{ A \in \mathbb{R}^{2\times 2} : (A, \det A) \in (\mathbf{M}K)^{**} \}, \quad \mathbf{M}K := \{ (A, \det A) : A \in K \},$$

also see Problem 9.4. Employing the conformal coordinates, we can thus represent $A \in K$ as the pair $(y, z) \in \mathbb{R}^2 \times \mathbb{R}^2$, where either $|y| = 1$ and $z = 0$, or $|z| = 1$ and $y = 0$. To determine $K^{\rm pc}$ we then need to compute the convex hull G_κ^{**} of the set

$$G_\kappa := \{ (y, 0, 1) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R} : |y| = 1 \} \cup \{ (0, z, \kappa) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R} : |z| = 1 \}$$

in the case $\kappa = \det U = 1$. Since this set is invariant under rotations of the matrices represented by y, z, it suffices to compute this convex hull with $\mathbb{R}^1$ in place of $\mathbb{R}^2$, which is elementary. We get

$$G_1^{**} = \{ (y, z, 1) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R} : |y| + |z| \le 1 \}, \tag{9.4}$$

whereby

$$K^{\rm pc} = \{ A = (y, z) \in K^{**} : |y| + |z| \le 1 \text{ and } \det A = 1 \}.$$

By Lemma 9.3 it remains to show that $K^{\rm pc} \subset K^{\rm lc}$.

Let $F \in K^{\rm pc} = K^{**} \cap \{\det = 1\}$, where here and in the following we simply write "$\{\det = 1\}$" for the set $\{ A \in \mathbb{R}^{2\times 2} : \det A = 1 \}$. The equality follows by considering G_1 again. We distinguish two cases:

(a) If $F = (\bar{y}, \bar{z}) \in \partial K^{**} \cap \{\det = 1\}$, then by (9.4) it holds that $|\bar{y}| + |\bar{z}| = 1$. For $\bar{y} = 0$ or $\bar{z} = 0$ we immediately get $A \in K \subset K^{\rm lc}$. If $\bar{y} \neq 0$ and $\bar{z} \neq 0$, we define

$$q(t) := \det\Big[t\Big(\frac{\bar{y}}{|\bar{y}|}, 0\Big) + (1-t)\Big(0, \frac{\bar{z}}{|\bar{z}|}\Big)\Big], \qquad t \in [0, 1].$$

Clearly, q is a quadratic function in t and $q(|\bar{y}|)$ is the determinant of the matrix with representation $(\bar{y}, \bar{z})$, that is, our F. Thus, $q(0) = q(1) = q(|\bar{y}|) = 1$ and hence $q \equiv 1$, whereby $\bar{y}/|\bar{y}|$ and $\bar{z}/|\bar{z}|$ are rank-one connected. We conclude that F lies on the rank-one line between the matrices $(\bar{y}/|\bar{y}|, 0) \in K$ and $(0, \bar{z}/|\bar{z}|) \in K$, that is, $F \in K^{\rm lc}$.

(b) Let now $F \in (K^{**} \setminus \partial K^{**}) \cap \{\det = 1\}$. It is always possible to find $a \in \mathbb{R}^2 \setminus \{0\}$, $n \in \mathbb{S}^{d-1}$ such that $\operatorname{cof} F : (a \otimes n) = a^T (\operatorname{cof} F) n = 0$. Then, by Jacobi's formula (see Appendix A.1),

$$t \mapsto \det[F + t(a \otimes n)] \equiv \text{const},$$

whereby the determinant constraint is preserved along the rank-one line $F + \mathbb{R}(a \otimes n)$. This line intersects the set $\partial(K^{**}) \cap \{\det = 1\}$ for precisely two values of t. Since we have already shown that $\partial(K^{**}) \cap \{\det = 1\} \subset K^{\mathrm{lc}}$, we also get that $F \in K^{\mathrm{lc}}$.

Ad (iii). By a similar reasoning as at the beginning of Step (ii), we get

$$G_\kappa^{**} = \left\{ (y, z, s) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R} \,:\, |y| \leq \tfrac{\kappa - s}{\kappa - 1} \text{ and } |z| \leq \tfrac{s-1}{\kappa-1} \right\},$$

where this time $\kappa = \det U > 1$. Then, denoting by $\det(y, z)$ the determinant of the matrix with conformal coordinates (y, z), we have

$$K^{\mathrm{pc}} = \left\{ (y, z) \in K^{**} \,:\, |y| \leq \tfrac{\det U - \det(y,z)}{\det U - 1} \text{ and } |z| \leq \tfrac{\det(y,z)-1}{\det U - 1} \right\}. \tag{9.5}$$

In the following we will show that $\partial K^{\mathrm{pc}} \subset K^{\mathrm{lc}}$. This suffices to prove the claim since every rank-one line through an interior point of K^{pc} intersects ∂K^{pc} in precisely two points.

So, let $F = (\bar{y}, \bar{z}) \in \partial K^{\mathrm{pc}}$. Define

$$\begin{aligned} f(y, z) &:= (\det U - 1)|y| - \det U + \det(y, z), \\ g(y, z) &:= (\det U - 1)|z| + 1 - \det(y, z). \end{aligned}$$

Then, (9.5) can be rewritten as

$$K^{\mathrm{pc}} = \left\{ (y, z) \in \mathbb{R}^{2\times 2} \,:\, f(y, z) \leq 0 \text{ and } g(y, z) \leq 0 \right\}.$$

As $F = (\bar{y}, \bar{z}) \in \partial K^{\mathrm{pc}}$, either $f(\bar{y}, \bar{z}) = 0$ or $g(\bar{y}, \bar{z}) = 0$. We distinguish three cases:

(a) If $f(\bar{y}, \bar{z}) = 0$ and $g(\bar{y}, \bar{z}) = 0$, then a simple calculation shows that $|\bar{y}| + |\bar{z}| = 1$. Along the line

$$M(t) := t\left(\frac{\bar{y}}{|\bar{y}|}, 0\right) + (1 - t)\left(0, \frac{\bar{z}}{|\bar{z}|}\right), \qquad t \in [0, 1],$$

the functions f, g are quadratic in t and have three zeros, at $t = 0, 1, |\bar{y}|$. Thus, f, g are identically zero along the line $M(t)$. Since $\det U > 1$, from the definition of f or g we conclude that $\det(y, z)$ grows linearly in $|t|$ and thus $M(t)$ must be a rank-one line. Consequently, $F = M(|\bar{y}|) \in K^{\mathrm{lc}}$.

(b) If $f(\bar{y}, \bar{z}) < 0$ and $g(\bar{y}, \bar{z}) = 0$, we first assume without loss of generality that $\bar{z}_2 = 0$ by the SO(2)-symmetry of K, K^{pc}. We also assume $\bar{z}_1 > 0$. The case $\bar{z}_1 < 0$ is similar. In the three-dimensional space $Y := \{(y_1, y_2, z_1, 0) \in \mathbb{R}^4\}$ the equation

$$
\begin{aligned}
0 &= g(y_1, y_2, z_1, 0) \\
&= (\det U - 1)z_1 + 1 - \det(y_1, y_2, z_1, 0) \\
&= (\alpha\beta - 1)z_1 + 1 - (y_1, y_2, z_1) \begin{pmatrix} 1 & & \frac{\alpha+\beta}{2} \\ & 1 & \\ \frac{\alpha+\beta}{2} & & \alpha\beta \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ z_1 \end{pmatrix}
\end{aligned}
$$

defines a one-sheeted hyperboloid H since the matrix has two positive eigenvalues and one negative eigenvalue as well as $\alpha\beta > 1$. Since every one-sheeted hyperboloid has the property that through every point on it there pass two lines lying entirely inside it, we can thus find a line $M(t)$ in H with $M(0) = F$. On $M(t)$, however, the determinant is affine by construction of H and thus $M(t)$ must be a rank-one line. Write $M(t) = (y(t), z(t))$. Then, choosing $t_0 \in \mathbb{R}$ such that $z_1(t_0) = 0$, we see that $M(t_0) = (y(t_0), 0) \in K$. We also observe

$$
f(M(t)) = f(M(t)) + g(M(t)) = (\det B - 1)(|y(t)| + |z(t)| - 1) \to \infty
$$

as $|t| \to \infty$. Combining this with the fact that $f(M(0)) = f(F) < 0$, there must be a $t_1 \in \mathbb{R}$ with the property that $f(M(t_1)) = g(M(t_1)) = 0$ and thus $M(t_1) \in K^{\mathrm{lc}}$ by (a). Consequently, since F lies on a rank-one line between $M(t_0) \in K$ and $M(t_1) \in K^{\mathrm{lc}}$, it holds that $F = M(0) \in K^{\mathrm{lc}}$.

(c) The case $f(y_0, z_0) = 0$ and $g(y_0, z_0) < 0$ is analogous to the previous one.

In all cases we have shown that $F \in K^{\mathrm{lc}}$ and thus $K^{\mathrm{pc}} \subset K^{\mathrm{lc}}$, which implies the claim. □

For more general inclusions in two dimensions, including the N-well problem, we quote the following result concerning the case of wells with the same determinant:

Theorem 9.6 **(Bhattacharya–Dolzmann 2001 [42]).** *Let $\delta > 0$ and assume that the set*

$$
K \subset \left\{ A \in \mathbb{R}^{2\times 2} \;:\; \det A = \delta \right\}
$$

is compact and left SO(2)*-invariant, i.e.,*

$$
K = \mathrm{SO}(2)K = \left\{ QA \;:\; Q \in \mathrm{SO}(2),\; A \in K \right\}.
$$

Then,

$$
\begin{aligned}
K^{\mathrm{lc}} &= K^{\mathrm{rc}} = K^{\mathrm{qc}} = K^{\mathrm{pc}} \\
&= \left\{ A \in \mathbb{R}^{2\times 2} \;:\; \det A = \delta \textit{ and } |An|^2 \le \max_{B\in K} |Bn|^2 \textit{ for all } n \in \mathbb{S}^1 \right\}.
\end{aligned}
$$

A proof can be found in Section 2.2 of [100] and in the original work [42]. For three dimensions, we only mention the following result.

Theorem 9.7 (Dolzmann–Kirchheim–Müller–Šverák 2000 [101]). *Let*

$$K := \mathrm{SO}(3) \cup \mathrm{SO}(3)U$$

with

$$U = \begin{pmatrix} \alpha & & \\ & \beta & \\ & & 1 \end{pmatrix}, \qquad 0 < \alpha \le 1 \le \beta, \qquad \alpha\beta \ge 1.$$

Then,

$$K^{\mathrm{lc}} = K^{\mathrm{rc}} = K^{\mathrm{qc}} = K^{\mathrm{pc}} = \left\{ Q \begin{pmatrix} \hat{A} & \\ & 1 \end{pmatrix} \;:\; Q \in \mathrm{SO}(3),\ \hat{A} \in \hat{K}^{\mathrm{qc}} \right\},$$

where

$$\hat{K} = \mathrm{SO}(2) \cup \mathrm{SO}(2) \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix},$$

and $\hat{K}^{\mathrm{qc}}$ can be explicitly calculated via Theorem 9.5.

9.3 Convex Integration

The considerations in Section 8.3 revealed that for a non-empty compact set $K \subset \mathbb{R}^{m\times d}$ the differential inclusion

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m), \quad u|_{\partial\Omega} = Fx, \\ \nabla u \in K \end{cases} \tag{9.6}$$

has an approximate solution if and only if $F \in K^{\mathrm{qc}}$. Now we investigate the finer question of which additional hypotheses are required to solve (9.6) *exactly*, that is, whether there exists a map $u \in \mathrm{W}^{1,\infty}_{Fx}(\Omega; \mathbb{R}^m)$ with

$$\nabla u(x) \in K \qquad \text{for a.e. } x \in \Omega.$$

We first illustrate why this question is quite delicate in general: Let K_{T4} be the T_4-configuration from Proposition 8.17. We showed in Proposition 9.4 that $K_{\mathrm{T4}}^{\mathrm{qc}} \supsetneq K_{\mathrm{T4}}$. On the other hand, we know from the Chlebík–Kirchheim Theorem 8.16 that (9.6) is rigid for exact solutions. Hence, only for $F \in K_{\mathrm{T4}}$ can we solve (9.6) exactly and then the solution is necessarily affine.

Let us also consider the non-rigid two-well inclusion in two dimensions,

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^2), \quad u|_{\partial\Omega} = Fx, \quad 0 < \alpha < 1 < \beta, \quad \alpha\beta \geq 1, \\ \nabla u \in K := \mathrm{SO}(2) \cup \mathrm{SO}(2)\begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}. \end{cases} \tag{9.7}$$

By Theorem 9.5 in conjunction with Lemma 8.25 (iii) we again have $K^{\mathrm{qc}} \supsetneq K$. In fact, by Lemma 8.25 (iii), for every $Q \in \mathrm{SO}(2)$ there are precisely two solutions $R_1, R_2 \in \mathrm{SO}(2)$, $i = 1, 2$, of

$$\operatorname{rank}\left(Q - R\begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}\right) \leq 1.$$

Consequently, two laminates are possible, but there seems to be no obvious way to combine them. As every simple laminate with linear (or affine) boundary values is trivial, at first sight it thus appears to be impossible to solve (9.6) exactly for $F \in K^{\mathrm{qc}} \setminus K$.

This conjecture turns out to be false, but the resulting exact solutions are necessarily highly irregular. The general approach to constructing such exact solutions is called **convex integration**. Its basic idea is to start with a map satisfying the given boundary conditions and to consecutively add oscillations with high frequency and low amplitude in a controlled fashion to move the gradient toward K.

The property of K that will allow us to implement this idea is the following: We say that the open and bounded set $G \subset \mathbb{R}^{m\times d}$ can be **piecewise affinely reduced** to a compact and non-empty set $K \subset \mathbb{R}^{m\times d}$ if for all $A \in G$ there exists a sequence $(\psi_j) \subset \mathrm{W}^{1,\infty}_0(B(0,1);\mathbb{R}^m)$ of countably piecewise affine maps such that

(i) $A + \nabla\psi_j(x) \in G$ for almost every $x \in B(0,1)$ and all $j \in \mathbb{N}$;

(ii) $\displaystyle\int_{B(0,1)} \operatorname{dist}(A + \nabla\psi_j(x), K)\,\mathrm{d}x \to 0$ as $j \to \infty$.

Recall that a map $\varphi\colon D \to \mathbb{R}^m$ is called **countably piecewise affine** if there exists a disjoint partition of D into countably many open sets D_k up to a negligible set, i.e., $D = Z \cup \bigcup_{k\in\mathbb{N}} D_k$, where $|Z| = 0$, such that $\varphi|_{D_k}$ is affine. Throughout this chapter, "piecewise affine" will always mean "countably piecewise affine".

It can be shown via a standard covering argument that in the above definition the unit ball can be replaced by any bounded Lipschitz domain; see Lemma 5.2 for a similar statement regarding the definition of quasiconvexity. Moreover, if G can be piecewise-affinely reduced to K, then every $A \in G$ must necessarily lie in K^{qc}, cf. the definition in Section 8.3.

Then we have the following (abstract) convex integration result, which in this form is due to Sychev, but builds crucially on earlier works by Gromov and Müller–Šverák.

Theorem 9.8 (Gromov 1986, Müller–Šverák 1996, Sychev 1998 [145, 205, 257]). *Let $K \subset \mathbb{R}^{m\times d}$ be a non-empty compact set and assume that the open and bounded set $G \subset \mathbb{R}^{m\times d}$ can be piecewise affinely reduced to K. Suppose that either $v \in \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^m)$ is piecewise affine or $v \in \mathrm{C}^1(\Omega;\mathbb{R}^m)$, and that*

$$\nabla v \in G \quad a.e.$$

Then, for every $\varepsilon > 0$ *there exists a solution to the exact differential inclusion*

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m), \quad u|_{\partial\Omega} = v|_{\partial\Omega}, \\ \nabla u \in K \quad a.e. \end{cases} \tag{9.8}$$

with $\|u - v\|_{\mathrm{L}^\infty} \le \varepsilon$.

Proof. Let $\varepsilon > 0$ be given. If v is already piecewise affine, then set $v_1 := v$. If v instead has C^1-regularity, then find a sequence of disjoint open polyhedral sets $\Omega_k \Subset \Omega$ such that $\Omega = Z \cup \bigcup_{k=1}^{\infty} \Omega_k$, where $|Z| = 0$. In each Ω_k we may approximate v on a fine triangulation with a (finitely) piecewise affine map $\tilde{v}_k$ such that $\|v - \tilde{v}_k\|_{\mathrm{L}^\infty} \le \varepsilon 2^{-(k+1)}$ and such that the $\tilde{v}_k$ agree where they meet over the boundaries of the Ω_k. Since G is open, we may also require $\nabla \tilde{v}_k \in G$. Combining these maps into a single piecewise affine map $\tilde{v}$, which necessarily satisfies $\tilde{v}|_{\partial\Omega} = v|_{\partial\Omega}$, we are again in the situation of a piecewise affine $v_1 := \tilde{v}$. So, we may assume that v_1 is piecewise affine and

$$\nabla v_1 \in G \quad \text{a.e.}, \qquad v_1|_{\partial\Omega} = v|_{\partial\Omega}, \qquad \|v - v_1\|_{\mathrm{L}^\infty} \le \frac{\varepsilon}{2}.$$

The idea of the proof is to use the piecewise affine reduction of G to K in conjunction with a "controlled" mollification to construct a sequence (v_j) that converges *strongly* in $\mathrm{W}^{1,1}$ to the sought solution u of (9.8).

Let $(\eta_\delta)_{\delta > 0}$ be a family of mollifiers and choose $\delta_1 > 0$ such that

$$\delta_1 \le \varepsilon \qquad \text{and} \qquad \|\nabla v_1 - \eta_{\delta_1} \star \nabla v_1\|_{\mathrm{L}^1} \le \frac{1}{2}.$$

Now, assuming that we have already defined v_j and δ_j, we use the assumption that G can be piecewise affinely reduced to K to find a piecewise affine map $v_{j+1} \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ with

$$\nabla v_{j+1} \in G \quad \text{a.e.}, \qquad v_{j+1}|_{\partial\Omega} = v|_{\partial\Omega}, \qquad \|v_{j+1} - v_j\|_{\mathrm{L}^\infty} \le \frac{\delta_j}{2^{j+1}}$$

and

$$\int_\Omega \mathrm{dist}(\nabla v_{j+1}(x), K) \, \mathrm{d}x \le \frac{1}{2^{j+1}}. \tag{9.9}$$

Indeed, we can construct this map separately in each subdomain D_k of Ω from the partition with respect to which v_j is piecewise affine, say $v_j(x) = z_k + A_k x$ for $x \in D_k$, where $z_k \in \mathbb{R}^m$, $A_k \in \mathbb{R}^{m \times d}$ ($k \in \mathbb{N}$). From the reduction property choose $\psi_{j+1}^{(k)} \in \mathrm{W}_0^{1,\infty}(D_k; \mathbb{R}^m)$ piecewise affine with $\nabla \psi_{j+1}^{(k)} \in G$ almost everywhere and

$$\int_{D_k} \operatorname{dist}(A_k + \nabla\psi^{(k)}_{j+1}(x), K)\,\mathrm{d}x \le \frac{1}{2^{j+k+1}}.$$

Moreover, by a procedure similar to the proof of the averaging principle for Young measures, Lemma 4.14, we may assume that $\|\psi^{(k)}_{j+1}\|_{\mathrm{L}^\infty} \le 2^{-(j+1)}\delta_j$. Then,

$$v_{j+1}(x) := z_k + A_k x + \psi^{(k)}_{j+1}(x) \qquad \text{if } x \in D_k \ (k \in \mathbb{N})$$

has the desired properties.

Also choose $0 < \delta_{j+1} \le \delta_j$ such that

$$\|\nabla v_{j+1} - \eta_{\delta_{j+1}} \star \nabla v_{j+1}\|_{\mathrm{L}^1} \le \frac{1}{2^{j+1}},$$

where we consider the map v_{j+1} to be continuously extended to all of $\mathbb{R}^d$. For all $l \ge j$ we get

$$\|v_l - v_j\|_{\mathrm{L}^\infty} \le \sum_{k=j}^{l-1} \|v_{k+1} - v_k\|_{\mathrm{L}^\infty} \le \sum_{k=j}^{\infty} \frac{\delta_k}{2^{k+1}} \le \frac{\delta_j}{2^j} \le \frac{\varepsilon}{2^j}.$$

In particular, (v_j) is an L^∞-Cauchy sequence and hence $v_j \to u$ in L^∞ for some $u \in \mathrm{L}^\infty(\Omega; \mathbb{R}^m)$ with

$$\|u - v_1\|_{\mathrm{L}^\infty} \le \frac{\varepsilon}{2} \qquad \text{and} \qquad \|u - v_j\|_{\mathrm{L}^\infty} \le \frac{\delta_j}{2^j}.$$

Furthermore, also the gradients of the v_j's converge *strongly* in L^1: To this end we observe that since $\|\nabla\eta_{\delta_j}\|_{\mathrm{L}^1} \le C/\delta_j$ for some constant $C > 0$,

$$\|\eta_{\delta_j} \star (\nabla u - \nabla v_j)\|_{\mathrm{L}^1} = \|\nabla\eta_{\delta_j} \star (u - v_j)\|_{\mathrm{L}^1} \le \frac{C}{\delta_j} \cdot \frac{\delta_j}{2^j} \to 0.$$

Then,

$$\begin{aligned}\|\nabla u - \nabla v_j\|_{\mathrm{L}^1} &\le \|\nabla u - \eta_{\delta_j} \star \nabla u\|_{\mathrm{L}^1} + \|\eta_{\delta_j} \star (\nabla u - \nabla v_j)\|_{\mathrm{L}^1} \\ &\qquad + \|\eta_{\delta_j} \star \nabla v_j - \nabla v_j\|_{\mathrm{L}^1} \\ &\to 0.\end{aligned}$$

Thus, $v_j \to u$ in $\mathrm{W}^{1,1}$ and, selecting a subsequence, also $\nabla v_j \to \nabla u$ almost everywhere. By (9.9), we arrive at the inclusion $\nabla u \in K$ almost everywhere. Moreover, $\|u - v\|_{\mathrm{L}^\infty} \le \varepsilon$ and $u|_{\partial\Omega} = v|_{\partial\Omega}$, so u is indeed the sought solution to the differential inclusion (9.8). □

Now we come to Gromov's original convex integration principle. For an open and bounded set $G \subset \mathbb{R}^{m \times d}$ define the **lamination-convex hull** just like for a compact set, that is,

$$G^{\mathrm{lc}} := \bigcup_{i=0}^{\infty} G^{\mathrm{lc},i},$$

where

$$G^{\mathrm{lc},0} := G,$$
$$G^{\mathrm{lc},i+1} := \left\{ \theta A + (1-\theta)B \;:\; A, B \in G^{\mathrm{lc},i},\ \operatorname{rank}(A-B) \le 1,\ \theta \in [0,1] \right\}.$$

Let $K \subset \mathbb{R}^{m \times d}$ be compact and non-empty. Then, a sequence $(G_k)_{k \in \mathbb{N}}$ of open, uniformly bounded sets $G_k \subset \mathbb{R}^{m \times d}$ is called an **in-approximation** for K if the following two conditions hold:

(i) $G_k \subset G_{k+1}^{\mathrm{lc}}$ for all $k \in \mathbb{N}$;
(ii) if $A_k \in G_k$ and $A_k \to A$ as $k \to \infty$, then $A \in K$.

Theorem 9.9 **(Gromov 1986 [145]).** *Let $K \subset \mathbb{R}^{m \times d}$ be a non-empty compact set and assume that there exists an in-approximation $(G_k)_{k \in \mathbb{N}}$ for K. Suppose that either $v \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ is piecewise affine or $v \in \mathrm{C}^1(\Omega; \mathbb{R}^m)$, and that*

$$\nabla v \in G_\infty := \bigcup_{l=1}^{\infty} G_l \quad \text{a.e.}$$

Then, for every $\varepsilon > 0$ there exist a solution to the exact differential inclusion

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m), & u|_{\partial\Omega} = v|_{\partial\Omega}, \\ \nabla u \in K & \text{a.e.} \end{cases}$$

with $\|u - v\|_{\mathrm{L}^\infty} \le \varepsilon$.

In order to apply the abstract Theorem 9.8, we need to show that, under the assumptions of Gromov's theorem, G_∞ can be piecewise affinely reduced to K. This will be accomplished by successive laminations.

Lemma 9.10. *Let $A, B \in \mathbb{R}^{m \times d}$ with $\operatorname{rank}(A-B) = 1$, $\theta \in (0,1)$ and let $D \subset \mathbb{R}^d$ be a bounded Lipschitz domain. Then, for every $\varepsilon > 0$ there exists a piecewise affine map $u \colon D \to \mathbb{R}^m$ with $u|_{\partial D} = Fx$, where $F := \theta A + (1-\theta)B$, and*

$$\sup_{x \in D} \left[|u(x) - Fx| + \operatorname{dist}(\nabla u(x), \{A, B\}) \right] \le \varepsilon. \tag{9.10}$$

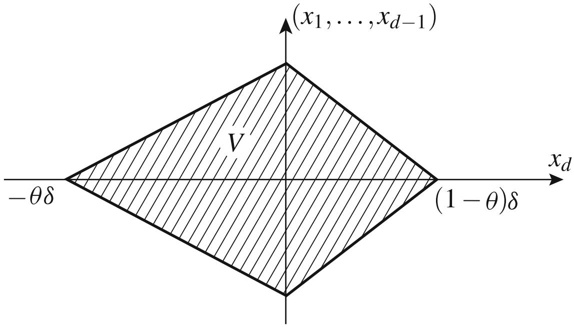


Fig. 9.4 The construction of the domain V

Proof. Let $\varepsilon > 0$ be given.

Step 1. By a change of variables (which only results in an affine transformation of the resulting u), we may assume that

$$A = -(1-\theta)a \otimes \mathrm{e}_d, \qquad B = \theta a \otimes \mathrm{e}_d, \qquad F = 0$$

for some $a \in \mathbb{R}^m \setminus \{0\}$. We will first construct a special open and convex domain $V \subset \mathbb{R}^d$ such that (9.10) holds for $D = V$. In the auxiliary domain

$$U := (-1, 1)^{d-1} \times (-\theta\delta, (1-\theta)\delta),$$

where $\delta > 0$ will be specified below, we define the functions $g, h \colon U \to \mathbb{R}$ by

$$g(x) := -\theta(1-\theta)\delta + \begin{cases} -(1-\theta)x_d & \text{if } x_d \le 0, \\ \theta x_d & \text{if } x_d > 0. \end{cases},$$

$$h(x) := \theta(1-\theta)\delta \sum_{i=1}^{d-1} |x_i|.$$

Then we let

$$V := \bigl\{ x \in U \ : \ g(x) + h(x) < 0 \bigr\},$$

which has the shape of a kite, see Figure 9.4. We will then show that the map

$$v(x) := (g(x) + h(x))a, \qquad x \in V,$$

has the desired properties (in V). Indeed, by construction it is clear that v is piecewise affine and vanishes on ∂V. Moreover, as $\nabla g \in \{-(1-\theta)\mathrm{e}_d, \theta\mathrm{e}_d\}$ in U, we can estimate

$$\sup_{x \in V} \operatorname{dist}(\nabla v(x), \{A, B\}) \le C\delta,$$

where $C > 0$. Thus, choosing $\delta > 0$ suitably, we can achieve (9.10) in $D = V$.

Step 2. From Vitali's Covering Theorem A.15 we find a covering

$$D = Z \cup \bigcup_{k=1}^{\infty} V(a_k, r_k),$$

where $V(a_k, r_k) = a_k + r_k V \subset D$ for $a_k \in D$, $0 < r_k \le \varepsilon/(1 + \|v\|_{L^\infty})$, and $|Z| = 0$. Then set

$$u(x) := r_k v\left(\frac{x - a_k}{r_k}\right) \qquad \text{if } x \in V(a_k, r_k)\ (k \in \mathbb{N}),$$

and observe

$$\nabla u(x) = \nabla v\left(\frac{x - a_k}{r_k}\right) \qquad \text{if } x \in V(a_k, r_k)\ (k \in \mathbb{N}).$$

Thus, u is piecewise affine, satisfies $u|_{\partial D} = 0$, and (9.10) holds. □

Lemma 9.11. *Let $G \subset \mathbb{R}^{m \times d}$ be open and bounded and let $v \in W^{1,\infty}(\Omega; \mathbb{R}^m)$ be piecewise affine with*

$$\nabla v \in G^{\mathrm{lc}} \quad \text{a.e.}$$

Then, for every $\varepsilon > 0$ there exists a piecewise affine map $u \in W^{1,\infty}(\Omega; \mathbb{R}^m)$ with

$$\nabla u \in G \quad \text{a.e.,} \qquad u|_{\partial\Omega} = v|_{\partial\Omega},$$

and $\|u - v\|_{L^\infty} \le \varepsilon$.

Proof. We assume that v is affine; otherwise we apply the following argument separately in every set of the decomposition of Ω, with respect to which v is piecewise affine. Moreover, we may add a constant to v to normalize to the situation where $v(x) = Fx$ for some $F \in \mathbb{R}^{m \times d}$.

By assumption, $F \in G^{\mathrm{lc}}$ and so, $F \in G^{\mathrm{lc},i}$ for some $i \in \mathbb{N} \cup \{0\}$. We proceed by induction over i. For $i = 0$ there is nothing to show. If $F \in G^{\mathrm{lc},i+1} \setminus G^{\mathrm{lc},i}$, then there are $A, B \in G^{\mathrm{lc},i}$ with $\operatorname{rank}(A - B) = 1$ and $\theta \in (0, 1)$ such that $F = \theta A + (1 - \theta)B$. By Lemma 9.10 there exists a piecewise affine map $w\colon \Omega \to \mathbb{R}^m$ with $w|_{\partial\Omega} = Fx = v|_{\partial\Omega}$ and

$$\|w - v\|_{L^\infty} < \frac{\varepsilon}{2}, \qquad \sup_{x \in \Omega} \operatorname{dist}(\nabla w(x), \{A, B\}) < \frac{\varepsilon}{2}.$$

Since $G^{\mathrm{lc},i}$ is open (which is easy to see), we may further assume that (potentially reducing ε in the previous conditions)

$$\nabla w \in G^{\mathrm{lc},i} \quad \text{a.e.}$$

Then, if $\Omega = Z \cup \bigcup_{k=1}^{\infty} D_k$, where $|Z| = 0$, is the decomposition of Ω with respect to which w is piecewise affine, we apply the induction hypothesis in every D_k to get piecewise affine maps $u_k \in \mathrm{W}^{1,\infty}(D_k; \mathbb{R}^m)$ with $u_k|_{\partial D_k} = w|_{\partial D_k}$,

$$\|u_k - w\|_{\mathrm{L}^\infty} < \frac{\varepsilon}{2} \qquad \text{and} \qquad \nabla u_k \in G \quad \text{a.e.}$$

Setting

$$u(x) := u_k(x) \qquad \text{if } x \in D_k \ (k \in \mathbb{N}),$$

we see that $u|_{\partial\Omega} = v|_{\partial\Omega}$, $\|u - v\|_{\mathrm{L}^\infty} \leq \varepsilon$, and $\nabla u \in G$ almost everywhere. We thus conclude the proof of the inductive step and hence the proof of the lemma. □

Proof of Theorem 9.9. By Theorem 9.8 we only need to show that $G_\infty = \bigcup_{l=1}^{\infty} G_l$ can be piecewise affinely reduced to K. Without loss of generality we assume $A \in G_1 \subset G_2^{\mathrm{lc}}$, where the inclusion holds by property (i) from the definition of an in-approximation. The proof for $A \in G_l$ is analogous.

Let $n \in \mathbb{N}$. Lemma 9.11 allows us to construct a piecewise affine map $\psi_2^{(n)} \in \mathrm{W}_0^{1,\infty}(B(0,1); \mathbb{R}^m)$ with

$$A + \nabla\psi_2^{(n)} \in G_2 \subset G_3^{\mathrm{lc}} \quad \text{a.e.} \qquad \text{and} \qquad \|\psi_2^{(n)}\|_{\mathrm{L}^\infty} \leq \frac{1}{2^{2-1}n}.$$

Since $\psi_2^{(n)}$ is piecewise affine, arguing separately in every piece, we may find $\psi_3^{(n)} \in \mathrm{W}_0^{1,\infty}(B(0,1); \mathbb{R}^m)$ with

$$A + \nabla\psi_2^{(n)} + \nabla\psi_3^{(n)} \in G_3 \subset G_4^{\mathrm{lc}} \quad \text{a.e.} \qquad \text{and} \qquad \|\psi_3^{(n)}\|_{\mathrm{L}^\infty} \leq \frac{1}{2^{3-1}n}.$$

Continuing this procedure, we construct a sequence $(\psi_k^{(n)})_k \subset \mathrm{W}^{1,\infty}(B(0,1); \mathbb{R}^m)$ of piecewise affine maps such that

$$A + \sum_{l=2}^{k} \nabla\psi_l^{(n)} \in G_k \subset G_{k+1}^{\mathrm{lc}} \quad \text{a.e.} \qquad \text{and} \qquad \|\psi_k^{(n)}\|_{\mathrm{L}^\infty} \leq \frac{1}{2^{k-1}n}.$$

For the piecewise affine map

$$\psi_j := \sum_{l=2}^{j} \psi_l^{(j)} \in \mathrm{W}^{1,\infty}(B(0,1); \mathbb{R}^m),$$

we see that

$$A + \nabla\psi_j \in G_j \subset G_\infty \quad \text{a.e.} \qquad \text{and} \qquad \|\psi_j\|_{\mathrm{L}^\infty} \leq \sum_{k=2}^{j} \frac{1}{2^{k-1}j} \leq \frac{1}{j}.$$

It only remains to observe that

$$\int_{B(0,1)} \operatorname{dist}(A + \nabla\psi_j(x), K)\,dx \to 0 \qquad \text{as } j \to \infty.$$

This follows immediately since the integrand is uniformly bounded (as the G_k are uniformly bounded) and $A + \nabla\psi_j(x)$ converges pointwise almost everywhere to an element of K by property (ii) from the definition of an in-approximation. Thus, we have shown that G_1 can be affinely reduced to K and Theorem 9.8 yields the conclusion. □

We now apply Gromov's convex integration principle to the two-well problem in two dimensions, see (9.7). We only consider the case $\det U > 1$ (see [206] for the case $\det U = 1$). That is, we will find exact solutions to the differential inclusion

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^2), \quad u|_{\partial\Omega} = Fx, \quad 0 < \alpha < 1 < \beta, \quad \alpha\beta > 1, \\ \nabla u \in K := \mathrm{SO}(2) \cup \mathrm{SO}(2)U, \quad U = \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}. \end{cases} \tag{9.11}$$

Theorem 9.12 **(Müller–Šverák 1993/1996 [205, 254]).** *Let*

$$F \in \operatorname{int} K^{\mathrm{lc}} = \left\{ A = (y,z) \in K^{**} \;:\; |y| < \tfrac{\det U - \det A}{\det U - 1} \text{ and } |z| < \tfrac{\det A - 1}{\det U - 1} \right\}. \tag{9.12}$$

Then, there exists an exact solution to (9.11).

We remark that similar results have also been established by Dacorogna and Marcellini [79], see the notes at the end of this chapter.

Proof. Step 1. We first justify the formula for $\operatorname{int} K^{\mathrm{lc}}$. From Theorem 9.5 (iii) we know that

$$K^{\mathrm{lc}} = \left\{ A = (y,z) \in K^{**} \;:\; |y| \le \tfrac{\det U - \det A}{\det U - 1} \text{ and } |z| \le \tfrac{\det A - 1}{\det U - 1} \right\}.$$

By Problem 9.5, the conformal coordinate map $(y,z) \mapsto A$ is a diffeomorphism and hence maps boundaries to boundaries. This immediately implies the formula (9.12).

Step 2. Define for any $J, V \in \mathbb{R}^{2\times 2}$ with $\det J, \det V > 1$ the set

$$L(J,V) := \left\{ A = (y,z) \in K^{**} \;:\; |y| \le \tfrac{\det V - \det A}{\det V - \det J} \text{ and } |z| \le \tfrac{\det A - \det J}{\det V - \det J} \right\}.$$

Note that $L(\mathrm{Id}, U) = K^{\mathrm{lc}}$. The following continuity property holds: Let $(J_j, V_j) \subset \mathbb{R}^{2\times 2} \times \mathbb{R}^{2\times 2}$ be a converging sequence such that $(J_j, V_j) \to (J, V)$ as well as

$$0 < \det J_j < \det V_j \qquad \text{and} \qquad \lambda_*(V_j J_j^{-1}) < 1 < \lambda^*(V_j J_j^{-1}),$$

where $\lambda_*(V_j J_j^{-1})$, $\lambda^*(V_j J_j^{-1})$ are the smaller and the larger eigenvalue of $V_j J_j^{-1}$, respectively. Let $P \subset \operatorname{int} L(J, V)$ be compact. Then, $P \subset L(J_j, V_j)$ for all sufficiently large j. Indeed, this follows by Step 1 and the fact that the conformal coordinates map is a diffeomorphism.

Step 3. We now prove the assertion of the theorem, for which we construct an in-approximation, starting with a neighborhood $G_1 \Subset K^{\mathrm{lc}}$ of $F \in \operatorname{int} K^{\mathrm{lc}}$. Assume that we have already constructed $G_k \Subset K^{\mathrm{lc}}$. We will show that we can find $G_{k+1} \Subset K^{\mathrm{lc}}$ with $G_k \subset G_{k+1}^{\mathrm{lc}}$ and such that $\sup_{H \in G_{k+1}} \operatorname{dist}(H, K) \leq 1/k$. Indeed, since $G_k \Subset K^{\mathrm{lc}} = L(\mathrm{Id}, U)$, by the continuity property from the previous step we may find $J_{k+1}, U_{k+1} \in \operatorname{int} K^{\mathrm{lc}}$ such that

$$|J_{k+1} - \mathrm{Id}|, \ |U_{k+1} - U| \leq \frac{1}{2k}$$

and

$$\overline{G_k} \subset L(J_{k+1}, U_{k+1}) = \big(\mathrm{SO}(2)J_{k+1} \cup \mathrm{SO}(2)U_{k+1}\big)^{\mathrm{lc}}.$$

Here, the last equality follows from Theorem 9.5 (iii) and we note that $\mathrm{Id}, U \in \partial K^{\mathrm{lc}}$.

We have that

$$\mathrm{SO}(2)J_{k+1} \cup \mathrm{SO}(2)U_{k+1} \subset \operatorname{int} K^{\mathrm{lc}}$$

since $J_{k+1}, U_{k+1} \in \operatorname{int} K^{\mathrm{lc}} = L(\mathrm{Id}, U)$ and the expression (9.12) for K^{lc} shows that this set is SO(2)-invariant. Now take G_{k+1} to be an (at most) $1/(2k)$-neighborhood of $\mathrm{SO}(2)J_{k+1} \cup \mathrm{SO}(2)U_{k+1}$ that is compactly contained in $\operatorname{int} K^{\mathrm{lc}}$, i.e., $G_{k+1} \Subset K^{\mathrm{lc}}$. Then, $G_k \subset G_{k+1}^{\mathrm{lc}}$ and for all $H \in G_{k+1}$ we have

$$\operatorname{dist}(H, K) \leq \frac{1}{2k} + \operatorname{dist}\big(\mathrm{SO}(2)J_{k+1} \cup \mathrm{SO}(2)U_{k+1}, K\big) \leq \frac{1}{k}.$$

This completes the induction.

The claim of the theorem then follows from Gromov's Convex Integration Theorem 9.9. □

We close this section by quoting the following result, which shows that exact solutions to (9.11) are *necessarily* very complicated. For the statement of the theorem we need a notion of boundary regularity: A set $E \subset \Omega$ is called a **set of finite perimeter in** Ω if the distributional derivative of the indicator function $\mathbb{1}_E$ in Ω is a finite Borel measure, i.e.,

$$\operatorname{Per}_\Omega(E) := \sup \left\{ \int_E \operatorname{div} \varphi \, \mathrm{d}x \ : \ \varphi \in \mathrm{C}_c^1(\Omega; \mathbb{R}^d), \ \|\varphi\|_{\mathrm{L}^\infty} \leq 1 \right\} < \infty. \quad (9.13)$$

For sets E with smooth boundary we have $\operatorname{Per}_\Omega(E) = \mathscr{H}^{d-1}(E \cap \Omega)$, which follows from the divergence theorem.

Theorem 9.13 **(Dolzmann–Müller 1995 [103]).** *Let* $u \in \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^d)$ *satisfy*

$$\nabla u \in K := \mathrm{SO}(d) \cup \mathrm{SO}(d)U$$

such that $\det U > 0$ *and every matrix in* K *is rank-one connected to precisely two other matrices in* K. *Assume furthermore that*

$$E := \{\, x \in \Omega \;:\; \nabla u(x) \in \mathrm{SO}(d) \,\}$$

has finite perimeter in Ω. *Then,* u *is locally a simple laminate and if* u *is affine on the boundary* $\partial\Omega$*, then* u *is an affine map.*

9.4 Infinite-Order Laminates

In this section we quote an extension of the Gromov Convex Integration Theorem 9.9 that also works in situations where infinite-order laminates are necessary. For this, let again $K \subset \mathbb{R}^{m\times d}$ be compact and non-empty (closed sets are also possible if property (ii) below holds in a stronger form). Then, a sequence $(G_k)_{k\in\mathbb{N}}$ of open, uniformly bounded sets $G_k \subset \mathbb{R}^{m\times d}$ is called an **RC-in-approximation** for K if the following two conditions hold:

(i) $G_k \subset G_{k+1}^{\mathrm{rc}} := \bigcup\{\, S^{\mathrm{rc}} \;:\; S \subset G_{k+1} \text{ compact} \,\}$ for all $k \in N$;
(ii) if $A_k \in G_k$ and $A_k \to A$ as $k \to \infty$, then $A \in K$.

Theorem 9.14 **(Müller–Šverák 2003 [207]).** *Let* $K \subset \mathbb{R}^{m\times d}$ *be a non-empty compact set and assume that there exists an RC-in-approximation* $(G_k)_{k\in\mathbb{N}}$ *for* K. *Suppose that either* $v \in \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^m)$ *is piecewise affine or* $v \in \mathrm{C}^1(\Omega;\mathbb{R}^m)$*, and that*

$$\nabla v \in G_\infty := \bigcup_{l=1}^{\infty} G_l \quad \textit{a.e.}$$

Then, for every $\varepsilon > 0$ *there exists a solution to the exact differential inclusion*

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^m), \quad u|_{\partial\Omega} = v|_{\partial\Omega}, \\ \nabla u \in K \quad \textit{a.e.} \end{cases}$$

with $\|u - v\|_{\mathrm{L}^\infty} \le \varepsilon$.

The proof follows essentially the same strategy as Theorem 9.9, but there is an additional step to approximate infinite-order laminates with finite-order laminates, without changing the barycenter; see [207] for details.

Recall that the Evans Partial Regularity Theorem 5.22 established partial regularity for minimizers: Let $f\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ be a smooth, strongly quasiconvex integrand such that with some $M > 0$ it holds that

$$D^2 f(A)[B, B] \leq M|B|^2, \qquad A, B \in \mathbb{R}^{m \times d}. \tag{9.14}$$

Then, for every solution u of

$$\begin{cases} \text{Minimize} & \mathscr{F}[u] := \int_\Omega f(\nabla u(x)) \, dx \\ \text{over all} & u \in W^{1,2}(\Omega; \mathbb{R}^m) \text{ with } u|_{\partial\Omega} = g \in W^{1/2,2}(\partial\Omega; \mathbb{R}^m), \end{cases}$$

there exists a relatively closed singular set $\Sigma_u \subset \Omega$ with $|\Sigma_u| = 0$ such that $u \in C^{1,\alpha}_{\mathrm{loc}}(\Omega \setminus \Sigma_u)$ for all $\alpha \in (0, 1)$.

It turns out that for weak solutions to the Euler–Lagrange equation that are not minimizers, the corresponding statement is *false*. In fact, there are many such pathological weak solutions. This statement can be proved using convex integration:

Theorem 9.15 (Müller–Šverák 2003 [207]). *Let $\Omega \subset \mathbb{R}^2$ be a bounded Lipschitz domain. There exists a smooth and strongly quasiconvex integrand $f_{\mathrm{MS}}\colon \mathbb{R}^{2\times 2} \to \mathbb{R}$ satisfying (9.14) such that the following property holds: For every $v \in C^1(\Omega; \mathbb{R}^2)$ and every $\varepsilon > 0$ there exists a map $u \in W^{1,\infty}(\Omega; \mathbb{R}^2)$ solving*

$$-\operatorname{div}[D f_{\mathrm{MS}}(\nabla u)] = 0 \quad \text{in } \Omega \text{ (weakly)}$$

and $\|u - v\|_{L^\infty} \leq \varepsilon$, but $u \notin C^1(U; \mathbb{R}^2)$ on any open subset $U \subset \Omega$.

For the proof (which makes use of generalized T_4-configurations) see [207]. There is also a (strongly) polyconvex integrand with the same property as shown in [264].

9.5 Crystalline Microstructure in 3D

The construction of exact solutions to multi-well differential inclusions in three space dimensions is an active but largely unfinished area of research. We only quote the following result.

Theorem 9.16 (Conti–Dolzmann–Kirchheim 2007 [69]). *Let*

$$K := \bigcup_{i=1}^{6} \mathrm{SO}(3) U_i,$$

where

$$U_1 = \begin{pmatrix} \alpha_1 & & \\ & \alpha_2 & \\ & & \alpha_3 \end{pmatrix}, \qquad 0 < \alpha_1 < \alpha_2 \leq \alpha_3, \qquad \alpha_1\alpha_2\alpha_3 = 1,$$

and the other matrices $U_2, \dots, U_6$ are given by permuting the three values on the diagonal (if $\alpha_2 = \alpha_3$, the six-well problem reduces to a three-well problem). Then, *there exists an* $R > 0$ *such that for all* $v \in \mathrm{C}^{1,\gamma}(\Omega;\mathbb{R}^3)$ *for some* $\gamma \in (0,1)$ *that satisfy*

$$|\nabla v(x) - \mathrm{Id}| < R, \quad \det \nabla v(x) = 1 \qquad \text{for all } x \in \Omega,$$

the differential inclusion

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^3), \quad u|_{\partial\Omega} = v|_{\partial\Omega}, \\ \nabla u \in K \end{cases}$$

has at least one exact solution. Moreover, for any given $\varepsilon > 0$ *this solution can be chosen such that* $\|u - v\|_{\mathrm{L}^\infty} \le \varepsilon$.

Example 9.17. With regard to the shape-memory effect observed in the NiAl alloy, which is described in Section 1.8, we have discussed in Example 8.10 that the material will try to satisfy the differential inclusion

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^3), \ \ \Omega \subset \mathbb{R}^3, \\ \nabla u \in K \quad \text{in } \Omega \end{cases}$$

as closely as possible, where K is the pointwise minimizer set of the integrand in the governing energy functional. We have

$$K = \begin{cases} \mathrm{SO}(3) & \text{above the critical temperature,} \\ \mathrm{SO}(3)U_1 \cup \mathrm{SO}(3)U_2 \cup \mathrm{SO}(3)U_3 & \text{below the critical temperature,} \end{cases}$$

and the matrices $U_1, \ldots, U_3$ are given explicitly in Section 1.8. This *symmetry breaking* from the cubic to the tetragonal phase explains the shape-memory effect: Above the critical temperature, the microstructure is rigid (see Reshetnyak's Rigidity Theorem 8.20) and so, plastic deformations are reflected in changes in the crystal lattice. Once the material is cooled below the critical temperature, however, K turns out to have many rank-one connections, K^{lc} is large (but it is often unknown *how* large, see Problem 16 in [28]), and the material can deform in some ways without changing the crystalline structure. By a convex integration result like the one from Theorem 9.16 above (we remarked in Section 1.8 that $\det U_1 = \det U_2 = \det U_3 \approx 1$, so the above theorem is *almost* applicable) one could then show that for given (suitably constrained) boundary values many exact solutions to the above differential inclusion exist. This is the mathematical manifestation of this flexibility. For the cubic-to-orthorhombic phase transition (a six-well problem with *non-diagonal* $U_1, \ldots, U_6$) of CuAlNi there is currently no applicable convex integration theorem.

We finally remark that it is currently unclear whether the convex integration solutions are *admissible* in real-world problems from material science. This admissibility would entail that these exact solutions are in fact limits of approximate solutions, for which the phases have finite perimeter. In view of Theorem 9.13, this is a non-trivial question. We refer to [28] for open problems in this area.

9.6 Stability of Gradient Distributions

The theory of convex integration as presented in Section 9.3 is based on fairly concrete constructions. One iteratively sums up very fast oscillations with small amplitudes and shows that this yields a strongly converging sequence, whose limit has the desired properties. There is also another, more abstract, approach to solving differential inclusions, which is based on Baire category theory. This idea was first explored in this context by Dacorogna and Marcellini [79, 80] and then put into its final form by Kirchheim, who introduced the notion of "stability" for gradient distributions.

Let $K \subset \mathbb{R}^{m\times d}$ be compact and non-empty. As before, we aim to find exact solutions to

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^m), \quad u|_{\partial\Omega} = g, \\ \nabla u \in K, \end{cases} \tag{9.15}$$

where $g \in \mathrm{W}^{1,1/2}(\partial\Omega;\mathbb{R}^m)$.

We say that K is **(piecewise affinely) stable** with respect to a bounded open set $G \subset \mathbb{R}^{m\times d}$ if for all $\eta > 0$ there exists a $\delta_\eta > 0$ such that for all $A \in G$ with $\operatorname{dist}(A, K) > \eta$ there is a (countably) piecewise affine map $\psi \in \mathrm{W}^{1,\infty}_0(B(0,1);\mathbb{R}^m)$ such that

(i) $A + \nabla\psi \in G$ for almost every $x \in B(0,1)$;

(ii) $\displaystyle⨍_{B(0,1)} |\nabla\psi(x)|\,\mathrm{d}x > \delta_\eta$.

From a standard covering argument we realize that in the above definition the unit ball can be replaced by any bounded Lipschitz domain. The main difference to the piecewise affine reduction property from Section 9.3 is that here one does not need to show that it is possible to push the gradient arbitrarily close to K.

We define the set

$$X_0 := \left\{ u \in \mathrm{W}^{1,2}_g(\Omega;\mathbb{R}^m) \;:\; u \text{ piecewise affine and } \nabla u \in G \text{ a.e.} \right\}$$

and assume that it is non-empty (which is a condition on G). Let furthermore

$$X := \text{w-clos}_{\mathrm{W}^{1,2}} X_0, \tag{9.16}$$

that is, X is the closure of X_0 with respect to the weak topology in $\mathrm{W}^{1,2}(\Omega;\mathbb{R}^m)$. The set X so defined together with the weak $\mathrm{W}^{1,2}$-topology is a complete metric space since the weak topology in $\mathrm{W}^{1,2}$ is metrizable on norm-bounded sets (G is bounded).

Then, a very general "convex integration" principle reads as follows:

Theorem 9.18 (Kirchheim 2003 [160]). *Assume that K is (piecewise affinely) stable with respect to G. Then, the set of exact solutions of (9.15) is dense in X.*

In order to establish this theorem, we first recall the following fundamental result from functional analysis, which we prove for the sake of completeness.

Theorem 9.19 **(Baire category theorem).** *Let X be a complete metric space.*

(i) If the sets $U_k \subset X$, $k \in \mathbb{N}$, are open and dense in X, then $\bigcap_{k=1}^\infty U_k$ is also dense in X.

(ii) If $f : X \to \mathbb{R}$ is a ***Baire-one function****, that is, f is the pointwise limit of continuous functions, then the set of continuity points of f is dense in X.*

Proof. Ad (i). Let $V \subset X$ be open. We need to show that there is a $u \in V$ with $u \in \bigcap_{k=1}^\infty U_k$. Since U_1 is dense in X it must intersect V. So, there is a ball

$$\overline{B(u_1, r_1)} \subset V \cap U_1 \qquad \text{for some } u_1 \in U_1 \text{ and } r_1 > 0.$$

We iterate this procedure to get $u_j \in X$ and $r_j \in (0, 1/j)$ such that

$$\overline{B(u_j, r_j)} \subset B(u_{j-1}, r_{j-1}) \cap U_j.$$

Since $u_j \in B(u_i, r_i)$ if $j > i$, the sequence (u_j) is Cauchy and thus converges to some $u \in X$. We have that $u \in B(u_i, r_i) \cap U_i$ for all i and thus $u \in V \cap \bigcap_{k=1}^\infty U_k$.

Ad (ii). Define for the function f and $u \in X$ the quantity

$$\omega(u) := \lim_{\delta \downarrow 0} \left[\sup_{v \in B(u,\delta)} f(v) - \inf_{v \in B(u,\delta)} f(v) \right].$$

Let $\varepsilon > 0$ and assume that there is an open set $V \subset X$ such that $\omega(u) > 4\varepsilon$ for all $u \in \overline{V}$. In the remainder of the proof we will lead this to a contradiction. Thus, ω vanishes on a dense set in X, which is the set of points where f is continuous.

Suppose that $f_j \to f$ pointwise in X and define

$$E_n := \bigcap_{i,j \geq n} \left\{ u \in X \,:\, |f_i(u) - f_j(u)| \leq \varepsilon \right\}, \qquad n \in \mathbb{N},$$

which are closed, increasing sets with $\bigcup_n E_n = X$ since $f_j \to f$ pointwise. Thus,

$$\bigcap_{n=1}^\infty (\overline{V} \setminus E_n) = \emptyset.$$

By (i) not all of the relatively open sets $\overline{V} \setminus E_n$ can be dense in the complete metric space $\overline{V}$. Hence, there exists an $n \in \mathbb{N}$ such that $V \cap E_n$ contains a non-empty open set W. In particular, for all $w \in W$ (letting $i \to \infty$ and $j := n$),

$$|f(w) - f_n(w)| \leq \varepsilon.$$

Moreover, for any $u_0 \in W$ choose a ball $B(u_0, r) \subset W$ with the property that

$$|f_n(w) - f_n(u_0)| \leq \varepsilon \qquad \text{for all } w \in B(u_0, r).$$

This is possible by the continuity of f_n. Then, for all $u, v \in B(u_0, r)$,

$$\begin{aligned}|f(u) - f(v)| &\le |f(u) - f_n(u)| + |f_n(u) - f_n(u_0)| + |f_n(u_0) - f_n(v)| \\ &\quad + |f_n(v) - f(v)| \\ &\le 4\varepsilon.\end{aligned}$$

Thus, $\omega(u_0) \le 4\varepsilon$ for all $u_0 \in W \subset V$. This is the sought contradiction. □

Returning to X as defined in (9.16), we denote by $S \subset X$ the set of **(weak-to-strong) stability points**, i.e., those $u \in X$ such that for any sequence $(u_j) \subset X$ with $u_j \rightharpoonup u$ in $\mathrm{W}^{1,2}$ it automatically holds that also $u_j \to u$ strongly in $\mathrm{W}^{1,2}$. A priori, this appears to be a very strong requirement and it is not clear whether this set is non-empty. However, we have the following result:

Lemma 9.20. *The set S is dense in X.*

Proof. Let $(\eta_\delta)_{\delta>0} \subset \mathrm{C}_c^\infty(\mathbb{R}^d)$ be a family of mollifiers. We set

$$\mathscr{G}[u] := \int_\Omega |\nabla u|^2 \,\mathrm{d}x, \qquad u \in \mathrm{W}_g^{1,2}(\Omega; \mathbb{R}^m),$$

and

$$\mathscr{G}_n[u] := \int_\Omega |\eta_{1/n} \star \nabla u|^2 \,\mathrm{d}x, \qquad u \in \mathrm{W}_g^{1,2}(\Omega; \mathbb{R}^m),$$

where we consider u to be extended by some fixed extension of g outside of Ω. Then, all the $\mathscr{G}_n$ are continuous on X with respect to the weak convergence in $\mathrm{W}^{1,2}$. Indeed, let $(u_j) \subset X$ with $u_j \rightharpoonup u$ in $\mathrm{W}^{1,2}$. Then, for all $x \in \Omega$,

$$\begin{aligned}(\eta_{1/n} \star \nabla u_j)(x) &= \int \eta_{1/n}(x - y)\nabla u_j(y) \,\mathrm{d}y \\ &\to \int \eta_{1/n}(x - y)\nabla u(y) \,\mathrm{d}y = (\eta_{1/n} \star \nabla u)(x)\end{aligned}$$

and, by Young's inequality for convolutions, see Lemma A.32,

$$\|\eta_{1/n} \star \nabla u_j\|_{\mathrm{L}^4} \le \|\eta_{1/n}\|_{\mathrm{L}^{4/3}} \cdot \|\nabla u_j\|_{\mathrm{L}^2}.$$

So, for fixed n, the family $\{\eta_{1/n} \star \nabla u_j\}_j$ is L^2-equiintegrable and via Vitali's Convergence Theorem A.11 we may conclude that $\eta_{1/n} \star \nabla u_j \to \eta_{1/n} \star \nabla u$ in L^2 as $j \to \infty$ (with n held fixed). In particular, $\mathscr{G}_n[u_j] \to \mathscr{G}_n[u]$.

Next, observe that for all $u \in \mathrm{W}^{1,2}(\Omega; \mathbb{R}^m)$ we have

$$\mathscr{G}_n[u] \to \mathscr{G}[u] \qquad \text{as } n \to \infty.$$

This means that $\mathscr{G}$ is the pointwise limit of X-continuous functionals, i.e., a Baire-one functional. From the Baire Category Theorem 9.19 (ii) it follows that the set of continuity points of $\mathscr{G}$ is dense in X. For such a continuity point u we have that

$u_j \rightharpoonup u$ in $\mathrm{W}^{1,2}$ implies $\|\nabla u_j\|_{\mathrm{L}^2} \to \|\nabla u\|_{\mathrm{L}^2}$, whereby $u_j \to u$ strongly in $\mathrm{W}^{1,2}$ (see the Radon–Riesz Theorem A.14). Thus, the set of continuity points of $\mathscr{G}$ is contained in S (in fact, these two sets are the same) and the claim follows. □

Proof of Theorem 9.18. *Step 1.* Let $u \in S$. We will show that $\nabla u \in K$ almost everywhere, which will imply the claim by Lemma 9.20. Assume to the contrary that

$$\mathscr{K}[u] := \int_\Omega \operatorname{dist}(\nabla u(x), K) \,\mathrm{d}x > 0$$

and take a sequence $(v_j) \subset X_0$ with $v_j \rightharpoonup u$ in $\mathrm{W}^{1,2}$. Then, since $u \in S$, it follows that $v_j \to u$ in $\mathrm{W}^{1,2}$. In particular, we may assume

$$\inf_{j\in\mathbb{N}} \mathscr{K}[v_j] > 0.$$

We will show below that for every v_j we can find a sequence $(v_{j,k})_k \subset X_0$ with $v_{j,k} \rightharpoonup v_j$ in $\mathrm{W}^{1,2}$ as $k \to \infty$ and

$$\|\nabla v_j - \nabla v_{j,k}\|_{\mathrm{L}^1} \geq \beta > 0,$$

where β is independent of j and k. Since all gradients $\nabla v_{j,k}$ are uniformly L^2-norm-bounded (because G is bounded), we may use the metrizability of the weak topology on norm-bounded sets to select a diagonal subsequence $(u_j) \subset X_0$ with $u_j \rightharpoonup u$ in $\mathrm{W}^{1,2}$ and

$$\|\nabla u_j - \nabla u\|_{\mathrm{L}^1} \geq \|\nabla u_j - \nabla v_j\|_{\mathrm{L}^1} - \|\nabla v_j - \nabla u\|_{\mathrm{L}^1} \geq \frac{\beta}{2} > 0$$

for j sufficiently large. However, using $u \in S$ again, we must also have $u_j \to u$ in $\mathrm{W}^{1,2}$, a contradiction.

Step 2. It remains to show that for any $v \in X_0$ with $\mathscr{K}[v] > 0$ and for any given $\varepsilon > 0$ we can find $w \in X_0$ with $\|v - w\|_{\mathrm{L}^2} \leq \varepsilon$ and

$$\|\nabla v - \nabla w\|_{\mathrm{L}^1} \geq \beta,$$

where $\beta > 0$ is a constant that does not depend on $\varepsilon > 0$ and that depends on v only through $\mathscr{K}[v]$.

Since $v \in X_0$ is piecewise affine, we can write

$$v(x) = \sum_{l=1}^{\infty} (z_l + A_l x) \mathbb{1}_{D_l}(x)$$

for some $z_l \in \mathbb{R}^m$, $A_l \in G$, and Lipschitz subdomains $D_l \subset \Omega$ that form a disjoint partition of Ω up to a negligible set. For $\eta > 0$ set

$$E_\eta := \left\{ x \in \Omega \;:\; \operatorname{dist}(\nabla v(x), K) > \eta \right\}$$

and estimate

$$\begin{aligned}\mathscr{K}[v] &= \int_{E_\eta} \operatorname{dist}(\nabla v(x), K)\,\mathrm{d}x + \int_{\Omega \setminus E_\eta} \operatorname{dist}(\nabla v(x), K)\,\mathrm{d}x \\ &\le (|G|_\infty + |K|_\infty) \cdot |E_\eta| + \eta |\Omega|.\end{aligned}$$

Here, $|G|_\infty := \sup\{\,|A| \,:\, A \in G\,\}$ and likewise for $|K|_\infty$. Thus, choosing $\eta := \mathscr{K}[v]/(2|\Omega|) > 0$, we get

$$|E_\eta| \ge \frac{\mathscr{K}[v]}{2(|G|_\infty + |K|_\infty)} =: 2\alpha > 0.$$

Refining the piecewise affine partition for v if necessary, we may assume that

$$\sum_{D_l \subset E_\eta} |D_l| \ge \alpha.$$

Then, use the (piecewise affine) stability of K with respect to G to find in each $D_l \subset E_\eta$ a piecewise affine map $\psi_l \in \mathrm{W}_0^{1,\infty}(D_l; \mathbb{R}^m)$ with

(a) $A_l + \nabla \psi_l(x) \in G$ for almost every $x \in D_l$;
(b) $\displaystyle\int_{D_l} |\nabla \psi_l(x)|\,\mathrm{d}x \ge \delta_\eta |D_l|$, where $\delta_\eta > 0$ is independent of l and ε;
(c) $\displaystyle\|\psi_l\|_{\mathrm{L}^2(D_l)} \le \varepsilon \frac{|D_l|}{|\Omega|}$.

The last condition can be established via a "homogenization" argument as in Lemma 4.14. For $D_l \not\subset E_\eta$ set $\psi_l := 0$. Define

$$w(x) := \sum_{l=1}^{\infty} (z_l + A_l x + \psi_l(x)) \mathbb{1}_{D_l}(x), \qquad x \in \Omega.$$

Then, $\|v - w\|_{\mathrm{L}^2} \le \varepsilon$ and

$$\|\nabla v - \nabla w\|_{\mathrm{L}^1} \ge \sum_{D_l \subset E_\eta} \int_{D_l} |\psi_l(x)|\,\mathrm{d}x \ge \alpha \delta_\eta =: \beta.$$

Since β is independent of ε and depends on v only through $\mathscr{K}[v]$, the claim of the theorem follows. □

Applications of this approach, in particular to *inhomogeneous* differential inclusions, can be found in Chapters 3 and 4 of [160].

The theory of convex integration via Baire category theory has several advantages. First, checking stability can be easier than showing the existence of a piecewise affine reduction or in-approximations as in the previous sections. One only needs to show that away from K one can perturb affine maps in a suitable way. Second, the analysis of the differential inclusion $\nabla u \in K$ can be refined by introducing a suitable notion of extremal points for K, see Chapter 3 of [160] for details. This in fact explains the underlying workings of the whole Baire convex integration strategy: The functional $\mathscr{G}$ can only be continuous at $u \in X$ with respect to the weak convergence in $\mathrm{W}^{1,2}$ if ∇u is already "trapped" in the extreme points of K. Otherwise, the (piecewise affine) stability property would imply that a perturbation is possible, violating the continuity.

Finally, we remark that the approach presented in this section is in fact equivalent to the strategy from Section 9.3, as was observed by Sychev [261, 262].

9.7 Non-laminate Microstructures

All homogeneous Young measures that we have explicitly constructed so far, like the T_4-configuration, have been laminates (perhaps infinite-order ones). So, a natural question is whether there are microstructures that cannot be realized as laminates. By the Kinderlehrer–Pedregal Theorem 7.15 and Pedregal's characterization of laminates in Theorem 9.2, this question is intimately tied to Morrey's Conjecture 7.9. Thanks to Šverák's Example 7.10 of a rank-one convex, but not quasiconvex, function, we can also construct examples of gradient Young measures that are not laminates:

Example 9.21. Let $\varphi \in \mathrm{W}^{1,\infty}_{\mathrm{per}}((0,1)^2;\mathbb{R}^3)$ be the function constructed in Example 7.10. Define the gradient Young measure $\overline{\delta[\nabla\varphi]} \in \mathbf{GY}^\infty((0,1)^2;\mathbb{R}^{3\times 2})$ via the Riemann–Lebesgue Lemma 4.15 (φ has periodic boundary values), i.e.,

$$\big\langle h, \overline{\delta[\nabla\varphi]}\big\rangle = \fint_{(0,1)^2} h(\nabla\varphi(x))\,\mathrm{d}x, \qquad h \in \mathrm{C}_0(\mathbb{R}^{3\times 2}).$$

However, $\overline{\delta[\nabla\varphi]}$ is not a laminate: In Example 7.10 a function $\tilde{h} = h_{\alpha,\beta}\colon \mathbb{R}^{3\times 2} \to \mathbb{R}$ was constructed that is rank-one convex, but not quasiconvex at $[\overline{\delta[\nabla\varphi]}] = 0$. More precisely, we showed in (7.13) that

$$\int_{(0,1)^2} \tilde{h}(\nabla\varphi)\,\mathrm{d}x < 0 = \tilde{h}(0) = \tilde{h}\big([\overline{\delta[\nabla\varphi]}]\big).$$

Then, Pedregal's Theorem 9.2 implies that $\overline{\delta[\nabla\varphi]}$ cannot be a laminate.

Example 9.22 (James). The following is a more concrete example of a gradient Young measure that is not a laminate; it is still based on a variant of Šverák's construction. Let $\varphi\colon \mathbb{R} \to \mathbb{R}$ be the periodic sawtooth function

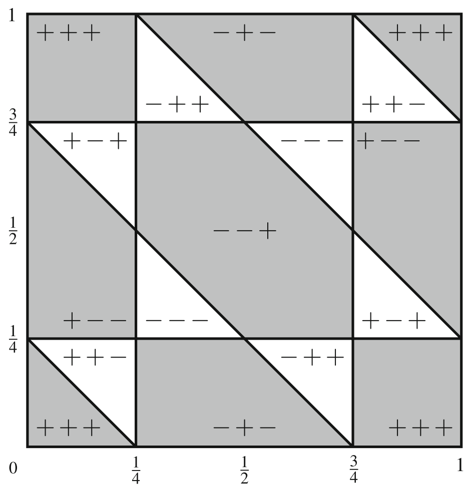


Fig. 9.5 James' example of a gradient Young measure that is not a laminate

$$\varphi(t) := \begin{cases} t & \text{if } t - \lfloor t \rfloor \in [0, 1/4), \\ \frac{1}{2} - t & \text{if } t - \lfloor t \rfloor \in [1/4, 3/4), \\ t - 1 & \text{if } t - \lfloor t \rfloor \in [3/4, 1), \end{cases}$$

and define $u \in \mathrm{W}^{1,\infty}(\mathbb{R}^2; \mathbb{R}^3)$ via

$$u(x_1, x_2) := \begin{pmatrix} \varphi(x_1) \\ \varphi(x_2) \\ \varphi(x_1 + x_2) \end{pmatrix}.$$

Then, for $u_j(x) := u(jx)/j$, where $x \in (0, 1)^2$, the gradients ∇u_j take values in the set

$$K := \left\{ \begin{pmatrix} x & 0 \\ 0 & y \\ z & z \end{pmatrix} \in \mathbb{R}^{3\times 2} \;:\; x, y, z \in \{-1, +1\} \right\}.$$

We refer to the eight elements of K by the signs of x, y, z; for instance, the matrix corresponding to $(x, y, z) = (+1, -1, +1)$ is simply denoted by "$+ - +$". See Figure 9.5 for an illustration of ∇u. We have that $\nabla u_j \overset{\mathbf{Y}}{\to} \nu \in \mathbf{GY}^\infty((0, 1)^2; \mathbb{R}^{3\times 2})$ and a simple counting argument gives

$$\nu = \frac{3}{16}\big[\delta_{+++} + \delta_{+--} + \delta_{-+-} + \delta_{--+}\big] + \frac{1}{16}\big[\delta_{++-} + \delta_{+-+} + \delta_{-++} + \delta_{---}\big].$$

It is also straightforward to compute that $[\nu] = 0$. Notice that the matrices in the first block all have positive *parity*, which for the matrix corresponding to the signs

x, y, z is given as xyz (i.e., the parity is positive if and only if the number of minuses is even). All matrices in the second block have negative parity.

For all $\alpha > 0$ there exists a $\beta > 0$ such that the function from Šverák's Example 7.10,

$$h_{\alpha,\beta}(A) := g(\mathbf{P}(A)) + \alpha\big(|A|^2 + |A|^4\big) + \beta|A - \mathbf{P}(A)|^2,$$

where

$$g\begin{pmatrix} x & 0 \\ 0 & y \\ z & z \end{pmatrix} := -xyz$$

and $\mathbf{P}\colon \mathbb{R}^{3\times 2} \to L$ is a linear projection onto

$$L := \left\{ \begin{pmatrix} x & 0 \\ 0 & y \\ z & z \end{pmatrix} \;:\; x, y, z \in \mathbb{R} \right\},$$

is rank-one convex. Thus, if ν was a laminate, Pedregal's Theorem 9.2 would imply

$$0 = h_{\alpha,\beta}([\nu]) \leq \int h_{\alpha,\beta}(A) \,\mathrm{d}\nu(A).$$

As all matrices in L have the same (Frobenius) norm 2, we then however get

$$\begin{aligned} 0 &\leq \left(\sum_{\substack{A\in L \text{ with} \\ \text{parity } +1}} -\nu(A) \quad + \quad \sum_{\substack{A\in L \text{ with} \\ \text{parity } -1}} \nu(A) \right) + 20\alpha \\ &= \left(-\frac{12}{16} + \frac{4}{16} \right) + 20\alpha \\ &= -\frac{1}{2} + 20\alpha. \end{aligned}$$

Since for α sufficiently small the right-hand side is negative, this is a contradiction. Hence, ν cannot be a laminate.

Unfortunately, while the preceding example shows that there exists a measure $\nu \in \mathscr{M}^{\mathrm{qc}}(K) \setminus \mathscr{M}^{\mathrm{rc}}(K)$ it turns out that nevertheless $K^{\mathrm{qc}} = K^{\mathrm{rc}}$, see Problem 9.7. Thus, it is reasonable to formulate a strong version of Morrey's conjecture:

Conjecture 9.23. There exists a compact set $K \subset \mathbb{R}^{m\times d}$ such that $K^{\mathrm{rc}} \neq K^{\mathrm{qc}}$.

Only some special cases are known. Milton showed that there is a compact set $K \subset \mathbb{R}^{3\times 12}$ with $K^{\mathrm{rc}} \neq K^{\mathrm{qc}}$ based on a modification of James' result, see [187]. There is also an argument of Šverák that shows that there is a compact set $K \subset \mathbb{R}^{6\times 2}$ with $K^{\mathrm{rc}} \neq K^{\mathrm{qc}}$, see Section 4.7 in [203].

9.8 Unbounded Microstructure

In this final section of the chapter we will show how some lamination methods can be extended to unbounded sets. In particular, we will show that the Friesecke–James–Müller Theorem 8.22 and Korn's inequality 8.20 do not hold in L^1.

Theorem 9.24 (Conti–Faraco–Maggi 2005 [71]). *For all $N \in \mathbb{N}$ there exists a map $u \in \mathrm{W}^{1,\infty}(B(0,1); \mathbb{R}^d)$ with $u|_{\partial B(0,1)} = x$ such that*

$$\inf_{Q \in \mathrm{SO}(d)} \int_{B(0,1)} |\nabla u - Q| \,\mathrm{d}x \geq N \int_{B(0,1)} \mathrm{dist}(\nabla u, \mathrm{SO}(d)) \,\mathrm{d}x.$$

The proof is based on the following explicit construction, where by $\mathscr{M}^{\mathrm{lc}}(\mathbb{R}^{m\times d})$ we denote the set of **unbounded (finite-order) laminates**, defined in complete analogy to $\mathscr{M}^{\mathrm{lc}}(K)$ for K compact, only without the restriction on the support.

Lemma 9.25. *There exists a sequence of finite-order laminates $(\nu_j) \subset \mathscr{M}^{\mathrm{lc}}(\mathbb{R}^{2\times 2})$ with*

$$[\nu_j] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \int \left| \frac{A + A^T}{2} \right| \mathrm{d}\nu_j(A) = \sqrt{2},$$

and

$$\lim_{j\to\infty} \int |A| \,\mathrm{d}\nu_j(A) = \infty.$$

Proof. Step 1. We first construct for every $k \in \mathbb{N}$ a laminate $\gamma_k \in \mathscr{M}^{\mathrm{lc}}(\mathbb{R}^{2\times 2})$ such that

$$[\gamma_k] = \begin{pmatrix} 0 & k \\ k & 0 \end{pmatrix}, \qquad \int \left| \frac{A + A^T}{2} \right| \mathrm{d}\gamma_k(A) = \sqrt{2}k, \qquad \int |A| \,\mathrm{d}\gamma_k(A) = \frac{5\sqrt{2}}{3}k.$$

Let $\delta_{\alpha,\beta}$ denote the Dirac mass at $M_{\alpha,\beta} := \left(\begin{smallmatrix} 0 & \alpha \\ \beta & 0 \end{smallmatrix}\right)$. For $k \in \mathbb{N}$ define

$$\gamma_k' := \frac{1}{3}\delta_{k,-k} + \frac{2}{3}\delta_{k,2k}.$$

Then, $[\gamma_k'] = M_{k,k}$ and γ_k' is a laminate since $M_{k,-k}$ and $M_{k,2k}$ are rank-one connected. Next, we set

$$\gamma_k'' := \frac{1}{4}\delta_{-2k,2k} + \frac{3}{4}\delta_{2k,2k},$$

for which $[\gamma_k''] = M_{k,2k}$. Again, γ_k'' is a laminate. Replacing the $\delta_{k,2k}$-term in the definition of γ_k' by γ_k'', we arrive at the second-order laminate

$$\gamma_k := \frac{1}{3}\delta_{k,-k} + \frac{1}{6}\delta_{-2k,2k} + \frac{1}{2}\delta_{2k,2k}.$$

It is easy to compute that γ_k has all the required properties.

Step 2. Let us now define $\nu_j \in \mathscr{M}^{\mathrm{lc}}(\mathbb{R}^{2\times 2})$. We set $\nu_1 := \gamma_1$. For $j = 2, 3, \dots$ the ν_j are defined iteratively. By construction, ν_1 contains the term $\frac{1}{2}\delta_{2,2}$. We apply the first step to this term and replace it by a laminate γ_2, yielding ν_2. Generally, ν_j contains a term of the form $2^{-j}\delta_{2^j,2^j}$, and we may replace the mass $\delta_{2^j,2^j}$ via the first step of the proof with a laminate γ_{2^j}, giving ν_{j+1}. Clearly, this iteration does not change the barycenter of the ν_j, so $[\nu_j] = M_{1,1}$. Moreover, for the absolute symmetric first moment we get

$$\int \left| \frac{A + A^T}{2} \right| \, \mathrm{d}\nu_j(A) = |M_{1,1}| = \sqrt{2}$$

since replacing $\delta_{2^j,2^j}$ with γ_{2^j} does not change the absolute symmetric first moment. If we unwind the recursion, we get

$$\nu_j = \sum_{n=1}^{j} \left(\frac{2^{1-n}}{3} \delta_{2^{n-1},-2^{n-1}} + \frac{2^{1-n}}{6} \delta_{-2^n,2^n} \right) + 2^{-n}\delta_{2^n,2^n}.$$

Thus, we may estimate

$$\int |A| \, \mathrm{d}\nu_j(A) \geq \sum_{n=1}^{j} \frac{2^{1-n}}{3} |M_{2^{n-1},-2^{n-1}}| = j\frac{\sqrt{2}}{3}.$$

Consequently,

$$\lim_{j\to\infty} \int |A| \, \mathrm{d}\nu_j(A) = \infty.$$

This finishes the proof. □

We next prove the following linearized version of Theorem 9.24, which is called **Ornstein's non-inequality** (or, more precisely, a special case of this quite general statement, see [162, 220]). It shows that an analogue of Korn's inequality (8.20) does not hold in $\mathrm{W}^{1,1}$.

Theorem 9.26 (Ornstein 1962 [220]). *For all $N \in \mathbb{N}$ there exists a map $u \in \mathrm{W}^{1,\infty}_0(B(0,1); \mathbb{R}^d)$ such that*

$$\inf_{F\in\mathbb{R}^{d\times d}} \int_{B(0,1)} |\nabla u - F| \, \mathrm{d}x \geq N \int_{B(0,1)} \left| \frac{\nabla u + \nabla u^T}{2} \right| \mathrm{d}x.$$

Proof. We only show the assertion for $d = 2$, the higher-dimensional cases follow from a natural embedding.

Step 1. We first prove that there exists a map $v \in \mathrm{W}^{1,\infty}_0(B(0,1); \mathbb{R}^2)$ with

$$\int_{B(0,1)} |\nabla v| \,\mathrm{d}x \geq N \int_{B(0,1)} \left| \frac{\nabla v + \nabla v^T}{2} \right| \mathrm{d}x. \tag{9.17}$$

By the preceding lemma, we may select a finite-order laminate $\nu \in \mathscr{M}^{\mathrm{lc}}(\mathbb{R}^{2\times 2})$ such that

$$\int |A| \,\mathrm{d}\nu(A) \geq 8\sqrt{2}N + \sqrt{2} \qquad \text{and} \qquad \int \left| \frac{A + A^T}{2} \right| \mathrm{d}\nu(A) = \sqrt{2}.$$

Then, by Lemma 9.3, ν is a homogeneous gradient Young measure on the domain $B(0,1)$ (the fact that we may choose the domain $B(0,1)$ follows from a covering argument as in Lemma 4.14). Thus, there exists a sequence $(v_j) \subset \mathrm{W}^{1,\infty}_{M_{1,1}x}(B(0,1);\mathbb{R}^2)$ with $\nabla v_j \overset{\mathbf{Y}}{\to} \nu$; in fact, the proof of Lemma 9.3 constructs this sequence explicitly. Setting $v(x) := v_j(x) - M_{1,1}x$ for j large enough, we may assume

$$\int_{B(0,1)} |\nabla v| \,\mathrm{d}x \geq \frac{1}{2} \int |A - M_{1,1}| \,\mathrm{d}\nu(A)$$

and

$$\int \left| \frac{A + A^T}{2} - M_{1,1} \right| \mathrm{d}\nu(A) \geq \frac{1}{2} \int_{B(0,1)} \left| \frac{\nabla v + \nabla v^T}{2} \right| \mathrm{d}x.$$

Then,

$$\begin{aligned}
\int_{B(0,1)} |\nabla v| \,\mathrm{d}x &\geq \frac{1}{2} \int |A - M_{1,1}| \,\mathrm{d}\nu(A) \\
&\geq \frac{1}{2} \int |A| \,\mathrm{d}\nu(A) - \frac{\sqrt{2}}{2} \\
&\geq 4\sqrt{2}N \\
&= 2N \int \left| \frac{A + A^T}{2} \right| \mathrm{d}\nu(A) + 2\sqrt{2}N \\
&\geq 2N \int \left| \frac{A + A^T}{2} - M_{1,1} \right| \mathrm{d}\nu(A) \\
&\geq N \int_{B(0,1)} \left| \frac{\nabla v + \nabla v^T}{2} \right| \mathrm{d}x.
\end{aligned}$$

Thus, (9.17) follows.

Step 2. Let $v \in \mathrm{W}^{1,\infty}_0(B(0,1);\mathbb{R}^2)$ satisfy (9.17). We now show Ornstein's claim by scaling. Set for $0 < r < 1$,

$$v_r(x) := r v\left(\frac{x}{r}\right), \qquad x \in B(0,1),$$

where we consider v to be extended by zero to all of $\mathbb{R}^2$. We calculate for $r \le 2^{-1/d}$ that

$$\begin{aligned}
&\inf_{F\in\mathbb{R}^{2\times 2}} \int_{B(0,1)} |\nabla v_r - F|\,\mathrm{d}x \\
&\quad\ge \inf_{F\in\mathbb{R}^{2\times 2}} \left(\int_{B(0,r)} |\nabla v_r|\,\mathrm{d}x - \int_{B(0,r)} |F|\,\mathrm{d}x + \int_{B(0,1)\setminus B(0,r)} |F|\,\mathrm{d}x \right) \\
&\quad\ge \int_{B(0,r)} |\nabla v_r|\,\mathrm{d}x + \inf_{F\in\mathbb{R}^{2\times 2}} \bigl(\omega_d(1-2r^d)|F|\bigr) \\
&\quad\ge \int_{B(0,1)} |\nabla v_r|\,\mathrm{d}x.
\end{aligned}$$

Since (9.17) is invariant under the scaling, the assertion of the theorem follows with $u := v_r$. □

Proof of Theorem 9.24. Again, we only consider the case $d = 2$. Let u be the function from Ornstein's Theorem 9.26 for the constant $2N$. Set

$$v(x) := x + \varepsilon u(x), \qquad x \in B(0,1),$$

for an $\varepsilon > 0$ to be determined later. We also observe, see (A.2), that

$$\operatorname{dist}(\mathrm{Id} + A, \mathrm{SO}(2)) \le \frac{1}{2}|A + A^T| + C|A|^2.$$

Thus, since $\nabla v = \mathrm{Id} + \varepsilon \nabla u$,

$$\begin{aligned}
\int_{B(0,1)} \operatorname{dist}(\nabla v, \mathrm{SO}(2))\,\mathrm{d}x \le \varepsilon \int_{B(0,1)} \left|\frac{\nabla u + \nabla u^T}{2}\right| \mathrm{d}x \\
+ \varepsilon^2 C \int_{B(0,1)} |\nabla u|^2\,\mathrm{d}x.
\end{aligned}$$

Now choose $\varepsilon > 0$ so small that the second term is less than or equal to the first. Then,

$$\begin{aligned}
N \int_{B(0,1)} \operatorname{dist}(\nabla v, \mathrm{SO}(2))\,\mathrm{d}x &\le 2N\varepsilon \int_{B(0,1)} \left|\frac{\nabla u + \nabla u^T}{2}\right| \mathrm{d}x \\
&\le \varepsilon \inf_{F\in\mathbb{R}^{2\times 2}} \int_{B(0,1)} |\nabla u - F|\,\mathrm{d}x \\
&= \inf_{G\in\mathbb{R}^{2\times 2}} \int_{B(0,1)} |\nabla v - G|\,\mathrm{d}x \\
&\le \inf_{Q\in\mathrm{SO}(2)} \int_{B(0,1)} |\nabla v - Q|\,\mathrm{d}x.
\end{aligned}$$

This proves the claim. □

Notes and Historical Remarks

The task to compute quasiconvex hulls first arose in the theory of mathematical material science, in particular in the work of Ball and James [30, 31]. However, the explicit definitions and study of the rank-one convex hull and polyconvex hull as lower and upper bounds, respectively, only seem to have appeared later (even though the methods to compute them are older). Dacorogna contributed many results investigating the properties of the various hulls, see [76] for an overview and pointers to the literature. One can also define notions of quasiconvex, polyconvex, and rank-one convex sets, not just envelopes as we have done, and develop a corresponding theory, see Chapter 7 in [76].

For most of the material on the application to the two-well problems we follow [203, 205]. The proof of Proposition 9.4 is based on ideas from [251]. Theorem 9.5 is from [254].

The basic idea of convex integration dates back to the famous Nash–Kuiper Theorem in differential geometry [173, 212]. These arguments were developed further until they culminated in Gromov's "h-principle", which is detailed in his treatise [145] (also see [144]). There, the Lipschitz case, which is the one relevant to us, is also briefly mentioned; in particular, the term "in-approximation" is due to Gromov. This work was transferred by Müller and Šverák [205, 207] to the multi-well inclusions that are of interest in material sciences. Simultaneously, Dacorogna and Marcellini [79, 80] adapted Cellina's Baire method [20, 58, 59] for ordinary differential inclusions to partial differential inclusions. Sychev [257, 260, 262] developed the method further to prove an existence theorem that only assumed the existence of what we call a "piecewise affine reduction". He also showed that the situation with an in-approximation (or RC-in-approximation) can be reduced to this abstract result. We remark that Sychev's original result, Theorem 1.1 in [260], is in fact a bit more general than our version. Finally, Kirchheim's approach based on Baire-one functionals unified the different strands of development, see [160], and further relates convex integration to the Banach–Mazur game. Our presentation mostly follows [160], but we also took some inspiration from [263]. We also remark that the term "convex integration" sometimes only designates the approach via an (RI-)in-approximation. Kirchheim speaks of "affine synthesis" in [160].

The technical Lemma 9.10 is from [207], which itself is a variant of Theorem 2.4 in [120], but the result is already mentioned in Gromov's book [145]. While Theorem 9.12 as stated is due to Müller and Šverák [205], at the same time the Baire approach by Dacorogna and Marcellini [79] also resulted in similar results.

For convex integration in the inhomogeneous case, that is, where $K = K(x)$, we refer to [208]. We also mention [206], where additionally a uniform determinant constraint is respected in the convex integration procedure. Convex integration has also been applied to the Euler equation to show the existence of very irregular solutions [91, 152, 239, 243]. One important and recurring theme in convex integration theory is that the *regularity* we require of solutions determines whether we can find solutions at all. The task is then to precisely identify the *threshold* between rigidity and non-rigidity.

Ornstein's non-inequality is in fact a much more general result that shows that many singular integral operators are not $(\mathrm{L}^1 \mapsto \mathrm{L}^1)$-bounded. An abstract approach to these questions is in [162].

Another aspect of microstructure which was not touched upon in this chapter is that of *uniqueness* and *stability*. Here, a microstructure $\nu \in \mathscr{M}^{\mathrm{qc}}(K)$ is called **unique** (in K) if $\hat{\nu} \in \mathscr{M}^{\mathrm{qc}}(K)$ with $[\hat{\nu}] = [\nu]$ implies $\hat{\nu} = \nu$. Many ideas in this area go back to Luskin; a nice modern exposition, which also incorporates some new ideas with a view toward numerics is in Chapter 4 of [100].

Problems

9.1. Let $K \subset \mathbb{R}^{m \times d}$ be a non-empty compact set. Show that K^{lc} is the smallest set $M \subset \mathbb{R}^{m \times d}$ that is closed under laminations, i.e., if $A, B \in M$ and $\operatorname{rank}(A - B) = 1$, then $\theta A + (1 - \theta)B \in M$ for all $\theta \in (0, 1)$.

9.2. Let $(t_k, A_k)_{k=1,\ldots,n} \subset [0, 1] \times \mathbb{R}^{m \times d}$ with $\sum_k t_k = 1$. We define:

(i) If $n = 2$ we say that $(t_k, A_k)_k$ satisfies the (H_2)**-condition** if $\operatorname{rank}(A_1 - A_2) \le 1$.
(ii) If $n > 2$ we say that $(t_k, A_k)_k$ satisfies the (H_n)**-condition** if possibly after a permutation of indices,

$$\operatorname{rank}(A_1 - A_2) \le 1$$

and with

$$\begin{aligned} s_1 &= t_1 + t_2, & B_1 &= \frac{t_1}{s_1} A_1 + \frac{t_2}{s_1} A_2, \\ s_k &= t_{k+1}, & B_k &= A_{k+1} \quad \text{for} \quad k = 2, 3, \ldots, n, \end{aligned}$$

the collection $(s_k, B_k)_{k=1,\ldots,n-1}$ satisfies the (H_{n-1})-condition.

Prove that

$$\mathscr{M}^{\mathrm{lc}}(K) = \Biggl\{ \mu \in \mathscr{M}^1(K) \;:\; \mu = \sum_{k=1}^n t_k \delta_{A_k}, \;\; (t_k, A_k)_k \text{ satisfies the } (H_n)\text{-condition} \\ \text{for some } n \in \mathbb{N} \Biggr\}.$$

9.3. Show that for all non-empty compact sets $K \subset \mathbb{R}^{m \times d}$ it holds that

$$K^{\square} = \bigl\{ A \in \mathbb{R}^{m \times d} \;:\; h(A) \le \inf_K h \text{ for all } h \colon \mathbb{R}^{m \times d} \to \mathbb{R} \text{ such that } h = h^{\square} \bigr\},$$

where $\square \in \{\text{rc}, \text{qc}, \text{pc}, **\}$, and $h^\square$ is the respective envelope of h (i.e., $h = h^\square$ means that h has the respective convexity property).

9.4. Show that for a compact set $K \subset \mathbb{R}^{3\times 3}$ it holds that

$$K^{\text{pc}} = \left\{ A \in \mathbb{R}^{3\times 3} \;:\; (A, \operatorname{cof} A, \det A) \in (\mathbf{M}K)^{**} \right\},$$

where

$$\mathbf{M}K := \left\{ (A, \operatorname{cof} A, \det A) \;:\; A \in K \right\}.$$

What is the corresponding formula for general dimensions?

9.5. Let

$$K := \mathrm{SO}(2) \cup \mathrm{SO}(2)U, \quad U := \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}, \quad 0 < \alpha \le 1 \le \beta, \quad \alpha\beta \ge 1.$$

Show that for every $A \in K^{**}$ we can find unique $y, z \in \mathbb{R}^2$ with

$$|y|^2 = y_1^2 + y_2^2 \le 1 \qquad \text{and} \qquad |z|^2 = z_1^2 + z_2^2 \le 1$$

such that

$$A = \begin{pmatrix} y_1 & -y_2 \\ y_2 & y_1 \end{pmatrix} + \begin{pmatrix} z_1 & -z_2 \\ z_2 & z_1 \end{pmatrix} U.$$

Show furthermore that the mapping $(y, z) \mapsto A$ is a diffeomorphism.

9.6. Show that for all sufficiently small open sets $U \supset K_{\mathrm{T4}}$ there are no rank-one connections in U.

9.7. Show that for the set

$$K := \left\{ \begin{pmatrix} x & 0 \\ 0 & y \\ z & z \end{pmatrix} \;:\; x, y, z \in \{-1, 1\} \right\}$$

from Example 9.22 it holds that

$$K^{\text{lc}} = K^{\text{rc}} = K^{\text{qc}} = K^{\text{pc}} = K^{**} = \left\{ \begin{pmatrix} x & 0 \\ 0 & y \\ z & z \end{pmatrix} \;:\; x, y, z \in [-1, 1] \right\}.$$

9.8. For an open set $G \subset \mathbb{R}^{m\times d}$ we define

$$G^{\text{qc}} := \bigcup \left\{ K^{\text{qc}} \;:\; K \subset G \text{ compact} \right\}.$$

Assume that $K \subset G^{\text{qc}}$ is compact. Show that there exists a subset $G_0 \Subset G$ with $K \subset G_0^{\text{qc}}$.

9.9. Use a convex integration procedure to show that there are continuous, bounded functions on the interval [0, 1] that are nowhere differentiable.

9.10. Show that if a compact set $K \subset \mathbb{R}^{m\times d}$ has an in-approximation consisting of open, uniformly bounded sets $G_k \subset \mathbb{R}^{m\times d}$ then K is piecewise affinely stable with respect to $G_\infty := \bigcup_{l=1}^{\infty} G_l$.

Chapter 10
Singularities

All of the existence theorems for minimizers of integral functionals defined on Sobolev spaces $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ that we have seen so far required that $p > 1$. Extending the existence theory to the *linear-growth* case $p = 1$ turns out to be quite intricate and necessitates the development of new tools. In order to illustrate this, consider the following minimization problem:

$$\begin{cases} \text{Minimize} \quad \mathscr{F}[u] := \int_\Omega f(x, \nabla u(x)) \, \mathrm{d}x \\ \text{over all} \quad\; u \in \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m) \text{ with } u|_{\partial\Omega} = g. \end{cases} \tag{10.1}$$

Here, $f \colon \Omega \times \mathbb{R}^{m\times d} \to \mathbb{R}$ is a Carathéodory integrand and $g \in \mathrm{L}^1(\partial\Omega;\mathbb{R}^m)$, which is the trace space for $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ (see Appendix A.5). Our hypothesis of *linear growth* on f means that there exists a constant $M > 0$ such that

$$|f(x, A)| \le M(1 + |A|), \qquad (x, A) \in \Omega \times \mathbb{R}^{m\times d}.$$

If for the moment we also assume coercivity in the form

$$\mu|A| \le f(x, A), \qquad (x, A) \in \Omega \times \mathbb{R}^{m\times d},$$

where $\mu > 0$, then a minimizing sequence $(u_j) \subset \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ satisfies

$$\limsup_{j\to\infty} \|u_j\|_{\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)} < \infty.$$

Here we also used the Poincaré inequality, see Theorem A.26 (i), in conjunction with the fixed boundary values.

However, when trying to apply the Direct Method, it turns out that (10.1) is badly behaved: It is a well-known observation that in $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ a uniform norm-bound

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8_10

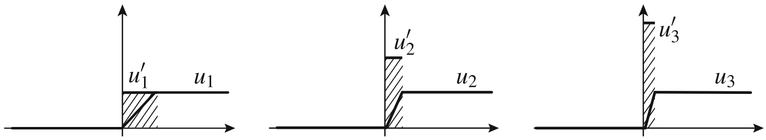


Fig. 10.1 A concentrating sequence

is *not* enough to deduce weak compactness, which is due to the non-reflexivity of $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$. For instance, take

$$u_j(x) := jx\mathbb{1}_{(0,1/j)}(x) + \mathbb{1}_{(1/j,1)}(x) \in \mathrm{W}^{1,1}(-1, 1)$$

and observe that the u_j converge to $u = \mathbb{1}_{(0,1)} \notin \mathrm{W}^{1,1}(-1, 1)$ pointwise and in L^1, see Figure 10.1. The decisive feature of the u_j's is that the (weak) derivatives u_j' *concentrate*, meaning that the family $\{u_j'\}_j$ is not equiintegrable. In fact, the measures $u_j' \mathscr{L}^d \llcorner (-1, 1)$ converge weakly* (in the sense of measures) to a measure that is not absolutely continuous with respect to Lebesgue measure, namely the Dirac mass δ_0.

We conclude that the space $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$ is too small to serve as the set of candidate functions for our minimization problems (10.1). Clearly, we need a space containing $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$ such that a uniform norm-bound implies precompactness for a suitable (weak) convergence. Moreover, in order to retain the connection to our original problem, it is desirable that $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$ is dense with respect to another notion of convergence that is strong enough to make the above functional $\mathscr{F}$ continuous. All these requirements are fulfilled by the space $\mathrm{BV}(\Omega; \mathbb{R}^m)$ of *functions of bounded variation*, as we will see in this and the next chapter.

Before we can delve into the theory of integral functionals defined on BV-maps in the next chapter, and in particular consider how to extend (10.1) to this space, we first need to study *singularities* in measures and BV-functions. This is the topic of this chapter. After introducing the so-called strict convergence of measures, we define tangent measures, which allow us to study the local shape of measures. Then, we turn to an analysis of the structure of singularities in BV-maps, and, more generally, in PDE-constrained measures. Finally, we present a surprising result about convexity at singularities, which will become useful in Chapter 12.

10.1 Strict Convergence of Measures

In all of this and the next two chapters we will make constant use of the theory of vector measures, which is recalled in Appendix A.4.

We say that a sequence of (vector) measures $(\mu_j) \subset \mathscr{M}(\Omega; \mathbb{R}^N)$ converges **strictly** to $\mu \in \mathscr{M}(\Omega; \mathbb{R}^N)$, written as "$\mu_j \to \mu$ strictly", if

$$\mu_j \overset{*}{\rightharpoonup} \mu \quad \text{in } \mathscr{M}(\Omega;\mathbb{R}^N) \qquad \text{and} \qquad |\mu_j|(\Omega) \to |\mu|(\Omega).$$

Here we recall that $|\mu| \in \mathscr{M}^+(\Omega)$ denotes the total variation measure of the vector measure μ, see Appendix A.4.

Lemma 10.1. *Let $\lambda_j \to \lambda$ strictly in $\mathscr{M}^+(\mathbb{R}^d)$. Then,*

$$\int g \, d\lambda_j \to \int g \, d\lambda \tag{10.2}$$

for all continuous and bounded functions $g \colon \mathbb{R}^d \to \mathbb{R}$. The same conclusion holds if instead of strict convergence we only assume $\liminf_{j\to\infty} \lambda_j(U) \geq \lambda(U)$ for all open sets $U \subset \mathbb{R}^d$, and $\lambda_j(\mathbb{R}^d) \to \lambda(\mathbb{R}^d)$.

Proof. By considering $\alpha + \beta g$ for $\alpha, \beta \in \mathbb{R}$, we may without loss of generality assume that $g(x) \in [0,1]$ for all $x \in \mathbb{R}^d$. Let $E_t := \{ x \in \mathbb{R}^d : g(x) > t \}$ for $t \in [0,1]$ and observe via Fubini's theorem that

$$\int g \, d\lambda = \int \int_0^1 \mathbb{1}_{E_t}(x) \, dt \, d\lambda(x) = \int_0^1 \lambda(E_t) \, dt.$$

Since E_t is open, $\lambda_j \overset{*}{\rightharpoonup} \lambda$ implies (via Lemma A.19)

$$\liminf_{j\to\infty} \lambda_j(E_t) \geq \lambda(E_t) \qquad \text{for all } t \in [0,1].$$

Thus, by Fatou's lemma,

$$\liminf_{j\to\infty} \int g \, d\lambda_j = \liminf_{j\to\infty} \int_0^1 \lambda_j(E_t) \, dt \geq \int_0^1 \lambda(E_t) \, dt = \int g \, d\lambda.$$

Let

$$a_j := \int g \, d\lambda_j, \qquad a := \int g \, d\lambda, \qquad b_j := \int 1 - g \, d\lambda_j, \qquad b := \int 1 - g \, d\lambda.$$

Then, the arguments above yield

$$\liminf_{j\to\infty} a_j \geq a, \qquad \liminf_{j\to\infty} b_j \geq b.$$

Moreover, since $\lambda_j(\mathbb{R}^d) \to \lambda(\mathbb{R}^d)$,

$$\lim_{j\to\infty} (a_j + b_j) = a + b.$$

Thus, it follows by elementary means that $a_j \to a$, that is, (10.2).

The additional statement is clear since we only used that $\liminf_{j\to\infty} \lambda_j(U) \geq \lambda(U)$ for open sets $U \subset \mathbb{R}^d$ and $\lambda_j(\mathbb{R}^d) \to \lambda(\mathbb{R}^d)$. □

Corollary 10.2. *If $\mu_j \to \mu$ strictly in $\mathscr{M}(\Omega;\mathbb{R}^N)$, then also $|\mu_j| \to |\mu|$ strictly.*

Proof. We need to show that $|\mu_j| \overset{*}{\rightharpoonup} |\mu|$ in $\mathscr{M}^+(\Omega)$.

It holds that $\liminf_{j\to\infty} |\mu_j|(U) \geq |\mu|(U)$ for all open sets $U \subset \mathbb{R}^d$ by the weak* convergence $\mu_j \overset{*}{\rightharpoonup} \mu$ and the fact that the total variation for open sets can be written as the supremum of weakly*-continuous functions, see (A.3), hence it must be lower semicontinuous with respect to the weak* convergence. Since also $|\mu_j|(\mathbb{R}^d) \to |\mu|(\mathbb{R}^d)$ by assumption, we may apply the preceding lemma for $\lambda_j := |\mu_j|$, $\lambda := |\mu|$. Then, (10.2) for $g \in \mathrm{C}_0(\mathbb{R}^d)$ implies that indeed $|\mu_j| \overset{*}{\rightharpoonup} |\mu|$ in $\mathscr{M}^+(\Omega)$. □

The most important result about strict convergence is the following *Reshetnyak continuity theorem*. For this, denote by

$$P := \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|} \in \mathrm{L}^1(\mathbb{R}^d, |\mu|; \mathbb{R}^N)$$

the **polar** of μ, i.e., the Radon–Nikodým density of μ with respect to it own total variation measure $|\mu|$ (see the Besicovitch Differentiation Theorem A.23), so that

$$\mu = P|\mu| \qquad \text{and} \qquad |P| = 1 \quad |\mu|\text{-a.e.}$$

Theorem 10.3 (**Reshetnyak 1967 [225]**). *Let $\Omega \subset \mathbb{R}^d$ be open and let $g \in \mathrm{C}(\overline{\Omega} \times \mathbb{S}^{N-1})$. For any sequence $(\mu_j) \subset \mathscr{M}(\Omega;\mathbb{R}^N)$ with $\mu_j \to \mu$ strictly it holds that*

$$\int g\left(x, \frac{\mathrm{d}\mu_j}{\mathrm{d}|\mu_j|}(x)\right) \mathrm{d}|\mu_j|(x) \to \int g\left(x, \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x)\right) \mathrm{d}|\mu|(x).$$

Proof. Define

$$\lambda_j := |\mu_j|(\mathrm{d}x) \otimes \delta_{P_j(x)} \in \mathscr{M}^+(\Omega \times \mathbb{S}^{N-1}), \qquad \text{where} \qquad P_j(x) := \frac{\mathrm{d}\mu_j}{\mathrm{d}|\mu_j|}(x),$$

that is, for $f \in \mathrm{C}_0(\Omega \times \mathbb{S}^{N-1})$,

$$\int_{\Omega \times \mathbb{S}^{N-1}} f(x, A)\, \mathrm{d}\lambda_j(x, A) = \int_\Omega f\left(x, \frac{\mathrm{d}\mu_j}{\mathrm{d}|\mu_j|}(x)\right) \mathrm{d}|\mu_j|(x).$$

We have

$$\sup_{j\in\mathbb{N}} |\lambda_j|(\Omega \times \mathbb{S}^{N-1}) = \sup_{j\in\mathbb{N}} |\mu_j|(\Omega) < \infty.$$

Thus, we may select a subsequence (not explicitly labeled) with the property that

$$\lambda_j \overset{*}{\rightharpoonup} \lambda \quad \text{in} \mathscr{M}^+(\Omega \times \mathbb{S}^{N-1}).$$

Let $\pi : \Omega \times \mathbb{R}^N \to \mathbb{R}^d$ be the projection onto the first argument, $\pi(x, A) := x$. On the one hand, we get $\pi_\#\lambda_j \overset{*}{\rightharpoonup} \pi_\#\lambda$, where $\pi_\#\lambda := \lambda \circ \pi^{-1}$ is the push-forward of λ under π, see Appendix A.4. On the other hand, Corollary 10.2 implies

$$\pi_\#\lambda_j = |\mu_j| \overset{*}{\rightharpoonup} |\mu|.$$

Hence, $\pi_\#\lambda = |\mu|$ and by the Disintegration Theorem 4.4 there exists a weakly* measurable family $(\nu_x)_{x\in\Omega} \subset \mathscr{M}^1(\mathbb{S}^{N-1})$ such that

$$\lambda = |\mu|(\mathrm{d}x) \otimes \nu_x.$$

Next, test the convergence $\lambda_j \overset{*}{\rightharpoonup} \lambda$ with functions of the form $(x, A) \mapsto \psi(x)A$, where $\psi \in \mathrm{C}_0(\Omega)$ (note that these functions lie in $\mathrm{C}_0(\Omega \times \mathbb{S}^{N-1})$) to see that

$$\begin{aligned}
\int_\Omega \psi(x) \int A \;\mathrm{d}\nu_x(A)\;\mathrm{d}|\mu|(x) &= \int_{\Omega\times\mathbb{S}^{N-1}} \psi(x)A \;\mathrm{d}\lambda(x, A) \\
&= \lim_{j\to\infty} \int_{\Omega\times\mathbb{S}^{N-1}} \psi(x)A \;\mathrm{d}\lambda_j(x, A) \\
&= \lim_{j\to\infty} \int_\Omega \psi(x) \frac{\mathrm{d}\mu_j}{\mathrm{d}|\mu_j|}(x)\;\mathrm{d}|\mu_j|(x) \\
&= \lim_{j\to\infty} \int_\Omega \psi(x)\;\mathrm{d}\mu_j(x) \\
&= \int_\Omega \psi(x)\;\mathrm{d}\mu(x) \\
&= \int_\Omega \psi(x) \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x)\;\mathrm{d}|\mu|(x).
\end{aligned}$$

Consequently, varying ψ, for the barycenter $[\nu_x]$ of ν_x we get

$$[\nu_x] = \int A \;\mathrm{d}\nu_x(A) = \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x) \qquad \text{for } |\mu|\text{-a.e. } x \in \Omega.$$

Then, at $|\mu|$-almost every $x \in \Omega$, it holds that

$$\begin{aligned}
&\frac{1}{2}\int_{\mathbb{S}^{N-1}} \left| A - \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x)\right|^2 \mathrm{d}\nu_x(A) \\
&\quad = \frac{1}{2}\int_{\mathbb{S}^{N-1}} |A|^2 - 2\frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x)\cdot A + \left|\frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x)\right|^2 \mathrm{d}\nu_x(A) \\
&\quad = 1 - \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x)\cdot [\nu_x] \\
&\quad = 0.
\end{aligned}$$

Hence,

$$\nu_x = \delta_{P(x)} \qquad \text{for } |\mu|\text{-a.e. } x \in \Omega, \qquad \text{where} \qquad P(x) := \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x).$$

Finally, $\lambda_j \to \lambda$ strictly since

$$|\lambda_j|(\Omega \times \mathbb{S}^{N-1}) = |\mu_j|(\Omega) \to |\mu|(\Omega) = |\lambda|(\Omega \times \mathbb{S}^{N-1}).$$

By Lemma 10.1 we thus get for every g as in the statement of the theorem that

$$\begin{aligned}\int g\Big(x, \frac{\mathrm{d}\mu_j}{\mathrm{d}|\mu_j|}(x)\Big)\, \mathrm{d}|\mu_j|(x) &= \int g(x, A)\, \mathrm{d}\lambda_j(x, A) \\ &\to \int g(x, A)\, \mathrm{d}\lambda(x, A) \\ &= \int g\Big(x, \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x)\Big)\, \mathrm{d}|\mu|(x)\end{aligned}$$

as $j \to \infty$. This yields the sought assertion since it identifies the limit for *every* subsequence. □

10.2 Tangent Measures

We now introduce a tool to study the local structure of measures. For $x_0 \in \mathbb{R}^d$ and $r > 0$ define the rescaling map

$$T^{(x_0,r)}(x) := \frac{x - x_0}{r}, \qquad x \in \mathbb{R}^d.$$

For a vector-valued (local) Radon measure $\mu \in \mathscr{M}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$ and $x_0 \in \mathbb{R}^d$, a **tangent measure** to μ at x_0 is any (local) weak* limit in the space $\mathscr{M}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$ of the rescaled measures

$$c_n T^{(x_0,r_n)}_{\#}\mu$$

for some sequence $r_n \downarrow 0$ of radii and rescaling constants $c_n > 0$. The definition of the push-forward $T^{(x_0,r)}_{\#}\mu$ (see Appendix A.4) here expands to

$$[T^{(x_0,r_n)}_{\#}\mu](B) = \mu(x_0 + r_n B) \qquad \text{for any Borel set } B \subset \mathbb{R}^d.$$

The sequence $(c_n T^{(x_0,r_n)}_{\#}\mu)_n$ is called a **blow-up sequence**. We collect all tangent measures to μ at a point x_0 in the set $\mathrm{Tan}(\mu, x_0)$. Clearly, $0 \in \mathrm{Tan}(\mu, x_0)$ for all $x_0 \in \mathbb{R}^d$ and $\mathrm{Tan}(\mu, x_0) = \{0\}$ for all $x_0 \notin \operatorname{supp}\mu$. We will see in Proposition 10.5 below that $\mathrm{Tan}(\mu, x_0)$ contains a non-zero measure for $|\mu|$-almost every $x_0 \in \mathbb{R}^d$.

For any non-zero $\tau \in \mathrm{Tan}(\mu, x_0)$ it turns out that, up to taking subsequences, we may always choose the rescaling constants c_n in the blow-up sequence $c_n T_\#^{(x_0,r_n)} \mu \overset{*}{\rightharpoonup} \tau$ as

$$\tilde{c}_n := c\big[|\mu|(x_0 + r_n\overline{U})\big]^{-1} \tag{10.3}$$

for any (fixed) bounded open set $U \subset \mathbb{R}^d$ containing the origin such that $|\tau|(U) > 0$, and some constant $c > 0$ (which, of course, depends on τ and U). Indeed,

$$0 < |\tau|(U) \le \liminf_{n\to\infty} c_n|\mu|(x_0 + r_n U) \le \limsup_{n\to\infty} c_n|\mu|(x_0 + r_n\overline{U}) < \infty$$

by the (local) weak* convergence of $c_n T_\#^{(x_0,r_n)}|\mu|$. After the selection of a subsequence (not relabeled) we may assume $c_n|\mu|(x_0 + r_n\overline{U}) \to c$, and hence that $\tilde{c}_n T_\#^{(x_0,r_n)}\mu \overset{*}{\rightharpoonup} \tau$.

The first fact we prove about tangent measures is a very useful (measure-theoretic) continuity property:

Lemma 10.4. *Let $\mu \in \mathscr{M}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$. At $|\mu|$-almost every point $x_0 \in \mathbb{R}^d$ and for all blow-up sequences $(c_n T_\#^{(x_0,r_n)}\mu)_n$, where $r_n \downarrow 0$, $c_n > 0$, it holds that*

$$\tau = \mathrm{w}^*\text{-}\lim_{n\to\infty} c_n T_\#^{(x_0,r_n)}\mu \qquad \Longleftrightarrow \qquad |\tau| = \mathrm{w}^*\text{-}\lim_{n\to\infty} c_n T_\#^{(x_0,r_n)}|\mu|.$$

In this case,

$$\tau = P_0|\tau| \quad \textit{with} \quad P_0 = \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x_0).$$

Thus, $\mathrm{Tan}(\mu, x_0) = \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x_0) \cdot \mathrm{Tan}(|\mu|, x_0)$ *for $|\mu|$-almost every $x_0 \in \mathbb{R}^d$.*

Proof. Let $x_0 \in \mathbb{R}^d$ be a Lebesgue point of the polar $P := \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}$ of μ with respect to $|\mu|$, that is,

$$\fint_{B(x_0,r_nR)} |P(x) - P_0| \,\mathrm{d}|\mu|(x) \to 0 \qquad \text{as } n \to \infty.$$

By general measure theory results, $|\mu|$-almost every $x_0 \in \mathbb{R}^d$ has this property, see Theorem A.20.

Let $\varphi \in \mathrm{C}_c(\mathbb{R}^d)$. By (10.3) we may assume that $c_n = c[|\mu|(\overline{B(x_0, r_nR)})]^{-1}$ for some $c \ge 0$ and a large ball $B(0, R) \Supset \operatorname{supp}\varphi$ $(R > 0)$. We have

$$\begin{aligned}
&\int \varphi \,\mathrm{d}T_\#^{(x_0,r_n)}|\mu| - \int \varphi\, P_0 \cdot \mathrm{d}T_\#^{(x_0,r_n)}\mu \\
&\quad = \int \varphi(y)\big[1 - P_0 \cdot P(x_0 + r_n y)\big] \,\mathrm{d}T_\#^{(x_0,r_n)}|\mu|(y) \\
&\quad = \int \varphi\Big(\frac{x - x_0}{r_n}\Big)\big[1 - P_0 \cdot P(x)\big] \,\mathrm{d}|\mu|(x).
\end{aligned}$$

Since

$$|1 - P_0 \cdot P(x)| \le |P(x) - P_0|,$$

the Lebesgue point property of x_0 yields

$$\begin{aligned} c_n \left| \int \varphi \, \mathrm{d}T_{\#}^{(x_0,r_n)}|\mu| - \int \varphi \, P_0 \cdot \mathrm{d}T_{\#}^{(x_0,r_n)}\mu \right| \\ \le c\|\varphi\|_\infty \fint_{B(x_0,r_nR)} |P(x) - P_0| \, \mathrm{d}|\mu|(x) \\ \to 0 \qquad \text{as } n \to \infty. \end{aligned}$$

Therefore, if $c_n T_{\#}^{(x_0,r_n)}\mu$ converges weakly* to $\tau \in \mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$, then $c_n T_{\#}^{(x_0,r_n)}|\mu|$ converges weakly* to $P_0 \cdot \tau =: \sigma$. If we write $\tau = G|\tau|$ with the polar Radon–Nikodým density $G \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d, |\tau|; \mathbb{S}^{d-1})$, then

$$|\tau| \le \operatorname*{w^*\text{-}lim}_{n\to\infty} c_n T_{\#}^{(x_0,r_n)}|\mu| = \sigma = P_0 \cdot \tau = (P_0 \cdot G)|\tau| \le |\tau|,$$

whereby $G = P_0$ almost everywhere with respect to $|\tau|$, i.e., $\sigma = |\tau|$.

On the other hand, if $c_n T_{\#}^{(x_0,r_n)}|\mu|$ converges weakly* to a measure $\sigma \in \mathcal{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$, then we also get (selecting a subsequence if necessary) that

$$c_n T_{\#}^{(x_0,r_n)}\mu \overset{*}{\rightharpoonup} \tau \in \mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N).$$

Immediately, the previous argument applies and again yields $\sigma = |\tau|$. □

The following result shows that $|\mu|$-almost everywhere there exists a non-zero tangent measure.

Proposition 10.5. *Let $\mu \in \mathcal{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$. At $|\mu|$-almost every $x_0 \in \mathbb{R}^d$ there exists a non-zero tangent measure $\tau \in \mathrm{Tan}(\mu, x_0)$. Moreover, τ can be chosen to have polynomial growth of order $d+2$, i.e., there exists a constant $C > 0$ (depending on τ) such that $|\tau|(B(0,r)) \le C(1 + r^{d+2})$ for all $r > 0$.*

Proof. Restricting to a ball centered at the origin (tangent measures are local), we may assume without loss of generality that μ is finite. Further, by the previous lemma we may suppose without loss of generality that μ is a positive measure.

Fix $\varepsilon > 0$ and set

$$\beta_k := \frac{2^d(k+1)^d k^2}{\varepsilon}\mu(\mathbb{R}^d) \qquad \text{for } k = 2, 3, \ldots.$$

Define for $k = 2, 3, \ldots$ and $r > 0$ the set

$$A_{k,r} := \left\{ x \in \mathbb{R}^d \;:\; \mu(B(x,kr)) \ge \beta_k \mu(B(x,r)) \right\}.$$

Assume that for some ball $B(x, r/2) \subset \mathbb{R}^d$ it holds that $B(x, r/2) \cap A_{k,r} \neq \emptyset$. Then, take $z \in B(x, r/2) \cap A_{k,r}$ and estimate

$$\beta_k \mu(B(x, r/2)) \le \beta_k \mu(B(z, r)) \le \mu(B(z, kr)) \le \mu(B(x, (k+1)r)).$$

Hence, using the formula

$$\mu(B) = \frac{1}{\omega_d r^d} \int \mu(B \cap B(x, r)) \, \mathrm{d}x \qquad \text{for any Borel set} B \subset \mathbb{R}^d,$$

which follows from Fubini's theorem (recall that $\omega_d := |B(0, 1)|$), we get

$$\begin{aligned} \mu(A_{k,r}) &= \frac{1}{\omega_d (r/2)^d} \int \mu(A_{k,r} \cap B(x, r/2)) \, \mathrm{d}x \\ &\le \frac{2^d (k+1)^d}{\beta_k} \cdot \frac{1}{\omega_d (k+1)^d r^d} \int \mu(B(x, (k+1)r)) \, \mathrm{d}x \\ &= \frac{2^d (k+1)^d}{\beta_k} \mu(\mathbb{R}^d) \\ &= \frac{\varepsilon}{k^2}. \end{aligned}$$

Next, for

$$B_r := \left\{ x \in \mathbb{R}^d \ : \ \text{there exists a } k \in \{2, 3, \ldots\} \text{ such that } x \in A_{k,r} \right\}$$

we have

$$\mu(B_r) \le \sum_{k=2}^{\infty} \frac{\varepsilon}{k^2} \le \frac{\pi^2}{6} \varepsilon.$$

Defining

$$B := \bigcup_{n=1}^{\infty} \bigcap_{l=n}^{\infty} B_{1/l},$$

we see from the continuity of μ from below that

$$\mu(B) \le \frac{\pi^2}{6} \varepsilon.$$

Let $x_0 \in \operatorname{supp} \mu \setminus B$. For all $n \in \mathbb{N}$ by definition there exists an $l_n \ge n$ such that

$$\mu(B(x_0, k/l_n)) < \beta_k \mu(B(x_0, 1/l_n)) \qquad \text{for all } k = 2, 3, \ldots.$$

Set r_n to be $1/l_n$ for the l_n so chosen at n. Then,

$$\limsup_{n\to\infty} \frac{\mu(B(x_0, kr_n))}{\mu(B(x_0, r_n))} \le \beta_k = \frac{2^d(k+1)^d k^2}{\varepsilon}\mu(\mathbb{R}^d)$$

for all $k \in \mathbb{N}$. Consequently, with $c_n := \mu(B(x_0, r_n))^{-1}$ we have

$$\limsup_{n\to\infty} (c_n T_\#^{(x_0,r_n)}\mu)(B(0,k)) \le C_\varepsilon (k+1)^{d+2}$$

for some ε-dependent constant $C_\varepsilon > 0$. First, this shows that we may select a subsequence of the r_n's (not explicitly labeled) such that $c_n T_\#^{(x_0,r_n)}\mu \overset{*}{\rightharpoonup} \tau \in \operatorname{Tan}(\mu, x_0)$ in $\mathscr{M}_{\mathrm{loc}}(\mathbb{R}^d)$. Second, τ is non-zero since

$$\tau(\overline{B(0,1)}) \ge \limsup_{n\to\infty} (c_n T_\#^{(x_0,r_n)}\mu)(\overline{B(0,1)}) \ge 1.$$

Third,

$$\tau(B(0,k)) \le \liminf_{n\to\infty} (c_n T_\#^{(x_0,r_n)}\mu)(B(0,k)) \le C_\varepsilon (k+1)^{d+2}.$$

In conclusion, we have shown that τ is a non-zero tangent measure to μ in x_0 with polynomial growth.

The existence of a τ as above holds at μ-almost every $x_0 \in \mathbb{R}^d$ since we can make $\varepsilon > 0$ arbitrarily small, whereby the exceptional set must be μ-negligible. □

The following observation shows that we may even find blow-up sequences that converge *strictly*.

Lemma 10.6. *Let $\mu \in \mathscr{M}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$. At $|\mu|$-almost every $x_0 \in \mathbb{R}^d$ and for every open, bounded, and convex set $C \subset \mathbb{R}^d$ the following assertions hold:*

(i) There exists a tangent measure $\tau \in \operatorname{Tan}(\mu_0, x_0)$ with $|\tau|(C) = 1$, $|\tau|(\partial C) = 0$.
(ii) There exists a blow-up sequence $\gamma_n := c_n T_\#^{(x_0,r_n)}\mu$ such that $\gamma_n \to \tau$ strictly in $\mathscr{M}(\overline{C}; \mathbb{R}^N)$.

Proof. Choose $\eta > 0$, $K \in \mathbb{N}$ such that $B(0,1) \Subset \eta C \Subset B(0,K)$. At $|\mu|$-almost every point $x_0 \in \operatorname{supp}\mu$ there exists a sequence $r_n \downarrow 0$ such that

$$\limsup_{n\to\infty} \frac{|\mu|(B(x_0, Kr_n))}{|\mu|(B(x_0, r_n))} \le \beta_K \tag{10.4}$$

for a constant $\beta_K > 0$; this is proved in Proposition 10.5.

At such an x_0, let $b_n := |\mu|(B(x_0, r_n))^{-1}$. Then we have from (10.4) that, up to selecting a subsequence,

$$b_n T_\#^{(x_0,r_n)}\mu \overset{*}{\rightharpoonup} \sigma \quad \text{in } \mathscr{M}(B(0,K); \mathbb{R}^N).$$

Hence,

$$|\sigma|(\overline{\eta C}) \ge \limsup_{n\to\infty} (b_n T_\#^{(x_0,r_n)}|\mu|)(\overline{\eta C}) \ge 1.$$

From Lemma 10.4 we infer furthermore that

$$b_n T_\#^{(x_0,r_n)}|\mu| \overset{*}{\rightharpoonup} |\sigma|.$$

We can also assume $|\sigma|(\partial(\eta C)) = 0$ by increasing η slightly and without changing K since $|\sigma|$ is a finite measure, see Problem 10.1.

For the rescaled measure

$$\tau := \frac{1}{|\sigma|(\eta C)} T_\#^{(0,\eta)}\sigma = \frac{\sigma(\eta\,\bullet)}{|\sigma|(\eta C)}$$

it holds with $c_n := |\sigma|(\eta C)^{-1} b_n$ that

$$\gamma_n := c_n T_\#^{(x_0,\eta r_n)}\mu \overset{*}{\rightharpoonup} \tau \quad \text{in } \mathscr{M}(\overline{C};\mathbb{R}^N), \qquad c_n T_\#^{(x_0,\eta r_n)}|\mu| \overset{*}{\rightharpoonup} |\tau| \quad \text{in } \mathscr{M}^+(\overline{C}).$$

Since $|\tau|(\partial C) = |\sigma|(\eta C)^{-1}|\sigma|(\partial(\eta C)) = 0$, we get from standard results in measure theory (see Lemma A.22) that

$$|\gamma_n|(C) \to |\tau|(C) = 1.$$

Thus, $\tau \in \operatorname{Tan}(\mu, x_0)$ satisfies assertion (i) in the statement of the lemma and the blow-up sequence $\gamma_n = c_n T_\#^{(x_0,\eta r_n)}\mu$ satisfies assertion (ii). □

10.3 Functions of Bounded Variation

In this section we give an overview over some aspects of the theory of BV-functions. We refer to [15, 112, 176, 285] for a more thorough exposition and for proofs.

As usual, let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. The space $\mathrm{BV}(\Omega;\mathbb{R}^m)$ of functions of **bounded variation** is defined to contain all $u \in \mathrm{L}^1(\Omega;\mathbb{R}^m)$ such that there exists a matrix-valued Radon measure $Du = [Du]^i_j = [\partial_j u^i]^i_j \in \mathscr{M}(\Omega;\mathbb{R}^{m\times d})$, henceforth called the **derivative** of u, with the property that for all $i = 1,\dots,m$ and $j = 1,\dots,d$ the integration-by-parts formula

$$\int_\Omega u^i(x)\frac{\partial\psi}{\partial x_j}(x)\,\mathrm{d}x = -\int_\Omega \psi(x)\,\mathrm{d}[Du]^i_j(x), \qquad \psi \in \mathrm{C}^1_c(\Omega),$$

holds. One observes that $\mathrm{BV}(\Omega;\mathbb{R}^m)$ is a Banach space under the norm

$$\|u\|_{\mathrm{BV}(\Omega;\mathbb{R}^m)} := \|u\|_{\mathrm{L}^1(\Omega;\mathbb{R}^m)} + |Du|(\Omega),$$

but the norm topology turns out to be too strong for most purposes (for instance, smooth maps are not norm-dense in $\mathrm{BV}(\Omega;\mathbb{R}^m)$). Therefore, one is led to consider the following notions of convergence: We say that a sequence $(u_j) \subset \mathrm{BV}(\Omega;\mathbb{R}^m)$ **converges weakly*** to $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$, denoted as "$u_j \overset{*}{\rightharpoonup} u$", if

(i) $u_j \to u$ in L^1 and
(ii) $Du_j \overset{*}{\rightharpoonup} Du$ in $\mathscr{M}(\Omega;\mathbb{R}^{m\times d})$.

If we add the mass-conservation hypothesis

(iii) $|Du_j|(\Omega) \to |Du|(\Omega)$,

then the u_j are said to converge to u **strictly**, written as "$u_j \to u$ strictly"; in fact, (ii) then follows from (i) and (iii). An even stronger notion of convergence is the **area-strict convergence**, where (iii) is replaced by

(iii') $\langle Du_j\rangle(\Omega) \to \langle Du\rangle(\Omega)$.

Here, for $\mu \in \mathscr{M}(\Omega,\mathbb{R}^N)$ with Lebesgue–Radon–Nikodým decomposition $\mu = \frac{\mathrm{d}\mu}{\mathrm{d}\mathscr{L}^d}\mathscr{L}^d \llcorner \Omega + \mu^s$ (μ^s singular with respect to $\mathscr{L}^d$), the quantity $\langle\mu\rangle(\Omega)$ is the **(reduced) area-functional**, which is defined via

$$\langle\mu\rangle(B) := \int_B \sqrt{1 + \left|\frac{\mathrm{d}\mu}{\mathrm{d}\mathscr{L}^d}\right|^2}\,\mathrm{d}x + |\mu^s|(B)$$

for any Borel set $B \subset \Omega$. It can be shown that area-strict convergence implies strict convergence, see Problem 10.3 (ii). Smooth functions are area-strictly dense (hence also strictly dense) in $\mathrm{BV}(\Omega;\mathbb{R}^m)$, see Lemma 11.1 in the next chapter.

The **compactness theorem** for weak* convergence in $\mathrm{BV}(\Omega;\mathbb{R}^m)$ says that a sequence $(u_j) \subset \mathrm{BV}(\Omega;\mathbb{R}^m)$ with $\sup_j \|u_j\|_{\mathrm{BV}} < \infty$ has a weakly* converging subsequence.

The derivative Du of $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ has the Lebesgue–Radon–Nikodým decomposition

$$Du = D^a u + D^s u, \qquad D^a u = \nabla u\,\mathscr{L}^d \llcorner \Omega,$$

where the Lebesgue density $\nabla u \in \mathrm{L}^1(\Omega;\mathbb{R}^{m\times d})$ is called the **approximate gradient** of u and the measure $D^s u \in \mathscr{M}(\Omega;\mathbb{R}^{m\times d})$ is the **singular part** of Du. The latter further splits as

$$D^s u = D^j u + D^c u,$$

where $D^j u$ is the **jump part** and $D^c u$ is the **Cantor part**. The jump part is of the form

$$D^j u = (u^+ - u^-) \otimes n_{J_u}\,\mathscr{H}^{d-1} \llcorner J_u,$$

where $J_u \subset \Omega$ is the $\mathscr{H}^{d-1}$-rectifiable **jump set**, n_{J_u} is a normal on J_u, and $u^\pm$ are the one-sided traces of u on J_u (in positive and negative n_{J_u}-direction, respectively); we refer to [15] for the definition of rectifiability (which is not absolutely important for our purposes here). More precisely, J_u is the $\mathscr{H}^{d-1}$-rectifiable (Borel) set of points $x_0 \in \Omega$, where both the two one-sided traces $u^\pm(x_0)$, defined via

$$\lim_{r\downarrow 0}\frac{1}{r^d}\int_{\{x\in B(x_0,r)\,:\,x\cdot n_{J_u}\gtrless 0\}}|u(y)-u^{\pm}(x_0)|\,\mathrm{d}y=0,$$

exist, but differ.

It can be shown that if $\mathscr{H}^{d-1}(S)=0$, then also $Du(S)=0$.

For every $u\in \mathrm{BV}(\Omega;\mathbb{R}^m)$ there is an $\mathscr{L}^d$-negligible Borel set $S_u\subset\Omega$ such that every $x_0\in\Omega\setminus S_u$ is an **approximate continuity point** of u, that is, there exists a $\tilde{u}(x_0)\in\mathbb{R}^m$ such that

$$\lim_{r\downarrow 0}\frac{1}{r^d}\int_{B(x_0,r)}|u(y)-\tilde{u}(x_0)|\,\mathrm{d}y=0.$$

Consequently, S_u is called the **approximate discontinuity set** of u, and the map $\tilde{u}\colon\Omega\backslash S_u\to\mathbb{R}^m$ (which can be made Borel-measurable) is called the **precise representative** of u.

The Lebesgue density ∇u of the derivative Du can also be interpreted pointwise: The Calderón–Zygmund theorem entails that at $\mathscr{L}^d$-almost every approximate continuity point $x_0\in\Omega\setminus S_u$ the map u is **approximately differentiable**, that is,

$$\lim_{r\downarrow 0}\fint_{B(x_0,r)}\frac{|u(x)-\tilde{u}(x_0)-\nabla u(x_0)(x-x_0)|}{r}\,\mathrm{d}x=0.$$

From the Besicovitch Differentiation Theorem A.23 it further follows that

$$\nabla u(x_0)=\lim_{r\downarrow 0}\frac{Du(B(x_0,r))}{\omega_d r^d}=\lim_{r\downarrow 0}\frac{D^a u(B(x_0,r))}{\omega_d r^d}$$

at $\mathscr{L}^d$-almost every $x_0\in\Omega$. Denote by D_u the **approximate differentiability set**, i.e., the collection of all approximate differentiability points $x_0\in\Omega$ such that $\nabla u(x_0)$ satisfies the two conditions above. At all $x_0\in D_u$,

$$\lim_{r\downarrow 0}\frac{|D^s u|(B(x_0,r))}{r^d}=0. \tag{10.5}$$

On the other hand, $|D^s u|$-almost every $x_0\in\Omega$ satisfies

$$\lim_{r\downarrow 0}\frac{|Du|(B(x_0,r))}{r^d}=\infty, \tag{10.6}$$

again by the Besicovitch Differentiation Theorem A.23.

The **trace** $u|_{\partial\Omega}\colon\partial\Omega\to\mathbb{R}^m$ of $u\in\mathrm{BV}(\Omega;\mathbb{R}^m)$ on the boundary of Ω is defined for $\mathscr{H}^{d-1}$-almost every $x_0\in\partial\Omega$ via

$$\lim_{r\downarrow 0}\frac{1}{r^d}\int_{B(x_0,r)\cap\Omega}\left|u(y)-u|_{\partial\Omega}(x_0)\right|\,\mathrm{d}x=0.$$

It can be shown that $u|_{\partial\Omega} \in \mathrm{L}^1(\partial\Omega;\mathbb{R}^m)$.

We define the one-sided traces on Lipschitz subdomains $D \Subset \Omega$ in the same manner as the boundary traces. The inner and outer one-sided trace disagree on a $\mathscr{H}^{d-1}$-non-negligible subset of ∂D precisely if $|Du|(\partial D) > 0$. It is easy to see that the trace operator $u \mapsto u|_{\partial\Omega}$ is not continuous with respect to weak*-convergence in $\mathrm{BV}(\Omega;\mathbb{R}^m)$, but it is continuous with respect to the strict convergence and hence also with respect to the area-strict convergence.

We also remark that any bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ is a BV-extension domain in the sense that for every $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ there exists $\bar{u} \in \mathrm{BV}(\mathbb{R}^d;\mathbb{R}^m)$ such that $u = \bar{u}$ almost everywhere in Ω and $|D\bar{u}|(\partial\Omega) = 0$.

In BV, a **Poincaré inequality** holds: For all $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$,

$$\|u\|_{\mathrm{BV}} \leq C\biggl(|Du|(\Omega) + \int_{\partial\Omega} |u| \,\mathrm{d}\mathscr{H}^{d-1}\biggr), \tag{10.7}$$

where $C = C(\Omega) > 0$ is a domain-dependent constant, and in the last integral the values of u are to be understood in the sense of trace.

We will employ the following "gluing" procedure tacitly many times in the sequel: Given a Lipschitz subdomain $D \Subset \Omega$ as well as $u \in \mathrm{BV}(D;\mathbb{R}^m)$, $v \in \mathrm{BV}(\Omega \setminus \overline{D};\mathbb{R}^m)$, the "glued" map $w := u\mathbb{1}_D + v\mathbb{1}_{\Omega\setminus\overline{D}}$ lies in $\mathrm{BV}(\Omega;\mathbb{R}^m)$. In this case

$$Dw = Du \llcorner D + Dv \llcorner (\Omega \setminus \overline{D}) + (u|_{\partial D} - v|_{\partial(\mathbb{R}^d\setminus\overline{D})}) \otimes n_D \,\mathscr{H}^{d-1} \llcorner \partial D,$$

where n_D is the measure-theoretic unit inner normal to ∂D, that is, $n_D := \frac{\mathrm{d}D\mathbb{1}_D}{\mathrm{d}|D\mathbb{1}_D|}$.

For $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$, extended to all of $\mathbb{R}^d$, in the following we will often consider the push-forward maps $T^{(x_0,r)}_{\#}u$ under the rescaling $T^{(x_0,r)}(x) := (x - x_0)/r$ (where $x_0 \in \Omega$ and $r > 0$), that is,

$$(T^{(x_0,r)}_{\#}u)(y) := u(x_0 + ry), \qquad y \in \mathbb{R}^d.$$

Then,

$$D(T^{(x_0,r)}_{\#}u)(B) = \frac{(T^{(x_0,r)}_{\#}Du)(B)}{r^{d-1}} = \frac{Du(x_0 + rB)}{r^{d-1}} \tag{10.8}$$

for any Borel set $B \subset \Omega$, see Problem 10.2.

10.4 Structure of Singularities

Recall the decomposition of the derivative Du of $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$, namely

$$Du = \nabla u \,\mathscr{L}^d \llcorner \Omega + (u^+ - u^-) \otimes n \,\mathscr{H}^{d-1} \llcorner J_u + D^c u.$$

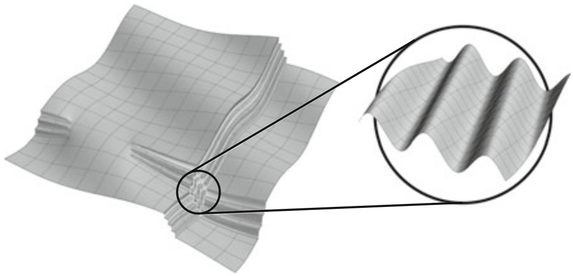


Fig. 10.2 The blow-up of a BV map at a singular point

Here, $(u^+ - u^-) \otimes n\, \mathscr{H}^{d-1} \llcorner J_u$ is the jump part of Du, as specified in the previous section.

For the jump part the local structure is clear. In particular, we have the (trivial) property that

$$\operatorname{rank}\left(\frac{\mathrm{d}D^j u}{\mathrm{d}|D^j u|}(x)\right) = 1 \qquad \text{for } |D^j u|\text{-a.e. } x \in \Omega.$$

Intuitively, "locally" around $x_0 \in J_u$ the map u is almost *one-directional* (locally, the set J_u is "almost straight").

A natural but deep question is whether the same property also holds for the Cantor part. This was conjectured by Ambrosio & De Giorgi [14] and proved first by Giovanni Alberti. The result is now usually called *Alberti's rank-one theorem.*

Theorem 10.7 (**Alberti 1993 [4]**). *Let* $u \in \mathrm{BV}(\Omega; \mathbb{R}^m)$. *Then,*

$$\operatorname{rank}\left(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\right) = 1 \qquad \textit{for } |D^s u| \textit{-a.e. } x \in \Omega.$$

As a consequence of this fundamental theorem, we get that even at points $x_0 \in \Omega$ around which $u \in \mathrm{BV}(\Omega; \mathbb{R}^m)$ has a Cantor-type (e.g. fractal) structure, the "slope" of u has a well-defined *direction*, given by the vector $n(x_0) \in \mathbb{S}^{d-1}$ from $\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x_0) = a(x_0) \otimes n(x_0)$ ($a(x_0) \in \mathbb{R}^{m \times d} \setminus \{0\}$). This is made precise in the following important consequence of Alberti's theorem.

Corollary 10.8. *Let* $u \in \mathrm{BV}(\Omega; \mathbb{R}^m)$. *Then, at* $|D^s u|$*-almost every* $x_0 \in \Omega$ *every tangent measure* $\tau \in \mathrm{Tan}(D^s u, x_0)$ *is one-directional in the sense that there exists a direction* $n \in \mathbb{S}^{d-1}$ *such that*

$$\tau(B + v) = \tau(B)$$

for all bounded Borel sets $B \subset \mathbb{R}^d$ *and all* $v \in \mathbb{R}^d$ *orthogonal to* n.

This corollary is illustrated in Figure 10.2.

Proof. By Lemma 10.4 and Alberti's Rank-One Theorem 10.7 we know that at $|D^s u|$-almost every $x_0 \in \Omega$,

$$\tau = P_0|\tau| \qquad \text{with} \qquad \operatorname{rank}\ P_0 = 1.$$

We write $P_0 = a \otimes n$ for some $a, n \in \mathbb{S}^{d-1}$. If $\gamma_n := c_n T_\#^{(x_0,r_n)} D^s u$ is a blow-up sequence for τ, then set

$$v_n(y) := r_n^{d-1} c_n u(x_0 + r_n y) + d_n, \qquad y \in R^d,$$

where the $d_n \in \mathbb{R}^m$ are chosen such that $[v_n] = \fint_{B(0,1)} v_n \, \mathrm{d}x = 0$. Then, $Dv_n = \gamma_n \overset{*}{\rightharpoonup} \tau$ in $\mathscr{M}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^{m \times d})$ and thus, by the Poincaré inequality in BV, $v_n \overset{*}{\rightharpoonup} w$ in $\mathrm{BV}_{\mathrm{loc}}$ for some $w \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^m)$, for which $Dw = \tau = P_0|\tau|$.

A mollification argument shows that we may assume that $Dw = \nabla w \mathscr{L}^d$ with a smooth, locally integrable map $w \in \mathrm{C}^\infty(\mathbb{R}^d; \mathbb{R}^m)$ such that

$$\nabla w(x) = P_0 |\nabla w(x)| \qquad \text{for a.e. } x \in \mathbb{R}^d.$$

Then we could conclude by the Ball–James Rigidity Theorem 5.13 (i) (b). However, for the sake of clarity let us write out the short argument here. For any $v \in \mathbb{R}^d$ that is orthogonal to n we have

$$\frac{\mathrm{d}}{\mathrm{d}t} w(x + tv)\bigg|_{t=0} = \nabla w(x)v = [an^T v]|\nabla w(x)| = 0.$$

This implies that w is constant in direction v. As v was an arbitrary orthogonal vector to n, $w(x)$ can only depend on $x \cdot n$. This directly yields the assertion of the lemma. □

Example 10.9. Let $S \subset \mathbb{R}^2$ be the Sierpiński triangle in the plane, which is constructed by repeatedly removing inverted maximal equilateral (solid) triangles from a given equilateral (solid) triangle, see Figure 10.3. Denote the sets in this construction by $S_0, S_1, \ldots$, whereby $S = \bigcap_{n=0}^\infty S_n$. It is shown, for instance, in Example 9.4 of [113] that S is of Hausdorff-dimension $s = \log 3/\log 2$ and $0 < \mathscr{H}^{\log 3/\log 2}(S) < \infty$. It is also easy to see that S is self-similar,

$$S = \bigcup_{n=1}^{3} \psi_n(S),$$

where the $\psi_1, \psi_2, \psi_3 \colon \mathbb{R}^2 \to \mathbb{R}^2$ are the homotheties of rate $1/2$ that keep one of the three vertices of S_0 fixed. Call S_∞ the result of covering all of $\mathbb{R}^2$ with copies of S. Define

$$\mu_S := \mathscr{H}^{\log 3/\log 2} \llcorner S_\infty$$

and observe that μ_S is not one-directional in the sense of Corollary 10.8. In fact, for the blowup-sequence $\gamma_n := c_n T_\#^{(x_0,r_n)} \mu_S$ with $r_n = 2^{-n}$ and $x_0 \in S_\infty$, we have that the γ_n are just translates of μ_S with uniformly bounded translation distances (because of the self-similarity). In particular, they are locally uniformly bounded in the total variation norm and so a subsequence converges weakly* to a tangent measure $\tau \in \mathrm{Tan}(\mu_S, x_0)$. However, this τ can only be a translation of μ_S and so

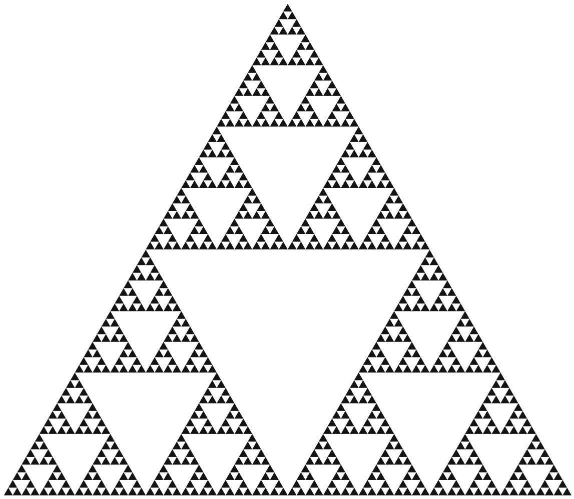


Fig. 10.3 The Sierpiński triangle

τ is not one-directional. Thus, from Corollary 10.8 we infer that there is no map $u \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^2; \mathbb{R}^m)$ for any $m \in \mathbb{N}$ with $|D^s u| = \mu_S$.

Here, we consider Alberti's rank-one theorem as a consequence of a more general result about singularities that may occur in measure solutions to linear PDEs, which is related to the compensated compactness theory presented in Section 8.8. Let $\Omega \subset \mathbb{R}^d$ be open and consider a weak solution $\mu \in \mathscr{M}(\Omega; \mathbb{R}^N)$ to the k'th-order linear constant-coefficient PDE

$$\mathscr{A}\mu := \sum_{|\alpha| \leq k} A_\alpha \partial^\alpha \mu = \sigma \qquad \text{in } \mathrm{C}_c^\infty(\Omega; \mathbb{R}^M)^*, \tag{10.9}$$

where $A_\alpha \in \mathbb{R}^{M \times N}$, $\partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d}$ for each multi-index $\alpha = (\alpha_1, \ldots, \alpha_d) \in (\mathbb{N} \cup \{0\})^d$ with length $|\alpha| := \alpha_1 + \cdots + \alpha_d \leq k$ ($k \in \mathbb{N}$), and $\sigma \in \mathscr{M}(\Omega; \mathbb{R}^M)$. This means that we require

$$(-1)^{|\alpha|} \sum_{|\alpha| \leq k} \int_\Omega (A_\alpha^T \partial^\alpha \psi) \cdot \mathrm{d}\mu = \int_\Omega \psi \cdot \mathrm{d}\sigma, \qquad \psi \in \mathrm{C}_c^\infty(\Omega; \mathbb{R}^M).$$

For $\mathscr{A}$ we define the **wave cone** (cf. the definition in Section 8.8 for first-order operators)

$$\Lambda_{\mathscr{A}} := \bigcup_{\xi \in \mathbb{S}^{d-1}} \ker \mathbb{A}^k(\xi) \subset \mathbb{R}^N \qquad \text{with} \qquad \mathbb{A}^k(\xi) := (2\pi \mathrm{i})^k \sum_{|\alpha|=k} A_\alpha \xi^\alpha \in \mathbb{R}^{M \times N},$$

where $\xi^\alpha := \xi_1^{\alpha_1} \cdots \xi_d^{\alpha_d} \in \mathbb{R}$ and $\mathbb{A}^k$ is called the **principal symbol** of $\mathscr{A}$.

We have seen in Section 8.8 that the wave cone plays a fundamental role in the study of oscillations under the constraint (10.9) (there, however, we only considered uniformly L^∞-bounded sequences instead of measures). In this section we will see that the wave cone also determines the admissible *concentrations* in measures satisfying (10.9).

The underlying idea can best be understood via a heuristic argument: Assume for simplicity that $\sigma = 0$ and that $\mathscr{A}$ is a first-order homogeneous PDE operator, i.e. $\mathscr{A} = \sum_{l=1}^d A_l \partial_l$. If $\mathscr{A}\mu = 0$, then also $\mathscr{A}\tau = 0$ for all singular tangent measures $\tau \in \operatorname{Tan}(\mu^s, x_0)$ at $|\mu^s|$-almost every point $x_0 \in \Omega$, as can be calculated easily. By Lemma 10.4,

$$\tau = P_0|\tau| \qquad \text{with} \qquad P_0 = \frac{\mathrm{d}\mu^s}{\mathrm{d}|\mu^s|}(x_0).$$

We may formally Fourier-transform the PDE $\mathscr{A}\tau = 0$ to get (here, $\mathbb{A}(\xi) = \mathbb{A}^1(\xi)$)

$$\mathbb{A}(\xi)\widehat{\tau}(\xi) = \mathbb{A}(\xi)P_0\widehat{|\tau|}(\xi) = (2\pi\mathrm{i})\sum_{l=1}^d A_l\xi_l P_0\widehat{|\tau|}(\xi) = 0 \qquad \text{for all } \xi \in \mathbb{R}^d,$$

where $\widehat{|\tau|}$ denotes the Fourier transform of $|\tau|$ (which is only defined if τ happens to be a *tempered distribution*, see [138]). Thus,

$$P_0 \in \ker \mathbb{A}(\xi) \subset \Lambda_{\mathscr{A}} \qquad \text{for all } \xi \in \mathbb{R}^d \setminus \{0\} \text{ with } \widehat{|\tau|}(\xi) \neq 0.$$

In particular, if $P_0 \notin \Lambda_{\mathscr{A}}$, then

$$\widehat{|\tau|}(\xi) = 0 \qquad \text{for all } \xi \in \mathbb{R}^d \setminus \{0\},$$

whereby $\operatorname{supp}\widehat{|\tau|} \subset \{0\}$ and hence τ is absolutely continuous with respect to $\mathscr{L}^d$. This suggests that x_0 cannot be a singular point of μ. However, by itself the absolute continuity of τ with respect to $\mathscr{L}^d$ is not a contradiction to $\tau \in \operatorname{Tan}(\mu^s, x_0)$. Indeed, Preiss provided an example of a purely singular measure that has only multiples of Lebesgue measure as tangent measures, see Example 5.9 (1) in [224]. Yet, the singular polar $\frac{\mathrm{d}\mu^s}{\mathrm{d}|\mu^s|}$ indeed only attains values in the wave cone $\Lambda_{\mathscr{A}}$:

Theorem 10.10 (De Philippis–Rindler 2016 [92]). *Let $\mu \in \mathscr{M}(\Omega;\mathbb{R}^N)$ be a solution of* (10.9) *with Lebesgue–Radon–Nikodým decomposition*

$$\mu = \mu^a + \frac{\mathrm{d}\mu^s}{\mathrm{d}|\mu^s|}|\mu^s|, \qquad \mu^s \textit{ singular with respect to } \mathscr{L}^d.$$

Then,

$$\frac{\mathrm{d}\mu^s}{\mathrm{d}|\mu^s|}(x) \in \Lambda_{\mathscr{A}} \qquad \textit{for } |\mu^s|\textit{-a.e. } x \in \Omega.$$

In the proof we will need the following two auxiliary lemmas:

Lemma 10.11. *Define the* ***Bessel potential*** *of order* $s > 0$ *via*

$$(\mathrm{id} - \Delta)^{-s/2} u := \mathscr{F}^{-1}\big[(1 + 4\pi^2 |\xi|^2)^{-s/2} \widehat{u}(\xi)\big], \qquad u \in \mathrm{C}_c^\infty(\mathbb{R}^d).$$

Then, $(\mathrm{id} - \Delta)^{-s/2}$ *extends to a compact operator from* $\mathrm{L}^1(B(0,1))$ *to* $\mathrm{L}^p(\mathbb{R}^d)$ *whenever* $1 \le p < p(d,s)$*, where*

$$p(d,s) := \begin{cases} \frac{d}{d-s} & \text{if } s < d, \\ \infty & \text{if } s \ge d. \end{cases}$$

Here, as usual, the compactness of the operator $(\mathrm{id} - \Delta)^{-s/2}$ means that it maps norm-bounded sets into precompact sets.

Proof. For $u \in \mathrm{C}_c^\infty(\mathbb{R}^d)$ we can write the application of the Bessel potential via a convolution,

$$(\mathrm{id} - \Delta)^{-s/2} u = u \star b_{s,d}, \qquad \text{where} \qquad b_{s,d} := \mathscr{F}^{-1}[(1 + 4\pi^2|\xi|^2)^{-s/2}].$$

It can be computed, see, for instance, Section 6.1.2 of [139], that $b_{s,d}$ is smooth in $\mathbb{R}^d \setminus \{0\}$ and that

$$b_{s,d}(x) \approx \begin{cases} |x|^{-d+s} & \text{if } |x| \le 2, \\ \mathrm{e}^{-|x|/2} & \text{if } |x| \ge 2. \end{cases}$$

Thus, $b_{s,d} \in \mathrm{L}^p(\mathbb{R}^d)$ for $1 \le p < p(d,s)$. Young's inequality for convolutions, see Lemma A.32, then gives for $u \in \mathrm{L}^1(B(0,1))$ that

$$(\mathrm{id} - \Delta)^{-s/2} u \in \mathrm{L}^p(\mathbb{R}^d) \qquad \text{for } 1 \le p < p(d,s).$$

For every $\varepsilon > 0$ we can furthermore write

$$b_{s,d} = b_{1,\varepsilon} + b_{2,\varepsilon} \qquad \text{with} \qquad b_{1,\varepsilon} \in \mathrm{C}_c^1(\mathbb{R}^d) \quad \text{and} \quad \|b_{2,\varepsilon}\|_{\mathrm{L}^1} < \varepsilon.$$

Indeed, for a smooth radial cut-off function $\rho \in \mathrm{C}_c^\infty(B(0,1);[0,1])$ with $\rho \equiv 1$ in a neighborhood of the origin we may set

$$b_{1,\varepsilon} := \big(\rho(\delta x) - \rho(\delta^{-1}x)\big) b_{s,d}, \qquad b_{2,\varepsilon} := \big(1 - \rho(\delta x) + \rho(\delta^{-1}x)\big) b_{s,d},$$

where $\delta \in (0,1)$ is chosen sufficiently small. Then,

$$(\mathrm{id} - \Delta)^{-s/2} u = u \star b_{1,\varepsilon} + u \star b_{2,\varepsilon} =: T_\varepsilon^{(1)} u + T_\varepsilon^{(2)} u.$$

Because $b_{1,\varepsilon} \in \mathrm{C}_c^1(\mathbb{R}^d)$, the operator $T_\varepsilon^{(1)}$ is compact from $\mathrm{L}^1(B(0,1))$ to $\mathrm{L}^1(\mathbb{R}^d)$. Moreover, again by Young's inequality for convolutions,

$$\|(\mathrm{id}-\Delta)^{-s/2} - T_\varepsilon^{(1)}\|_{\mathrm{L}^1\to\mathrm{L}^1} = \|T_\varepsilon^{(2)}\|_{\mathrm{L}^1\to\mathrm{L}^1} \le \varepsilon,$$

so that $(\mathrm{id}-\Delta)^{-s/2}$ is the limit in the uniform topology of compact operators and thus compact as well. Indeed, the last statement can be seen as follows: Let $(u_j) \subset \mathrm{L}^1(B(0,1))$ be a norm-bounded sequence. The compactness of $T_\varepsilon^{(1)}$ means that for all $\varepsilon > 0$ the sequence $(T_\varepsilon^{(1)}u_j)_j$ is precompact in $\mathrm{L}^1(\mathbb{R}^d)$. Fix a sequence $\varepsilon_k \downarrow 0$ and let $(u_{1,j})_j$ be a subsequence of $(u_{0,j})_j := (u_j)_j$ such that $(T_{\varepsilon_1}^{(1)}u_{1,j})_j$ is a Cauchy sequence in $\mathrm{L}^1(B(0,1))$. In particular, we may require that $\|T_{\varepsilon_1}^{(1)}u_{1,1} - T_{\varepsilon_1}^{(1)}u_{1,l}\|_{\mathrm{L}^1} \le \varepsilon_1$ for all $l \in \mathbb{N}$. In this manner we construct for every $k \in \mathbb{N}$ a subsequence $(u_{k,j})_j$ of $(u_{k-1,j})_{j\ge k}$ such that $(T_{\varepsilon_k}^{(1)}u_{k,j})_j$ is Cauchy and $\|T_{\varepsilon_k}^{(1)}u_{k,k} - T_{\varepsilon_k}^{(1)}u_{k,l}\|_{\mathrm{L}^1} \le \varepsilon_k$ for all $l \ge k$. Setting $v_j := u_{j,j}$, we then have for $l \ge k$ that

$$\begin{aligned}
&\|(\mathrm{id}-\Delta)^{-s/2}u_k - (\mathrm{id}-\Delta)^{-s/2}u_l\|_{\mathrm{L}^1} \\
&\quad\le \|(\mathrm{id}-\Delta)^{-s/2}u_k - T_{\varepsilon_k}^{(1)}u_k\|_{\mathrm{L}^1} + \|T_{\varepsilon_k}^{(1)}u_k - T_{\varepsilon_k}^{(1)}u_l\|_{\mathrm{L}^1} \\
&\qquad + \|T_{\varepsilon_k}^{(1)}u_l - (\mathrm{id}-\Delta)^{-s/2}u_l\|_{\mathrm{L}^1} \\
&\quad\le 2\varepsilon_k \sup_{j\in\mathbb{N}} \|u_j\|_{\mathrm{L}^1} + \varepsilon_k.
\end{aligned}$$

Thus, the sequence $((\mathrm{id}-\Delta)^{-s/2}u_j)_j$ is Cauchy in $\mathrm{L}^1(B(0,1))$ and hence $(\mathrm{id}-\Delta)^{-s/2}$ is compact.

The conclusion of the lemma now follows since strong L^1-convergence together with L^p-boundedness also implies strong L^q-convergence for all $1 \le q < p$ (this follows from Hölder's inequality). □

Lemma 10.12. *Let $(f_j) \subset \mathrm{L}^1(B(0,1))$ be such that*

(i) $f_j \overset{*}{\rightharpoonup} 0$ *in* $\mathrm{C}_c^\infty(B(0,1))^*$*;*
(ii) the negative parts $f_j^- := \max\{-f_j, 0\}$ *of the* f_j*'s converge to zero in measure, i.e.,*

$$\lim_{j\to\infty} \bigl|\bigl\{\, x \in B(0,1) \;:\; f_j^-(x) > \delta \,\bigr\}\bigr| = 0 \qquad \textit{for every } \delta > 0;$$

(iii) the family of negative parts $\{f_j^-\}$ *is equiintegrable.*

Then, $f_j \to 0$ *in* $\mathrm{L}^1_{\mathrm{loc}}(B(0,1))$.

Proof. Let $\varphi \in \mathrm{C}_c^\infty(B(0,1);[0,1])$. Then,

$$\int \varphi|f_j| \,\mathrm{d}x = \int \varphi f_j \,\mathrm{d}x + 2\int \varphi f_j^- \,\mathrm{d}x \le \int \varphi f_j \,\mathrm{d}x + 2\int f_j^- \,\mathrm{d}x.$$

The first term on the right-hand side vanishes as $j \to \infty$ by assumption (i). Vitali's Convergence Theorem A.11 in conjunction with assumptions (ii) and (iii) further gives that the second term also tends to zero in the limit. □

Proof of Theorem 10.10. We will only show the theorem for the first-order homogeneous equation with zero on the right-hand side, i.e.,

$$\mathscr{A}\mu = \sum_{l=1}^{d} A_l \partial_l \mu = 0 \qquad \text{in } \mathrm{C}_c^\infty(\Omega; \mathbb{R}^M)^*$$

since this is the only case that is needed in the sequel. See [92] for the full proof. In this situation,

$$\Lambda_{\mathscr{A}} = \bigcup_{\xi \in \mathbb{S}^{d-1}} \ker \mathbb{A}(\xi), \qquad \mathbb{A}(\xi) = (2\pi \mathrm{i}) \sum_{l=1}^{d} A_l \xi_l.$$

Step 1. Assume to the contrary that there exists a point $x_0 \in \Omega$ and a sequence $r_n \downarrow 0$ such that

(a) $\displaystyle\lim_{n\to\infty} \frac{|\mu^a|(B(x_0, r_n))}{|\mu^s|(B(x_0, r_n))} = 0;$

(b) $\displaystyle\lim_{n\to\infty} \fint_{B(x_0,r_n)} \left| \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x_0) - \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x) \right| \mathrm{d}|\mu^s|(x) = 0;$

(c) there exists a positive Radon measure $\tau \in \mathscr{M}^+(B(0,1))$ with $\tau(B(0,1/2)) > 0$ and

$$\gamma_n := c_n T_\#^{(x_0,r_n)} |\mu^s| \overset{*}{\rightharpoonup} \tau \quad \text{in } \mathscr{M}(B(0,1)), \qquad c_n := |\mu^s|(B(x_0, r_n))^{-1};$$

(d) for the polar vector it holds that

$$P_0 := \frac{\mathrm{d}\mu^s}{\mathrm{d}|\mu^s|}(x_0) \notin \Lambda_{\mathscr{A}}$$

and hence there is a positive constant $c > 0$ such that $|\mathbb{A}(\xi)P_0| \geq c|\xi|$ for all $\xi \in \mathbb{R}^d$.

Indeed, (a), (b) hold at $|\mu^s|$-almost every point by the Besicovitch Differentiation Theorem A.23 and the fact that $|\mu^s|$-almost every $x_0 \in \Omega$ is a Lebesgue point of $\frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}$. Assertion (c) follows since for $|\mu^s|$-almost every $x \in \Omega$ the space of tangent measures $\mathrm{Tan}(|\mu^s|, x)$ to $|\mu^s|$ at $x \in \Omega$ is non-trivial, see Proposition 10.5. Note that by (10.3) we may define the c_n as above (choose the radii r_n such that $|\mu^s|(\partial B(x_0, r_n)) = 0$) and via a rescaling argument (if necessary) we can ensure that $\tau(B(0,1/2)) > 0$. Finally, (d) is the assumption that will lead to a contradiction.

We will show below that (a)–(d) imply the following two assertions:

(I) $\tau \llcorner B(0,1/2)$ is absolutely continuous with respect to $\mathscr{L}^d$;

(II) $\displaystyle\lim_{n\to\infty} |\gamma_n - \tau|(B(0,1/2)) = 0.$

Then, we may conclude as follows: The γ_n are singular with respect to $\mathscr{L}^d$ by construction. So, by (I) we may find Borel sets $E_n \subset B(0,1/2)$ with $\mathscr{L}^d(E_n) = \tau(E_n) = 0$ and $\gamma_n(E_n) = \gamma_n(B(0,1/2))$. Then, by (II),

$$\gamma_n(B(0,1/2)) = \gamma_n(E_n) \le |\gamma_n - \tau|(B(0,1/2)) + \tau(E_n) = |\gamma_n - \tau|(B(0,1/2)) \to 0.$$

Hence, $\tau(B(0,1/2)) = 0$, contradicting (c) above. We are thus left to prove (I) and (II).

Step 2. Since $\mathscr{A}$ is rescaling-invariant, we have for all $r > 0$ that $\mathscr{A}(T_\#^{(x_0,r)}\mu) = 0$. Therefore,

$$\mathscr{A}(P_0\gamma_n) = \mathscr{A}\big(P_0\gamma_n - c_n T_\#^{(x_0,r_n)}\mu\big). \tag{10.10}$$

Let $(\eta_\delta)_{\delta>0} \subset \mathrm{C}_c^\infty(\mathbb{R}^d)$ be a family of positive mollifiers. The total variation on open sets is lower semicontinuous, whereby

$$|\gamma_n - \tau|(B(0,1/2)) \le \liminf_{\delta\to 0} |\gamma_n \star \eta_\delta - \tau|(B(0,1/2)).$$

Thus, for every $n \in \mathbb{N}$ there exists a $\delta_n \in (0, 1/n)$ with

$$|\gamma_n - \tau|(B(0,1/2)) \le |\gamma_n \star \eta_{\delta_n} - \tau|(B(0,1/2)) + \frac{1}{n}. \tag{10.11}$$

We now convolve (10.10) with η_{δ_n}. In this way, with

$$u_n := \gamma_n \star \eta_{\delta_n} \overset{*}{\rightharpoonup} \tau \quad \text{in } \mathscr{M}_{\mathrm{loc}}(\mathbb{R}^d), \qquad V_n := \big[P_0\gamma_n - c_n T_\#^{(x_0,r_n)}\mu\big] \star \eta_{\delta_n},$$

we get

$$\mathscr{A}(P_0 u_n) = \mathscr{A}(V_n). \tag{10.12}$$

Note that u_n, V_n are smooth and $u_n \ge 0$ since our family of mollifiers was chosen to be positive. Let $\rho \in \mathrm{C}_c^\infty(B(0,3/4);[0,1])$ be a cut-off function with $\rho \equiv 1$ on $B(0,1/2)$. Then, from (10.12) we get

$$\mathscr{A}(P_0\rho u_n) = \rho\mathscr{A}(P_0 u_n) + \mathscr{A}(P_0\rho)u_n = \mathscr{A}(\rho V_n) + R_n, \tag{10.13}$$

where

$$R_n := \mathscr{A}(P_0\rho)u_n - \sum_{l=1}^d A_l V_n \partial_l \rho \in \mathrm{C}_c^\infty(B(0,1);\mathbb{R}^M).$$

For $\delta_n \le 1/4$ it holds that

$$\begin{aligned}
\int_{B(0,3/4)} |V_n| \,\mathrm{d}x &\le c_n \big| P_0\, T_\#^{(x_0,r_n)}|\mu^s| - T_\#^{(x_0,r_n)}\mu \big|(B(0,1)) \\
&\le \frac{|P_0|\mu^s| - \mu^s|(B(x_0,r_n))}{|\mu^s|(B(x_0,r_n))} + \frac{|\mu^a|(B(x_0,r_n))}{|\mu^s|(B(x_0,r_n))} \\
&= \fint_{B(x_0,r_n)} \left| \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x_0) - \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x) \right| \mathrm{d}|\mu^s|(x) + \frac{|\mu^a|(B(x_0,r_n))}{|\mu^s|(B(x_0,r_n))}.
\end{aligned}$$

From (a), (b) we deduce that

$$\lim_{n\to\infty} \int_{B(0,3/4)} |V_n| \, \mathrm{d}x = 0. \tag{10.14}$$

Hence, the remainder terms R_n are uniformly L^1-bounded.

Step 3. We now Fourier transform (10.13) to obtain

$$[\mathbb{A}(\xi)P_0]\widehat{\rho u_n}(\xi) = \mathbb{A}(\xi)\widehat{\rho V_n}(\xi) + \widehat{R}_n(\xi), \qquad \xi \in \mathbb{R}^d.$$

Multiplying on the left by $[\mathbb{A}(\xi)P_0]^* = \overline{[\mathbb{A}(\xi)P_0]^T}$ and adding $\widehat{\rho u_n}(\xi)$ to both sides of the above equation yields

$$(1 + |\mathbb{A}(\xi)P_0|^2)\widehat{\rho u_n}(\xi) = [\mathbb{A}(\xi)P_0]^*\mathbb{A}(\xi)\,\widehat{\rho V_n}(\xi) + \widehat{\rho u_n}(\xi) + [\mathbb{A}(\xi)P_0]^*\widehat{R}_n(\xi)$$

or, equivalently,

$$\begin{aligned}\widehat{\rho u_n}(\xi) &= \frac{[\mathbb{A}(\xi)P_0]^*\mathbb{A}(\xi)}{1+|\mathbb{A}(\xi)P_0|^2}\widehat{\rho V_n}(\xi) \\ &\quad + \frac{1+4\pi^2|\xi|^2}{1+|\mathbb{A}(\xi)P_0|^2}\cdot\frac{1}{1+4\pi^2|\xi|^2}\widehat{\rho u_n}(\xi) \\ &\quad + \frac{(1+4\pi^2|\xi|^2)^{1/2}[\mathbb{A}(\xi)P_0]^*}{1+|\mathbb{A}(\xi)P_0|^2}\cdot\frac{1}{(1+4\pi^2|\xi|^2)^{1/2}}\widehat{R}_n(\xi).\end{aligned}$$

Thus,

$$\begin{aligned}\rho u_n &= T_0[\rho V_n] + \bigl(T_1\circ(\mathrm{id}-\Delta)^{-1}\bigr)[\rho u_n] + \bigl(T_2\circ(\mathrm{id}-\Delta)^{-1/2}\bigr)[R_n] \\ &=: f_n + g_n + h_n\end{aligned} \tag{10.15}$$

with

$$\begin{aligned}T_0[V] &:= \mathscr{F}^{-1}\bigl[m_0(\xi)\widehat{V}(\xi)\bigr], & m_0(\xi) &= \frac{[\mathbb{A}(\xi)P_0]^*\mathbb{A}(\xi)}{1+|\mathbb{A}(\xi)P_0|^2}, \\ T_1[u] &:= \mathscr{F}^{-1}\bigl[m_1(\xi)\widehat{u}(\xi)\bigr], & m_1(\xi) &= \frac{1+4\pi^2|\xi|^2}{1+|\mathbb{A}(\xi)P_0|^2}, \\ T_2[R] &:= \mathscr{F}^{-1}\bigl[m_2(\xi)\widehat{R}(\xi)\bigr], & m_2(\xi) &= \frac{(1+4\pi^2|\xi|^2)^{1/2}[\mathbb{A}(\xi)P_0]^*}{1+|\mathbb{A}(\xi)P_0|^2},\end{aligned}$$

and

$$(\mathrm{id}-\Delta)^{-s/2}w = \mathscr{F}^{-1}\bigl[(1+4\pi^2|\xi|^2)^{-s/2}\widehat{w}(\xi)\bigr].$$

By the *ellipticity estimate*

$$|\mathbb{A}(\xi)P_0| \geq c|\xi| \qquad \text{for } \xi \in \mathbb{R}^d$$

from (d) above, T_0 is an operator associated with a Mihlin Fourier multiplier, see Theorem A.35. The weak-type estimate from the said theorem in conjunction with (10.14) gives

$$\sup_{t \geq 0} t \Big| \big\{ x \in \mathbb{R}^d \;:\; |f_n(x)| > t \big\} \Big| \leq C \|\rho V_n\|_{\mathrm{L}^1} \to 0. \tag{10.16}$$

By Lemma 10.11, for every $s > 0$ the operator $(\mathrm{id} - \Delta)^{-s/2}$ is compact from $\mathrm{L}^1(B(0,1))$ to $\mathrm{L}^p(\mathbb{R}^d)$ for $1 \leq p < p(d,s)$ and the operators T_1, T_2 are bounded from L^p to L^p by the Mihlin Multiplier Theorem A.35 and (d) again. Thus, the family

$$g_n + h_n = \big(T_1 \circ (\mathrm{id} - \Delta)^{-1}\big)[\rho u_n] + \big(T_2 \circ (\mathrm{id} - \Delta)^{-1/2}\big)[R_n]$$

is precompact in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$. Furthermore, since $\rho u_n \geq 0$, we conclude from (10.15) that

$$f_n^- := \max\{-f_n, 0\} \leq |g_n + h_n|.$$

Thus, the family $\{f_n^-\}$ is precompact in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$, hence equiintegrable by the Dunford–Pettis Theorem A.12. Now apply Lemma 10.12, using also (10.16), and, via (10.14),

$$\big\langle f_n, \varphi \big\rangle = \big\langle T_0[\rho V_n], \varphi \big\rangle = \big\langle \rho V_n, T_0^*[\varphi] \big\rangle \to 0 \qquad \text{for every } \varphi \in \mathrm{C}_c^\infty(\mathbb{R}^d),$$

where T_0^* is the adjoint of T_0. Consequently, $f_n \to 0$ in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$. We conclude from (10.15) that the family $\{\rho u_n\}_n$ is precompact in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$ and that

$$\rho u_n \to \rho \tau \quad \text{in } \mathrm{L}^1(\mathbb{R}^d),$$

in particular $\rho\tau \in \mathrm{L}^1(\mathbb{R}^d)$. This immediately yields (I) and, combining with (10.11), also (II). □

The proof of Alberti's rank-one theorem then follows by simple linear algebra:

Proof of Theorem 10.7. Let $u \in \mathrm{BV}(\Omega; \mathbb{R}^m)$. Then, one sees directly from the definition of weak derivatives that

$$\partial_i [Du]^k_j = \partial_i \partial_j u^k = \partial_j \partial_i u^k = \partial_j [Du]^k_i \qquad \text{for all } i, j = 1, \ldots, d;\, k = 1, \ldots, m,$$

where all derivatives are to be interpreted in the sense of $\mathrm{C}_c^\infty(\Omega)^*$. We showed in the proof of Corollary 8.31 that for (with an obvious abuse of notation)

$$\mathscr{A}\mu := \operatorname{curl} \mu := \big(\partial_j \mu^k_i - \partial_i \mu^k_j\big)^{k=1,\ldots,m}_{i,j=1,\ldots,d}$$

we have

$$\Lambda_{\mathrm{curl}} = \bigcup_{\xi \in \mathbb{S}^{d-1}} \ker \mathbb{A}^1(\xi) = \left\{ a \otimes \xi \in \mathbb{R}^{m \times d} \ : \ a \in \mathbb{R}^m, \xi \in \mathbb{S}^{d-1} \right\}.$$

Consequently, the assertion of Theorem 10.7 follows immediately from Theorem 10.10. □

10.5 Convexity at Singularities

The following result about positively 1-homogeneous and rank-one convex functions is very useful since in conjunction with Alberti's rank-one theorem it implies that at singularities we can work with *convexity*.

Theorem 10.13 (**Kirchheim–Kristensen 2016 [162]**). *Let* $h^\infty \colon \mathbb{R}^{m \times d} \to \mathbb{R}$ *be positively* 1*-homogeneous and rank-one convex. Then,* h^∞ *is convex at every matrix* $F \in \mathbb{R}^{m \times d}$ *with* $\operatorname{rank} F \leq 1$*, i.e., there exists a linear function* $a_F \colon \mathbb{R}^{m \times d} \to \mathbb{R}$ *with*

$$h^\infty(F) = a_F(F) \quad \text{and} \quad h^\infty \geq a_F.$$

Recall that the positive 1-homogeneity of h^∞ means that $h^\infty(\alpha A) = \alpha h^\infty(A)$ for all $A \in \mathbb{R}^{m \times d}$, $\alpha \geq 0$.

In the language of Section 3.5 the preceding result says that $\partial h^\infty(F) \neq \emptyset$ at every matrix $F \in \mathbb{R}^{m \times d}$ with $\operatorname{rank} F \leq 1$ (here we exceptionally also use the definition of subdifferential for non-convex functions). We also record that as an immediate consequence of this theorem, for all probability measures $\mu \in \mathscr{M}^1(\mathbb{R}^{m \times d})$ with the property that the barycenter $[\mu] = \langle \mathrm{id}, \mu \rangle$ is a matrix of rank at most one, the *Jensen-type inequality*

$$h^\infty([\mu]) \leq \int h^\infty \, \mathrm{d}\mu \tag{10.17}$$

holds.

To prove Theorem 10.13, we first show two auxiliary results.

Lemma 10.14. *Let* $g \colon \mathbb{R}^{m \times d} \to \mathbb{R}$ *be positively* 1*-homogeneous and rank-one convex. If* $X, Y \in \mathbb{R}^{m \times d}$ *with* $\operatorname{rank} X \leq 1$*, then*

$$g(X + Y) \leq g(X) + g(Y).$$

Proof. Observe by the rank-one convexity and positive 1-homogeneity of g that for all $t \geq 1$ we have

$$g(X + Y) - g(Y) \leq \frac{g(tX + Y) - g(Y)}{t} = g\left(X + \frac{Y}{t}\right) - \frac{g(Y)}{t}.$$

Since g is (globally) Lipschitz by Lemma 5.6, the right-hand side converges to $g(X)$ as $t \to \infty$, which proves the claim. □

Lemma 10.15. *Let $h^\infty\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ be positively 1-homogeneous and rank-one convex. Then, for every matrix $A \in \mathbb{R}^{m\times d}$ with* rank $A \le 1$ *there exists a positively 1-homogeneous and rank-one convex* ***subcone*** $g\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ *to h^∞ at A, that is,*

$$h^\infty(A) + g(B - A) \le h^\infty(B) \qquad \textit{for all } B \in \mathbb{R}^{m\times d}. \tag{10.18}$$

Proof. Throughout the proof of the lemma we assume without loss of generality that $h^\infty(A) = 0$. Indeed, we may add a linear function to h^∞ to achieve this.

Step 1. We first show that for all $B \in \mathbb{R}^{m\times d}$ with $B \neq A$ and all $\theta \in (0, 1)$ it holds that

$$\frac{h^\infty(A + \theta(B - A))}{\theta} \le h^\infty(B). \tag{10.19}$$

Applying Lemma 10.14 for $g = h^\infty$, $X = (\theta^{-1} - 1)A$, $Y = B$, we get

$$\frac{h^\infty(A + \theta(B - A))}{\theta} = h^\infty\biggl(\frac{A}{\theta} + (B - A)\biggr) \le h^\infty\bigl((\theta^{-1} - 1)A\bigr) + h^\infty(B) = h^\infty(B),$$

which is (10.19).

Step 2. Define for $s \ge 1$ the function

$$g_s(M) := sh^\infty\biggl(A + \frac{M}{s}\biggr), \qquad M \in \mathbb{R}^{m\times d},$$

which satisfies $g_s(0) = 0$. Moreover, for $M_1, M_2 \in \mathbb{R}^{m\times d}$ we have

$$|g_s(M_1) - g_s(M_2)| \le sL\biggl|\frac{M_1}{s} - \frac{M_2}{s}\biggr| = L|M_1 - M_2|,$$

where $L > 0$ is the Lipschitz constant of h^∞, which is finite by Lemma 5.6. So, the functions g_s are uniformly Lipschitz continuous.

Now let $t \ge s$. From (10.19) with $B := A + s^{-1}M$ and $\theta := s/t$ we get

$$g_t(M) = th^\infty\biggl(A + \frac{M}{t}\biggr) \le sh^\infty\biggl(A + \frac{M}{s}\biggr) = g_s(M).$$

Thus, $g_s(M)$ is monotonically decreasing in s for every fixed $M \in \mathbb{R}^{m\times d}$. Define

$$g(M) := \lim_{s\to\infty} g_s(M) = \inf_{s>0} g_s(M), \qquad M \in \mathbb{R}^{m\times d}.$$

As the pointwise limit of uniformly Lipschitz continuous rank-one convex functions, g is also Lipschitz continuous and rank-one convex. Moreover, g is positively 1-homogeneous since for all $\alpha \ge 0$ and $M \in \mathbb{R}^{m\times d}$ it holds that

$$\begin{aligned} g(\alpha M) &= \lim_{s\to\infty}\left[sh^\infty\left(A+\frac{\alpha M}{s}\right)\right] \\ &= \alpha \lim_{s\to\infty}\left[\frac{s}{\alpha}h^\infty\left(A+\frac{M}{s/\alpha}\right)\right] \\ &= \alpha g(M). \end{aligned}$$

Finally, for $M := B - A$, $B \in \mathbb{R}^{m\times d}$, we get

$$g(B-A) \le g_1(B-A) = h^\infty(B).$$

Thus, (10.18) holds and we have proved the lemma. □

Proof of Theorem 10.13. We assume without loss of generality that $F = 0$, whereby also $h^\infty(F) = h^\infty(0) = 0$. Let $A_1, \dots, A_{md}$ be an orthonormal basis of $\mathbb{R}^{m\times d}$ consisting of rank-one matrices (for instance, the collection $\{\mathrm{e}_i \otimes \mathrm{e}_j\}_{i,j}$). First, we construct a sequence

$$h^\infty =: g_{md} \ge g_{md-1} \ge \cdots \ge g_0$$

of positively 1-homogeneous rank-one convex functions $g_k\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ as follows: Set $g_{md} := h^\infty$, Then, given g_k, we inductively define $g_{k-1}\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ to be a subcone to g_k at A_k as in Lemma 10.15. From (10.18) we get with $h^\infty := g_k$, $g := g_{k-1}$, $A := A_k$, $B := A_k + M$, where $M \in \mathbb{R}^{m\times d}$, and Lemma 10.14 with $g := g_k$, $X := A_k$, $Y := M$ that

$$g_k(A_k) + g_{k-1}(M) \le g_k(A_k + M) \le g_k(A_k) + g_k(M),$$

Hence $g_k \ge g_{k-1}$.

Next, with all the g_k constructed, define linear maps

$$L_k\colon \operatorname{span}\{A_1, \dots, A_k\} \to \mathbb{R}$$

iteratively according to the following procedure: Let $L_0\colon \{0\} \to \mathbb{R}$ be the trivial linear map. At step k we assume

$$L_k \le g_k \quad \text{on } \operatorname{span}\{A_1, \dots, A_k\} \qquad \text{and} \qquad L_k(A_k) = g_k(A_k), \tag{10.20}$$

which is certainly true for $k = 0$. Then set

$$L_{k+1}(\alpha A_{k+1} + B) := \alpha g_{k+1}(A_{k+1}) + L_k(B), \qquad B \in \operatorname{span}\{A_1, \dots, A_k\},\ \alpha \in \mathbb{R}.$$

By (10.20) and the fact that g_k is a subcone to g_{k+1} at A_{k+1},

$$\begin{aligned} L_{k+1}(A_{k+1}+B) &= g_{k+1}(A_{k+1}) + L_k(B) \\ &\le g_{k+1}(A_{k+1}) + g_k(B) \\ &\le g_{k+1}(A_{k+1}+B). \end{aligned}$$

Since g_{k+1} is positively 1-homogeneous, this further implies for $\alpha \ge 0$ that

$$\begin{aligned} L_{k+1}(\alpha A_{k+1}+B) &= \alpha L_{k+1}(A_{k+1}+\alpha^{-1}B) \\ &\le \alpha g_{k+1}(A_{k+1}+\alpha^{-1}B) \\ &= g_{k+1}(\alpha\, A_{k+1}+B). \end{aligned}$$

For $\alpha \le 0$ we have by Lemma 10.14 with $g := g_{k+1}$, $X := (1-\alpha)A_{k+1}$, $Y := \alpha A_{k+1} + B$,

$$\begin{aligned} g_{k+1}(A_{k+1}) + L_k(B) &= L_{k+1}(A_{k+1}+B) \\ &\le g_{k+1}(A_{k+1}+B) \\ &\le g_{k+1}((1-\alpha)A_{k+1}) + g_{k+1}(\alpha A_{k+1}+B) \\ &= (1-\alpha)g_{k+1}(A_{k+1}) + g_{k+1}(\alpha A_{k+1}+B). \end{aligned}$$

Thus, by definition of L_{k+1}, also in the case $\alpha \le 0$ it holds that

$$L_{k+1}(\alpha A_{k+1}+B) = \alpha g_{k+1}(A_{k+1}) + L_k(B) \le g_{k+1}(\alpha A_{k+1}+B)$$

and we have established (10.20) at $k+1$. Then set $a_0 := L_{md} \le g_{md} = h^\infty$, for which $a_0(0) = L_{md}(0) = 0 = h^\infty(0)$. □

Notes and Historical Remarks

The Reshetnyak Continuity Theorem 10.3 was first proved in [226], whereas our argument is from [15]. The paper [94] contains more on different notions of convergence for measures.

Often, tangent measures are only considered as defined on the unit ball instead of the whole of the space (see, for example, Section 2.7 in [15]). This is, however, sometimes restrictive. Here, we use Preiss's original theory as developed in [224], also see Chapter 14 of [183]. The original definition, however, explicitly excluded the zero-measure from $\mathrm{Tan}(\mu, x_0)$, which here we include. Proposition 10.5 is a slight improvement of Preiss' existence theorem for non-zero tangent measures, see Theorem 2.5 in [224] or the appendix of [227]. The proof of Lemma 10.4 is adapted from Theorem 2.44 in [15]. Lemma 10.6 is originally due to Larsen, see Lemma 5.1 in [174].

We remark that the study of local properties of a measure via its tangent measures has its limits. Most strikingly, Preiss constructed a purely singular positive measure on a bounded interval (in particular a BV-derivative) such that each of its tangent measures is a multiple of Lebesgue measure, see Example 5.9 (1) in [224]. Also see [219] for a measure that has *every* local measure as a tangent measure at almost every point.

The notion of area-strict convergence in BV seems to be somewhat less well known than it deserves. However, as shown in the next chapter, it is the right one when considering integral functionals.

Alberti's original proof [4] of what is now called Alberti's rank-one theorem is via a "decomposition technique" together with the BV-coarea formula; a streamlined version of his proof can be found in [90]. Another proof in two dimensions was announced in [5]. There is also now a nice short geometric proof [181]. Our argument is more in the spirit of PDE theory and has several other implications, see [92]. For more on the wave cone we refer to [98, 209, 210, 229, 267, 268].

The Kirchheim–Kristensen Theorem 10.13 was already announced in 2011 [161] with a simpler proof in a special case. The full proof appeared in [162]. The theorem actually holds in more generality, see Problem 10.10.

Problems

10.1. Let $\Lambda \in \mathscr{M}^+(B(0,1))$ be a finite and positive Borel measure. Then, we have $\Lambda(\partial B(0,r)) = 0$ for all but finitely many $r \in (0,1)$. *Hint:* Consider the sets $E_n := \{ r \in (0,1) : \Lambda(\partial B(0,r)) > 1/n \}$, where $n \in \mathbb{N}$.

10.2. Prove (10.8) using the integration-by-parts definition of BV-functions.

10.3. Let $\Omega \subset \mathbb{R}^d$ be open.

(i) Show *Reshetnyak's lower semicontinuity theorem*: Let $f \colon \Omega \times \mathbb{S}^{N-1} \to [0,\infty]$ be lower semicontinuous and convex in the second argument. For any sequence $(\mu_j) \subset \mathscr{M}(\Omega;\mathbb{R}^N)$ with $\mu_j \overset{*}{\rightharpoonup} \mu$ it holds that

$$\int f\left(x, \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x)\right) \mathrm{d}|\mu|(x) \le \liminf_{j\to\infty} \int f\left(x, \frac{\mathrm{d}\mu_j}{\mathrm{d}|\mu_j|}(x)\right) \mathrm{d}|\mu_j|(x).$$

Hint: Adapt the strategy of the proof of Theorem 10.3.

(ii) Let $(\mu_j) \subset \mathscr{M}(\Omega;\mathbb{R}^N)$ with $\mu_j \to \mu$ area-strictly, i.e., $\mu_j \overset{*}{\rightharpoonup} \mu$ and $\langle \mu_j \rangle(\Omega) \to \langle \mu \rangle(\Omega)$. Show that then also $\mu_j \to \mu$ strictly. *Hint:* If $\mu = a\mathscr{L}^d + \mu^s$ is the Lebesgue–Radon–Nikodým decomposition of μ, then define $\tilde{\mu} \in \mathscr{M}(\Omega;\mathbb{R}^N \times \mathbb{R})$ as $\tilde{\mu}(B) := (a,0)\mathscr{L}^d + (0,1)\mu^s$ and use the Reshetnyak continuity theorem.

(iii) Show that the mapping $\mu \in \mathscr{M}(\Omega;\mathbb{R}^N) \mapsto \langle \mu \rangle(\Omega)$ is weakly* lower semicontinuous.

10.4. Let $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ and assume that $x_0 \in \Omega$ is such that

$$\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x_0) = a \otimes n \qquad \text{for some } a \in \mathbb{S}^{m-1}, n \in \mathbb{S}^{d-1}.$$

Set

$$u_r(y) := \frac{r^d}{|Du|(Q_n(x_0,r))} \cdot \frac{u(x_0+ry) - [u]_{B(x_0,r)}}{r}, \qquad y \in Q_n(x_0,r),\ r>0,$$

where $Q_n(x_0,r)$ is an open cube with midpoint $x_0 \in \Omega$, side-length r, and two faces orthogonal to n. Then, show that for $|D^s u|$-almost every such x_0 it holds that $u_r \overset{*}{\rightharpoonup} u_0$ in $\mathrm{BV}(Q_n(0,1);\mathbb{R}^m)$ with $u_0(y) = a\psi(y \cdot n)$ for a bounded function $\psi\colon (-1/2,1/2) \to \mathbb{R}$. Further show that ψ can be chosen to be *increasing*. *Hint:* Use Corollary 10.8.

10.5. Let $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$. Show that for $|Du|$-almost every $x_0 \in \Omega$, every $\tau \in \mathrm{Tan}(Du, x_0)$ is a BV-derivative, $\tau = Dv$ for some $v \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^m)$, and with $P_0 := \frac{\mathrm{d}Du}{\mathrm{d}|Du|}(x_0)$ it holds that:

(i) If rank $P_0 \geq 2$, then $v(x) = v_0 + \alpha\, P_0 x$, where $\alpha \in \mathbb{R}$, $v_0 \in \mathbb{R}^m$.
(ii) If $P_0 = a \otimes n$ ($a \in \mathbb{R}^m$, $n \in \mathbb{S}^{d-1}$), then there exist $h \in \mathrm{BV}(\mathbb{R})$, $v_0 \in \mathbb{R}^m$ such that $v(x) = v_0 + h(x \cdot n)a$.

What does Alberti's Rank-One Theorem imply about (i)? This problem is continued in Problem 12.7.

10.6. Let $\Omega \subset \mathbb{R}^d$ be an open set and let $u \in \mathrm{L}^1(\Omega;\mathbb{R}^m)$ with the property that the r'th (weak) derivative $D^r u \in \mathcal{M}(\Omega; \mathrm{SLin}^r(\mathbb{R}^d;\mathbb{R}^m))$ exists for some $r \in \mathbb{N}$, where $\mathrm{SLin}^r(\mathbb{R}^d;\mathbb{R}^m)$ contains all symmetric r-linear maps from $\mathbb{R}^d$ to $\mathbb{R}^m$. Then show that for $|(D^r u)^s|$-almost every $x \in \Omega$ there exist $a(x) \in \mathbb{R}^m \setminus \{0\}$, $n(x) \in \mathbb{S}^{d-1}$ such that

$$\frac{\mathrm{d}(D^r u)^s}{\mathrm{d}|(D^r u)^s|}(x) = a(x) \otimes \underbrace{n(x) \otimes \cdots \otimes n(x)}_{r \text{ times}}.$$

Here, the tensor on the right-hand side is the r-linear map

$$V(v_0, v_1, \ldots, v_r) := (a(x)\cdot v_0)\prod_{i=1}^{r}(n(x)\cdot v_i), \qquad v_0 \in \mathbb{R}^m,\ v_1,\ldots,v_r \in \mathbb{R}^d.$$

This is Alberti's theorem for higher-order gradients [4]. *Hint:* Use Theorem 10.10.

10.7. Let $\mu \in \mathcal{M}(\Omega;\mathbb{R}^{d\times d})$ be a matrix-valued measure such that (in the distributional, i.e. $\mathrm{C}_c^\infty(\Omega)^*$-weak sense)

$$\operatorname{div} \mu \in \mathcal{M}(\Omega;\mathbb{R}^d).$$

Prove that then

$$\operatorname{rank}\left(\frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x)\right) \le d-1 \qquad \text{for } |\mu^s|\text{-a.e. } x \in \Omega.$$

10.8. A map $u \in \mathrm{L}^1(\Omega;\mathbb{R}^d)$ ($\Omega \subset \mathbb{R}^d$ a bounded Lipschitz domain) is called a function of **bounded deformation** if the symmetric part of its distributional derivative is a measure,

$$Eu := \frac{Du + (Du)^T}{2} \in \mathscr{M}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}}).$$

We collect all these functions into the set $\mathrm{BD}(\Omega)$; see [12, 273, 274] for the theory of this space. Let $\mu = (\mu^k_j) \in \mathscr{M}(\Omega,\mathbb{R}^{d\times d}_{\mathrm{sym}})$ be the symmetrized derivative of some $u \in \mathrm{BD}(\Omega)$, $\mu = Eu$. Verify that then

$$0 = \mathscr{A}\mu := \left[\sum_{i=1}^{d} \partial_i\partial_k\mu^j_i + \partial_i\partial_j\mu^k_i - \partial_j\partial_k\mu^i_i - \partial_i\partial_i\mu^k_j\right]_{j,k=1,\dots,d}$$

in the sense of $\mathrm{C}^\infty_c(\Omega)^*$. These equations are often called the *Saint-Venant compatibility conditions* in applications.

10.9. In the situation of the previous problem and denoting the Lebesgue–Radon–Nikodým decomposition of the symmetrized derivative Eu of $u \in \mathrm{BD}(\Omega)$ by

$$Eu = \mathscr{E}u\,\mathscr{L}^d \llcorner \Omega + E^s u,$$

prove that for $|E^s u|$-almost every $x \in \Omega$, there exist $a(x), b(x) \in \mathbb{R}^d \setminus \{0\}$ such that

$$\frac{\mathrm{d}E^s u}{\mathrm{d}|E^s u|}(x) = a(x) \odot b(x),$$

where we define the symmetrized tensor product as $a \odot b := (a \otimes b + b \otimes a)/2$ for $a, b \in \mathbb{R}^d$. *Hint:* Use Theorem 10.10 and the previous problem. This result was one of the main motivations for Theorem 10.10.

10.10. Assume that $D \subset \mathbb{R}^N$ is a *balanced cone*, that is, for $v \in D$ and $t \in \mathbb{R}$ also $tv \in D$, and further assume that $\operatorname{span} D = \mathbb{R}^N$. Let $h^\infty \colon \mathbb{R}^N \to \mathbb{R}$ be positively 1-homogeneous and convex in all directions in D. Show that h^∞ is convex at every $F \in D$, i.e., there exists an affine function $a_F \colon \mathbb{R}^N \to \mathbb{R}$ with

$$h^\infty(F) = a_F(F) \qquad \text{and} \qquad h^\infty \ge a_F.$$

This is (almost) the main result of [162].

Chapter 11
Linear-Growth Functionals

After the preparations in the previous chapter, we now return to the task at hand, namely to analyze the following minimization problem for an integral functional with *linear growth*:

$$\begin{cases} \text{Minimize} & \mathscr{F}[u] := \displaystyle\int_\Omega f(x, \nabla u(x)) \,\mathrm{d}x \\ \text{over all} & u \in \mathrm{W}^{1,1}(\Omega; \mathbb{R}^m) \text{ with } u|_{\partial\Omega} = g. \end{cases}$$

As usual, we assume $\Omega \subset \mathbb{R}^d$ to be a bounded Lipschitz domain and $f : \Omega \times \mathbb{R}^{m \times d} \to \mathbb{R}$ to be a Carathéodory integrand. We also suppose that there exists a constant $M > 0$ such that

$$|f(x, A)| \le M(1 + |A|), \qquad (x, A) \in \Omega \times \mathbb{R}^{m \times d}.$$

Furthermore, we let $g \in \mathrm{L}^1(\partial\Omega; \mathbb{R}^m)$.

As we have seen in the introduction to the previous chapter, $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$ is too small to describe concentration effects. Our first task is thus to extend $\mathscr{F}$ in a natural way to a functional $\widetilde{\mathscr{F}}$ defined on $\mathrm{BV}(\Omega; \mathbb{R}^m)$ such that this extension is continuous with respect to a notion of convergence in which $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$ is dense in $\mathrm{BV}(\Omega; \mathbb{R}^m)$. As we will see shortly, this extension turns out to be

$$\widetilde{\mathscr{F}}[u] := \int_\Omega f(x, \nabla u(x)) \,\mathrm{d}x + \int_\Omega f^\infty\left(x, \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\right) \mathrm{d}|D^s u|(x),$$

where f^∞ is the *(strong) recession function*, defined in (11.8) below, which we assume to exist as a continuous function. The notion of convergence in which $\widetilde{\mathscr{F}}$ is continuous and in which $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$ is dense in $\mathrm{BV}(\Omega; \mathbb{R}^m)$ can be found in the area-strict convergence, which is weaker than convergence in norm, but stronger

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8_11

than weak* convergence in $\mathrm{BV}(\Omega;\mathbb{R}^m)$. In particular, we conclude that the infimum of $\mathscr{F}$ over $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ and the infimum of $\widetilde{\mathscr{F}}$ over $\mathrm{BV}(\Omega;\mathbb{R}^m)$ agree.

If we assume coercivity of f, then we obtain a uniform $\mathrm{W}^{1,1}$-norm bound on any minimizing sequence for $\mathscr{F}$. Invoking the compactness theorem for the weak* convergence in $\mathrm{BV}(\Omega;\mathbb{R}^m)$, we select a (not explicitly labeled) subsequence of our minimizing sequence $(u_j) \subset \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ with $u_j \overset{*}{\rightharpoonup} u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$. One observes that any minimizing sequence $(u_j) \subset \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ for $\mathscr{F}$ is still minimizing for $\widetilde{\mathscr{F}}$ (we could even choose $(u_j) \subset \mathrm{C}^\infty(\Omega;\mathbb{R}^m)$ by a mollification argument). We thus need to investigate the weak* lower semicontinuity in $\mathrm{BV}(\Omega;\mathbb{R}^m)$ of our extended functional $\widetilde{\mathscr{F}}$.

In this chapter, we first prove a result concerning the extension of our functional $\mathscr{F}$ to $\widetilde{\mathscr{F}}$. Then we consider the questions of lower-semicontinuity and relaxation. Here, we focus on the "classical" approach to these questions, whereas the next chapter present a more abstract perspective.

11.1 Extension of Functionals

Recall that for a sequence $(u_j) \subset \mathrm{BV}(\Omega;\mathbb{R}^m)$ we say that $u_j \to u$ area-strictly if $u_j \to u$ in L^1, $Du_j \overset{*}{\rightharpoonup} Du$ in $\mathscr{M}(\Omega;\mathbb{R}^{m\times d})$, and $\langle Du_j\rangle(\Omega) \to \langle Du\rangle(\Omega)$, where $\langle\,\bullet\,\rangle$ is the (reduced) area functional, which we defined in Section 10.3.

Lemma 11.1. *Let $\Omega \subset \mathbb{R}^d$ be a bounded open set. For each $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ there exists a sequence $(v_j) \subset (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(\Omega;\mathbb{R}^m)$ with $v_j|_{\partial\Omega} = u|_{\partial\Omega}$ and $v_j \to u$ area-strictly. If $u \in \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ we may additionally require that $v_j \to u$ in $\mathrm{W}^{1,1}$.*

Note that for this lemma we do not assume any boundary regularity on Ω (which is sometimes useful). See Problem 11.1 for a simpler proof in the case when Ω is a Lipschitz domain.

Proof. Let $\varepsilon > 0$ and take a collection of open sets $\Omega_i \subset \Omega$, where $i \in \mathbb{N} \cup \{0\}$, with

$$\Omega_i \Subset \Omega_{i+1} \qquad \text{and} \qquad \bigcup_{i=0}^{\infty} \Omega_i = \Omega.$$

Furthermore, we assume

$$|Du|(\Omega \setminus \Omega_0) < \varepsilon.$$

Let $U_0 := \Omega_1$ and set

$$U_i := \Omega_{i+1} \setminus \overline{\Omega_{i-1}} \qquad \text{for } i \in \mathbb{N}.$$

Then, the U_i, $i \in \mathbb{N} \cup \{0\}$, form an open cover of Ω such that every point of Ω lies in at most two of the sets U_i. Let $\{\rho_i\}_{i\in\mathbb{N}\cup\{0\}} \subset \mathrm{C}^\infty(\Omega;[0,1])$ be a smooth partition of unity subordinate to the cover $(U_i)_i$. In particular, $\rho_i \in C_c^\infty(U_i; [0, 1])$ and $\sum_{i=0}^\infty \rho_i = \mathbb{1}_\Omega$. The construction of such partitions of unity is detailed, for instance, in Lemma 2.3.1 of [285].

Now take a family of mollifiers $(\eta_\delta)_{\delta>0} \in C_c^\infty(B(0,1))$. Then, for each $i \in \mathbb{N} \cup \{0\}$ choose $\delta_i > 0$ such that

$$\operatorname{supp}\bigl(\eta_{\delta_i} \star (\rho_i u)\bigr) \subset U_i, \tag{11.1}$$

$$\int_\Omega |\eta_{\delta_i} \star (\rho_i u) - \rho_i u| \,\mathrm{d}x < \frac{\varepsilon}{2^{i+1}}, \tag{11.2}$$

$$\int_\Omega |\eta_{\delta_i} \star (u \otimes \nabla \rho_i) - u \otimes \nabla \rho_i| \,\mathrm{d}x < \frac{\varepsilon}{2^{i+1}}, \tag{11.3}$$

and

$$\int_\Omega \rho_i (\eta_{\delta_i} \star \langle \rho_i Du \rangle) \,\mathrm{d}x \le \int_\Omega \rho_i \,\mathrm{d}\langle \rho_i Du \rangle + \varepsilon \fint_\Omega \rho_i \,\mathrm{d}x, \tag{11.4}$$

where $\rho_i u$ and $u \otimes \nabla \rho_i$ are understood to be extended by zero to all of $\mathbb{R}^d$. Let

$$v_\varepsilon := \sum_{i=0}^\infty \eta_{\delta_i} \star (\rho_i u).$$

Then, $v_\varepsilon \in C^\infty(\Omega)$ and, since $u = \sum_{i=0}^\infty \rho_i u$, we get from (11.2) that

$$\int_\Omega |v_\varepsilon - u| \,\mathrm{d}x \le \sum_{i=0}^\infty \int_\Omega |\rho_i u - \eta_{\delta_i} \star (\rho_i u)| \,\mathrm{d}x < \varepsilon. \tag{11.5}$$

By the product rule,

$$\nabla v_\varepsilon = \sum_{i=0}^\infty \eta_{\delta_i} \star (\rho_i Du) + \sum_{i=0}^\infty \bigl[\eta_{\delta_i} \star (u \otimes \nabla \rho_i) - u \otimes \nabla \rho_i\bigr]$$

since $\sum_{i=0}^\infty \nabla \rho_i = 0$. For $x \in U_i$ at most two terms in each sum are non-zero, hence

$$\nabla v_\varepsilon = \eta_{\delta_i} \star (\rho_i Du) + E_i \qquad \text{on } U_i,$$

where E_i is an error term, which we may estimate using (11.3) as

$$\sum_{i=0}^\infty \int_{U_i} |E_i| \,\mathrm{d}x \le 2|Du|(\Omega \setminus \overline{\Omega_0}) + \varepsilon < 3\varepsilon. \tag{11.6}$$

Here we also used that $\rho_0 \equiv 1$ on $\overline{\Omega_0}$.

For $\langle A\rangle := \sqrt{1+|A|^2}$ the elementary estimate $\langle A+B\rangle \le \langle A\rangle + |B|$ holds for all $A, B \in \mathbb{R}^{m\times d}$. Thus,

$$\int_\Omega \rho_i \langle \nabla v_\varepsilon\rangle \,\mathrm{d}x \le \int_\Omega \rho_i \langle \eta_{\delta_i} \star (\rho_i Du)\rangle \,\mathrm{d}x + \sum_{i=0}^\infty \int_{U_i} |E_i| \,\mathrm{d}x.$$

Also, by Jensen's inequality and (11.4),

$$\begin{aligned}\int_\Omega \rho_i \langle \eta_{\delta_i} \star (\rho_i Du)\rangle \,\mathrm{d}x &\le \int_\Omega \rho_i (\eta_{\delta_i} \star \langle \rho_i Du\rangle) \,\mathrm{d}x \\ &\le \int_\Omega \rho_i \,\mathrm{d}\langle \rho_i Du\rangle + \varepsilon \fint_\Omega \rho_i \,\mathrm{d}x \\ &\le \int_\Omega \rho_i \,\mathrm{d}\langle Du\rangle + \varepsilon \fint_\Omega \rho_i \,\mathrm{d}x.\end{aligned}$$

Combining the last two estimates, summing over $i \in \mathbb{N} \cup \{0\}$, and employing (11.6), we arrive at

$$\langle Dv_\varepsilon\rangle(\Omega) \le \langle Du\rangle(\Omega) + 4\varepsilon.$$

Since the (reduced) area functional $\langle\bullet\rangle(\Omega)$ is lower semicontinuous with respect to the weak* convergence of measures (see Problem 10.3 (iii)), we furthermore have

$$\langle Du\rangle(\Omega) \le \liminf_{\varepsilon\downarrow 0} \langle Dv_\varepsilon\rangle(\Omega) \le \limsup_{\varepsilon\downarrow 0} \langle Dv_\varepsilon\rangle(\Omega) \le \langle Du\rangle(\Omega).$$

Hence, $\langle Dv_\varepsilon\rangle(\Omega) \to \langle Du\rangle(\Omega)$ as $\varepsilon \downarrow 0$. As we have already established that $\|u - v_\varepsilon\|_{\mathrm{L}^1} < \varepsilon$, see (11.5), this shows the area-strict convergence of v_ε to u.

If $u \in \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$, then in the preceding argument we choose δ_i so that in addition to (11.1)–(11.4) also

$$\int_{U_i} \bigl|\nabla(\rho_i u) - \nabla(\eta_{\delta_i} \star (\rho_i u))\bigr| \,\mathrm{d}x < \frac{\varepsilon}{2^{i+1}}$$

holds for each $i \in \mathbb{N} \cup \{0\}$. Then,

$$\int_\Omega |\nabla u - \nabla v_\varepsilon| \,\mathrm{d}x < \varepsilon$$

and the strong convergence $v_\varepsilon \to u$ in $\mathrm{W}^{1,1}$ follows.

It remains to show that $v_\varepsilon \in \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ with $v_\varepsilon|_{\partial\Omega} = u|_{\partial\Omega}$, which will complete the proof. By construction, $v_\varepsilon \in \mathrm{W}^{1,1}(D;\mathbb{R}^m)$ for all Lipschitz subdomains $D \Subset \Omega$. We will prove that

$$w := \begin{cases} v_\varepsilon - u & \text{on } \Omega, \\ 0 & \text{on } \mathbb{R}^d \setminus \overline{\Omega}, \end{cases}$$

lies in $\mathrm{BV}(\mathbb{R}^d;\mathbb{R}^m)$ and that $|Dw|(\partial\Omega) = 0$. Take any $\psi \in \mathrm{C}_c^\infty(\mathbb{R}^d)$. From (11.1) we derive for $j = 1, \ldots, d$ and $k = 1, \ldots, m$ that

$$\begin{aligned} \int_{\mathbb{R}^d} w^k \frac{\partial \psi}{\partial x_j} \, \mathrm{d}x &= \int_\Omega [v_\varepsilon - u]^k \frac{\partial \psi}{\partial x_j} \, \mathrm{d}x \\ &= \sum_{i=0}^\infty \int_\Omega \bigl[\eta_{\delta_i} \star (\rho_i u) - \rho_i u\bigr]^k \frac{\partial \psi}{\partial x_j} \, \mathrm{d}x \\ &= \sum_{i=0}^\infty \biggl[\int_{U_i} \rho_i \psi \, \mathrm{d}[Du]_j^k - \int_{U_i} \eta_{\delta_i} \star (\rho_i [Du]_j^k) \psi \, \mathrm{d}x\biggr] \\ &\qquad + \sum_{i=0}^\infty \int_{U_i} \psi \biggl[u^k \frac{\partial \rho_i}{\partial x_j} - \eta_{\delta_i} \star \Bigl(u^k \frac{\partial \rho_i}{\partial x_j}\Bigr)\biggr] \, \mathrm{d}x. \end{aligned} \tag{11.7}$$

Combining this with (11.3), we arrive at

$$\left| \int_{\mathbb{R}^d} w^k \frac{\partial \psi}{\partial x_j} \, \mathrm{d}x \right| \le C\bigl(|Du|(\Omega) + 1\bigr) \|\psi\|_\infty.$$

Thus, $w \in \mathrm{BV}(\mathbb{R}^d;\mathbb{R}^m)$. Next, assume that ψ is supported near $\partial\Omega$, say $\psi \equiv 0$ on U_i for all $i < i_0$. Then, by (11.7) we get

$$\begin{aligned} \int_{\mathbb{R}^d} w^k \frac{\partial \psi}{\partial x_j} \, \mathrm{d}x &= \sum_{i=i_0}^\infty \biggl[\int_{U_i} \rho_i \psi \, \mathrm{d}[Du]_j^k - \int_{U_i} \eta_{\delta_i} \star (\rho_i [Du]_j^k) \psi \, \mathrm{d}x\biggr] \\ &\qquad + \sum_{i=i_0}^\infty \int_{U_i} \psi \biggl[u^k \frac{\partial \rho_i}{\partial x_j} - \eta_{\delta_i} \star \Bigl(u^k \frac{\partial \rho_i}{\partial x_j}\Bigr)\biggr] \, \mathrm{d}x, \end{aligned}$$

and so, letting $i_0 \to \infty$, we conclude that $|Dw|(\partial\Omega) = 0$. □

If we want to determine the extension of the functional $\mathscr{F}$ from $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ to $\mathrm{BV}(\Omega;\mathbb{R}^m)$, we will need to deal with concentrations. In order to understand the behavior of the (linear-growth) integrand f "at infinity", we define the following notions: Let $f\colon \Omega \times \mathbb{R}^N \to \mathbb{R}$ be a Carathéodory integrand with linear growth. The **(strong) recession function** $f^\infty\colon \overline{\Omega} \times \mathbb{R}^N \to \mathbb{R}$ is

$$f^\infty(x, A) := \lim_{\substack{x' \to x \\ A' \to A \\ t \to \infty}} \frac{f(x', tA')}{t}, \qquad (x, A) \in \overline{\Omega} \times \mathbb{R}^N, \tag{11.8}$$

if this limit exists (in $\mathbb{R}$). The function f^∞ is easily seen to be positively 1-homogeneous in A, that is,

$$f^\infty(x, \alpha A) = \alpha f^\infty(x, A) \qquad \text{for all } \alpha \geq 0,\ (x, A) \in \overline{\Omega} \times \mathbb{R}^N.$$

If f is Lipschitz continuous in its second argument and the Lipschitz constant is uniform with respect to x, then the definition of f^∞ simplifies to

$$f^\infty(x, A) = \lim_{\substack{x' \to x \\ t \to \infty}} \frac{f(x', tA)}{t}.$$

Clearly, the recession function f^∞ does not necessarily exist (not even for quasiconvex f, see Theorem 2 of [200]). However, for f with linear growth, we can always define the **upper weak recession function** $f^\#\colon \overline{\Omega} \times \mathbb{R}^N \to \mathbb{R}$ and the **lower weak recession function** $f_\#\colon \overline{\Omega} \times \mathbb{R}^N \to \mathbb{R}$, respectively, by

$$f^\#(x, A) := \limsup_{\substack{x' \to x \\ A' \to A \\ t \to \infty}} \frac{f(x', tA')}{t}, \qquad f_\#(x, A) := \liminf_{\substack{x' \to x \\ A' \to A \\ t \to \infty}} \frac{f(x', tA')}{t},$$

where $x \in \overline{\Omega}$, $A \in \mathbb{R}^{m \times d}$. We remark that here we employed the usual definition of lim sup, lim inf for functions on metric spaces, i.e.,

$$f^\#(x, A) = \lim_{\varepsilon \downarrow 0} \sup \left\{ \frac{f(x', tA')}{t} : 0 < |x' - x| < \varepsilon,\ 0 < |A' - A| < \varepsilon,\ t > \frac{1}{\varepsilon} \right\},$$
$$f_\#(x, A) = \lim_{\varepsilon \downarrow 0} \inf \left\{ \frac{f(x', tA')}{t} : 0 < |x' - x| < \varepsilon,\ 0 < |A' - A| < \varepsilon,\ t > \frac{1}{\varepsilon} \right\}.$$

It is immediate that $f^\#$, $f_\#$ are finite and positively 1-homogeneous in their second argument. If f is Lipschitz continuous in its second argument and the Lipschitz constant is uniform with respect to x, then

$$f^\#(x, A) = \limsup_{\substack{x' \to x \\ t \to \infty}} \frac{f(x', tA)}{t}, \qquad f_\#(x, A) = \liminf_{\substack{x' \to x \\ t \to \infty}} \frac{f(x', tA)}{t},$$

where $x \in \overline{\Omega}$, $A \in \mathbb{R}^N$.

We remark that by the rank-one convexity in conjunction with Alberti's Rank-One Theorem 10.7, one may replace the upper limit in the definition of $h^\#$ by a proper limit at matrices of rank at most one.

The extension question for $\mathscr{F}$ is then settled by the following result.

Theorem 11.2. *Let $f\colon \Omega \times \mathbb{R}^{m\times d} \to \mathbb{R}$ be a continuous integrand with linear growth and such that the strong recession function f^∞ exists. Then, the area-strictly continuous extension of the functional*

$$\mathscr{F}[u] := \int_\Omega f(x, \nabla u(x))\,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m),$$

to $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ is given by

$$\widetilde{\mathscr{F}}[u] := \int_\Omega f(x, \nabla u(x))\,\mathrm{d}x + \int_\Omega f^\infty\biggl(x, \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\biggr)\,\mathrm{d}|D^s u|(x),$$

where

$$Du = \nabla u\,\mathscr{L}^d + \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}\,|D^s u|$$

is the Lebesgue–Radon–Nikodým decomposition of Du. In particular, $\widetilde{\mathscr{F}}$ is area-strictly continuous.

Remark 11.3. It turns out that the preceding result remains valid with the upper weak recession function $f^\#$ in place of f^∞ if f is continuous and rank-one convex (or rank-one concave) and moreover we assume

$$f^\#(x, A) = f_\#(x, A) = (f(x, \bullet))^\#(A) = \limsup_{\substack{A' \to A\\ t\to\infty}} \frac{f(x, tA')}{t}$$

for all $(x, A) \in \Omega \times \mathbb{R}^{m\times d}$ such that $\operatorname{rank} A \le 1$. The proof of this extension is the task of Problem 11.5; also see Problem 11.6 for a condition on how to verify the above equality between recession functions.

Remark 11.4. A version of the preceding theorem also holds true for a class of u-dependent integrands, see [231].

Proof of Theorem 11.2. We will prove the general statement that for any sequence $(\mu_j) \subset \mathscr{M}(\Omega;\mathbb{R}^N)$ with $\mu_j \to \mu$ **area-strictly**, i.e., $\mu_j \overset{*}{\rightharpoonup} \mu$ and $\langle \mu_j\rangle(\Omega) \to \langle \mu\rangle(\Omega)$, it holds that

$$\begin{aligned}\int_\Omega f\biggl(x, \frac{\mathrm{d}\mu_j}{\mathrm{d}\mathscr{L}^d}(x)\biggr)\,\mathrm{d}x + \int_\Omega f^\infty\biggl(x, \frac{\mathrm{d}\mu_j^s}{\mathrm{d}|\mu_j^s|}(x)\biggr)\,\mathrm{d}|\mu_j^s|(x)\\ \to \int_\Omega f\biggl(x, \frac{\mathrm{d}\mu}{\mathrm{d}\mathscr{L}^d}(x)\biggr)\,\mathrm{d}x + \int_\Omega f^\infty\biggl(x, \frac{\mathrm{d}\mu^s}{\mathrm{d}|\mu^s|}(x)\biggr)\,\mathrm{d}|\mu^s|(x).\end{aligned}$$

This can be seen as an extension of Reshetnyak's continuity theorem. This assertion together with Lemma 11.1 then immediately implies the theorem.

To see the claim, for an integrand f as in the statement of the theorem define the **perspective integrand** $Pf\colon \Omega \times \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}$ by

$$Pf(x, A, s) := \begin{cases} sf(x, s^{-1}A) & \text{if } s \neq 0, \\ f^\infty(x, A) & \text{if } s = 0, \end{cases} \qquad (x, A, s) \in \Omega \times \mathbb{R}^N \times \mathbb{R}.$$

Clearly, Pf is continuous and $(A, s) \mapsto f(x, A, s)$ is positively 1-homogeneous (jointly in (A, s)) for fixed $x \in \Omega$. Likewise, for $\mu \in \mathcal{M}(\Omega; \mathbb{R}^N)$ we define $P^*\mu \in \mathcal{M}(\Omega; \mathbb{R}^N \times \mathbb{R})$ via

$$P^*\mu(B) := \big(\mu(B), \mathscr{L}^d(B)\big) \qquad \text{for any Borel set } B \subset \Omega.$$

With the Lebesgue–Radon–Nikodým decomposition $\mu = a\mathscr{L}^d + \mu^s$ of μ we have

$$P^*\mu = (a, 1)\,\mathscr{L}^d + (1, 0)\,\mu^s.$$

Then,

$$|P^*\mu|(\Omega) = \int_\Omega |(a, 1)|\,\mathrm{d}x + |\mu^s|(\Omega) = \langle\mu\rangle(\Omega)$$

and hence the area-strict convergence of a sequence $(\mu_j) \subset \mathcal{M}(\Omega; \mathbb{R}^N)$ is equivalent to the strict convergence of $(P^*\mu_j) \subset \mathcal{M}(\Omega; \mathbb{R}^N \times \mathbb{R})$. Thus, by Reshetnyak's continuity theorem, the mapping

$$\mu \mapsto \mathscr{G}[\mu] := \int_\Omega Pf\left(x, \frac{\mathrm{d}P^*\mu}{\mathrm{d}|P^*\mu|}(x)\right) \mathrm{d}|P^*\mu|(x)$$

is area-strictly continuous. Since

$$\mathscr{G}[\mu] = \int_\Omega f\left(x, \frac{\mathrm{d}\mu}{\mathrm{d}\mathscr{L}^d}(x)\right) \mathrm{d}x + \int_\Omega f^\infty\left(x, \frac{\mathrm{d}\mu^s}{\mathrm{d}|\mu^s|}(x)\right) \mathrm{d}|\mu^s|(x),$$

the claim follows. □

Example 11.5. Let $w \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$ be the (shifted) staircase function

$$w(x) := \left\lfloor x - \frac{1}{2} \right\rfloor, \qquad x \in \mathbb{R}.$$

For the sequence $(u_j) \subset \mathrm{BV}(0, 1)$ defined as $u_j(x) := w(jx)/j$ for $x \in (0, 1)$, we have $u_j \overset{*}{\rightharpoonup} u$ with $u(x) = x$ and also

$$|Du_j|((0, 1)) = \langle Du_j\rangle((0, 1)) = 1, \qquad |Du|((0, 1)) = 1, \qquad \langle Du\rangle((0, 1)) = \sqrt{2},$$

whereby $u_j \to u$ strictly but the u_j do not converge area-strictly to u. Thus, the area-strict convergence cannot be replaced with mere strict convergence in the above theorem since strict continuity fails for $\widetilde{\mathscr{F}} := \langle\bullet\rangle(\Omega)$ (with integrand $f(A) := \sqrt{1 + |A|^2}$, which has the strong recession function $f^\infty(A) = |A|$).

Example 11.6. We also observe that the requirement that the strong recession function exists cannot be dispensed with in general: On $\Omega := (-1, 1)$ and $\mathbb{R}^{m\times d} = \mathbb{R}$ define $f(a) := |a| \sin a$, $a \in \mathbb{R}$. Then,

$$f^{\#}(\pm 1) = \limsup_{t\to\infty} \frac{f(\pm t)}{t} = 1, \qquad f_{\#}(\pm 1) = \liminf_{t\to\infty} \frac{f(\pm t)}{t} = -1.$$

Now set

$$u_j(x) := \beta_j x \mathbb{1}_{(0,\beta_j^{-1})}(x) + \mathbb{1}_{(\beta_j^{-1},1)}(x), \qquad \text{where} \qquad \beta_j := 2\pi j - \pi/2.$$

Denote by $\widetilde{\mathscr{F}}^{\#}$ the functional $\widetilde{\mathscr{F}}$ with f^{∞} replaced by $f^{\#}$. We compute

$$u_j' = \beta_j \mathbb{1}_{(0,\beta_j^{-1})} \to \delta_0 \quad \text{area-strictly,}$$

whereby $u_j \to \mathbb{1}_{(0,1)}$ area-strictly in BV, but

$$\mathscr{F}[u_j] = \widetilde{\mathscr{F}}^{\#}[u_j] = \int_0^{\beta_j^{-1}} \beta_j \sin \beta_j \,\mathrm{d}x = -1 \neq 1 = \widetilde{\mathscr{F}}^{\#}[\mathbb{1}_{(0,1)}].$$

Likewise, one can convince oneself with a similar example that also with $f_{\#}$ we do not get the extension property.

11.2 Lower Semicontinuity

For a Carathéodory integrand $f\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ with linear growth we now (re)define

$$\mathscr{F}[u] := \int_{\Omega} f(\nabla u(x)) \,\mathrm{d}x + \int_{\Omega} f^{\#}\left(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\right) \mathrm{d}|D^s u|(x), \qquad u \in \mathrm{BV}(\Omega; \mathbb{R}^m).$$

Note that here we use the upper weak recession function, which always exists. If even the strong recession function f^{∞} exists, then by Theorem 11.2 the $\mathscr{F}$ so defined is the area-strictly continuous extension of our original $\mathscr{F}$ (defined on $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$). We will show in the next section that the above $\mathscr{F}$, however, is always the *relaxation* of the extended-valued functional

$$\mathscr{F}^{\infty}[u] := \begin{cases} \displaystyle\int_{\Omega} f(\nabla u(x)) \,\mathrm{d}x & \text{if } u \in \mathrm{W}^{1,1}(\Omega; \mathbb{R}^m), \\ +\infty & \text{if } u \in (\mathrm{BV} \setminus \mathrm{W}^{1,1})(\Omega; \mathbb{R}^m). \end{cases}$$

Thus, our choice of extension for $\mathscr{F}$ is justified even if the strong recession function f^{∞} does not exist.

The following is one of the most well-known lower semicontinuity theorems in BV.

Theorem 11.7 (Ambrosio–Dal Maso 1992 & Fonseca–Müller 1993 [13, 124]). *Assume that* $f\colon \mathbb{R}^{m\times d} \to [0,\infty)$ *is a continuous and quasiconvex integrand with linear growth. Then,* $\mathscr{F}$ *is weakly* lower semicontinuous on the space* $\mathrm{BV}(\Omega;\mathbb{R}^m)$.

Hence, under a suitable coercivity hypothesis a minimization problem for $\mathscr{F}$ has a solution in $\mathrm{BV}(\Omega;\mathbb{R}^m)$:

Theorem 11.8. *Let* $f\colon \mathbb{R}^{m\times d} \to [0,\infty)$ *be a continuous integrand satisfying the coercivity and linear growth estimate*

$$\mu|A| \le f(A) \le M(1+|A|) \qquad A \in \mathbb{R}^{m\times d},$$

for some $\mu, M > 0$*. If* f *is quasiconvex, then the associated functional* $\mathscr{F}$ *has a minimizer over the space* $\mathrm{BV}(\Omega;\mathbb{R}^m)$.

This theorem follows directly from the lower semicontinuity result via the Direct Method and the usual coercivity arguments, in particular the Poincaré inequality in BV, see (10.7). We postpone the investigation into boundary conditions to Corollary 11.17 below.

In the following we only prove Theorem 11.7 under the additional assumption that the strong recession function f^∞ exists, see Remark 11.18 below for the general case.

We will analyze $\mathscr{F}$ through the "blow-up" behavior of the auxiliary functional

$$\mathscr{G}[u;D] := \inf\Big\{ \mathscr{F}[w;D] \;:\; w \in \mathrm{BV}(D;\mathbb{R}^m) \text{ with } w|_{\partial D} = u|_{\partial D} \Big\}, \tag{11.9}$$

where $D \subset \Omega$ is a Lipschitz subdomain of Ω and

$$\mathscr{F}[u;D] := \int_D f(\nabla u(x))\,\mathrm{d}x + \int_D f^\#\left(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\right)\mathrm{d}|D^s u|(x)$$

for $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$. A crucial step in the proof of Theorem 11.7 will be to show that $\mathscr{G}[u;D]$ and $\mathscr{F}[u;D]$ are close in value if D is a small ball or cube.

One may easily prove the transformation rule

$$\mathscr{G}\left[y \mapsto \frac{u(x_0+ry)}{r};D\right] = \frac{1}{r^d}\mathscr{G}[u;x_0+rD] \tag{11.10}$$

for all $x_0 \in \Omega$ and $r > 0$. An easy way to see this is by approximating u with smooth maps via Lemma 11.1 and employing a change of variables.

We start with the following simple BV-gluing lemma.

Lemma 11.9. *Let* $U, V \subset \mathbb{R}^d$ *be bounded Lipschitz domains with* $U \Subset V$ *and let* $u \in \mathrm{BV}(U;\mathbb{R}^m)$*,* $v \in \mathrm{BV}(V \setminus \overline{U};\mathbb{R}^m)$*. For*

$$w := u\mathbb{1}_U + v\mathbb{1}_{V\setminus\overline{U}} \in \mathrm{BV}(U;\mathbb{R}^m)$$

there exists a sequence $(w_j)_j \subset (\mathrm{W}^{1,1}\cap\mathrm{C}^\infty)(V;\mathbb{R}^m)$ *that converges area-strictly to* w *on* V*. Moreover, for all continuous linear-growth integrands* $f\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ *such that the strong recession function* f^∞ *exists, it holds that*

$$\lim_{j\to\infty} \mathscr{F}[w_j; V] = \mathscr{F}[u; U] + \mathscr{F}[v; V\setminus\overline{U}] + \int_{\partial U} f^\infty\left(\frac{u-v}{|u-v|}\otimes n_U\right)|u-v| \,\mathrm{d}\mathscr{H}^{d-1},$$

where the values of u *and* v *on* ∂U *are to be understood as one-sided traces (that is,* $u = u|_{\partial U}$ *and* $v = v|_{\partial(V\setminus\overline{U})}$*) and* n_U *denotes the (measure-theoretic) unit inner normal on* ∂U*.*

Proof. By the usual BV-theory (see Section 10.3), we have that $w = u\mathbb{1}_U + v\mathbb{1}_{V\setminus\overline{U}}$ lies in $\mathrm{BV}(V;\mathbb{R}^m)$ and satisfies

$$Dw = Du \llcorner U + Dv \llcorner (V\setminus\overline{U}) + (u-v)\otimes n_U\, \mathscr{H}^{d-1}\llcorner \partial U.$$

Hence, the conclusion follows immediately from Theorem 11.2 in conjunction with Lemma 11.1. □

Next, we present a characterization of quasiconvexity in BV.

Lemma 11.10. *Let* $h\colon \mathbb{R}^{m\times d}\to\mathbb{R}$ *be a Borel function with linear growth and such that the strong recession function* h^∞ *exists. Then,* h *is quasiconvex if and only if for one (hence all) bounded Lipschitz domains* $D\subset\mathbb{R}^d$*, all* $u\in\mathrm{BV}(D;\mathbb{R}^m)$*, and all affine maps* $a\colon\mathbb{R}^d\to\mathbb{R}^m$ *it holds that*

$$|D|h(\nabla a) \le \int_D h(\nabla u)\,\mathrm{d}x + \int_D h^\infty\left(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}\right)\mathrm{d}|D^s u| + \int_{\partial D} h^\infty\left(\frac{u-a}{|u-a|}\otimes n_D\right)|u-a|\,\mathrm{d}\mathscr{H}^{d-1},$$

where n_D *denotes the unit inner normal on* ∂D *and* $u = u|_{\partial D}$ *is the inner trace.*

Proof. First, we notice that for $u\in\mathrm{W}^{1,\infty}(D;\mathbb{R}^m)\subset\mathrm{BV}(D;\mathbb{R}^m)$ with $u|_{\partial D} = a|_{\partial D}$, the inequality in the lemma reads as

$$|D|h(\nabla a) \le \int_D h(\nabla u)\,\mathrm{d}x,$$

which is equivalent to the quasiconvexity of h.

Turning to the other implication, we assume that h is quasiconvex. Let $r\in(0,1)$ and define the bounded Lipschitz domain

$$D_r = \left\{ x \in \mathbb{R}^d \ : \ \mathrm{dist}(x; D) < r \right\}.$$

Take a sequence $r_n \downarrow 0$ such that

$$D \Subset D_{r_n} \qquad \text{and} \qquad \bigcap_{n\in\mathbb{N}} D_{r_n} = \overline{D}.$$

Via Lemma 11.9 with $U := D$, $V := D_{r_n}$, $u \in \mathrm{BV}(U, \mathbb{R}^m)$, and $v = a$ we find maps $w_n \in W^{1,\infty}_a(D_{r_n}; \mathbb{R}^m)$ with

$$\int_{D_{r_n}} h(\nabla w_n) \,\mathrm{d}x \leq \int_D h(\nabla u) \,\mathrm{d}x + \int_D h^\infty\left(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}\right) \mathrm{d}|D^s u| + \int_{D_{r_n}\setminus\overline{D}} h(\nabla a) \,\mathrm{d}x \\ + \int_{\partial D} h^\infty\left(\frac{u-a}{|u-a|} \otimes n_D\right)|u-a| \,\mathrm{d}\mathscr{H}^{d-1} + \frac{1}{n}.$$

The quasiconvexity of h then implies

$$|D_{r_n}| h(\nabla a) \leq \int_{D_{r_n}} h(\nabla w_n) \,\mathrm{d}x.$$

On the other hand,

$$\int_{D_{r_n}\setminus\overline{D}} h(\nabla a) \,\mathrm{d}x \leq M(1 + |\nabla a|)|D_{r_n} \setminus \overline{D}| \to 0 \qquad \text{as } n \to \infty.$$

Thus,

$$\begin{aligned} |D| h(\nabla a) &\leq \lim_{n\to\infty} \int_{D_{r_n}} h(\nabla w_n) \,\mathrm{d}x \\ &\leq \int_D h(\nabla u) \,\mathrm{d}x + \int_D h^\infty\left(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}\right) \mathrm{d}|D^s u| \\ &\quad + \int_{\partial D} h^\infty\left(\frac{u-a}{|u-a|} \otimes n_D\right)|u-a| \,\mathrm{d}\mathscr{H}^{d-1}. \end{aligned}$$

Hence, the assertion of the lemma follows. □

Next, we prove the following "maximum principle" for $\mathscr{J}$:

Lemma 11.11. *For all $u, v \in \mathrm{BV}(\Omega; \mathbb{R}^m)$ and all* convex *Lipschitz subdomains $D \subset \Omega$ it holds that*

$$\left| \mathscr{J}[u; D] - \mathscr{J}[v; D] \right| \leq M \int_{\partial D} |u - v| \,\mathrm{d}\mathscr{H}^{d-1},$$

where $M > 0$ is the linear-growth constant of f.

Proof. For reasons of notational simplicity we assume that $D = B$ is a ball, the general proof proceeds along the same lines.

Since the trace spaces for $\mathrm{BV}(B; \mathbb{R}^m)$ and $(\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(B; \mathbb{R}^m)$ are both equal to $\mathrm{L}^1(\partial B; \mathbb{R}^m)$ and since $\mathscr{J}$ only depends on boundary values, we may without loss of generality assume that $u, v \in (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(B; \mathbb{R}^m)$. Below we will establish that

$$\mathscr{J}[u; B] \le \mathscr{F}[w; B] + M \int_{\partial B} |u - v| \, \mathrm{d}\mathscr{H}^{d-1} \tag{11.11}$$

for all $w \in \mathrm{BV}(B; \mathbb{R}^m)$ with $w|_{\partial B} = v|_{\partial B}$. The conclusion of the lemma then follows by taking the infimum over all such w and also by switching the roles of u and v.

To show (11.11), we may assume that $w \in (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(B; \mathbb{R}^m)$ with $w|_{\partial B} = v|_{\partial B}$ by virtue of Theorem 11.2 in conjunction with Lemma 11.1. Let $B = B(x_0, R)$ for some $x_0 \in \mathbb{R}^d$ and $R > 0$. For $\delta \in (0, 1)$ denote the concentric subball with radius δR by $B_\delta := B(x_0, \delta R)$. Let $w_\delta \in \mathrm{BV}(B; \mathbb{R}^m)$ be defined as

$$w_\delta(x) := \begin{cases} w(x) & \text{if } x \in B_\delta, \\ u(x) & \text{if } x \in B \setminus \overline{B_\delta}, \end{cases}$$

for which $w_\delta|_{\partial B} = u|_{\partial B}$ and

$$Dw_\delta = Dw \llcorner B_\delta + Du \llcorner (B \setminus \overline{B_\delta}) + (w - u) \otimes n_{B_\delta} \, \mathscr{H}^{d-1} \llcorner \partial B_\delta,$$

where n_{B_δ} is the unit inner normal on ∂B_δ. Using the linear growth of f (with growth constant $M > 0$), we estimate

$$\begin{aligned} \mathscr{J}[u; B] &\le \mathscr{F}[w; B_\delta] + \mathscr{F}[u; B \setminus \overline{B_\delta}] \\ &\qquad + \int_{\partial B_\delta} f^\infty\Big(\frac{w - u}{|w - u|} \otimes n_{B_\delta} \Big) |w - u| \, \mathrm{d}\mathscr{H}^{d-1} \\ &\le \mathscr{F}[w; B] + M\big(2\mathscr{L}^d + |Dw| + |Du|\big)(B \setminus \overline{B_\delta}) \\ &\qquad + M \int_{\partial B_\delta} |w - u| \, \mathrm{d}\mathscr{H}^{d-1}. \end{aligned}$$

We then let $\delta \uparrow 1$, for which the second term vanishes and the surface integral tends to $\int_{\partial B} |w - u| \, \mathrm{d}\mathscr{H}^{d-1}$ (by smoothness). Thus, (11.11) follows. □

The following is a local lower semicontinuity property of $\mathscr{J}$:

Lemma 11.12. *Let $u_j \overset{*}{\rightharpoonup} u$ in $\mathrm{BV}(\Omega; \mathbb{R}^m)$ and $|Du_j| \overset{*}{\rightharpoonup} \Lambda$ in $\mathscr{M}^+(\overline{\Omega})$. Then, for every Lipschitz subdomain $D \subset \Omega$ with $\Lambda(\partial D) = 0$ it holds that*

$$\mathscr{J}[u; D] \le \liminf_{j \to \infty} \mathscr{F}[u_j; D].$$

Proof. Again we only consider the case of a ball, $D = B(x_0, R) \subset \Omega$ for some $x_0 \in \Omega$, $R > 0$. In view of Theorem 11.2 in conjunction with Lemma 11.1 we may without loss of generality suppose that $u_j \in (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(\Omega; \mathbb{R}^m)$.

Take a sequence of radii $r_k \uparrow R$ such that $\Lambda(\partial B(x_0, r_k)) = 0$ and, possibly after selecting a subsequence (not explicitly labeled),

$$\int_{\partial B(x_0,r_k)} |u_j - u| \,\mathrm{d}\mathscr{H}^{d-1} \to 0 \qquad \text{as } j \to \infty$$

for all $k \in \mathbb{N}$. The existence of such radii can be seen as follows: By a curvilinear version of Fubini's theorem (coarea formula),

$$\int_0^R \int_{\partial B(x_0,r)} |u_j - u| \,\mathrm{d}\mathscr{H}^{d-1} \,\mathrm{d}r = \int_{B(x_0,R)} |u_j - u| \,\mathrm{d}x \to 0.$$

Thus, we may select a subsequence of the u_j's such that

$$\int_{\partial B(x_0,r)} |u_j - u| \,\mathrm{d}\mathscr{H}^{d-1} \to 0 \qquad \text{for a.e. } r \in (0, R),$$

and then choose appropriate radii r_k; we remark that the set of $r \in (0, R)$ with the property $\Lambda(\partial B(x_0, r)) \neq 0$ is at most countable because Λ is a finite measure (see Problem 10.1). Due to the inequality $|Du| \leq \text{w*-lim}_{j\to\infty}|Du_j| = \Lambda$, it also holds that $|Du|(\partial B(x_0, r_k)) = 0$ and therefore the inner and outer one-sided traces on $\partial B(x_0, r_k)$ coincide,

$$u|_{\partial B(x_0,r_k)} = u|_{\partial(\Omega\setminus\overline{B(x_0,r_k)})} \quad \mathscr{H}^{d-1}\text{-a.e.}$$

We fix k, set $B := B(x_0, R)$, $B_k := B(x_0, r_k)$, and define

$$w_j(x) := \begin{cases} u_j(x) & \text{if } x \in B_k, \\ u(x) & \text{if } x \in B \setminus \overline{B_k}, \end{cases}$$

which lies in $\mathrm{BV}(B; \mathbb{R}^m)$, satisfies $w_j|_{\partial B} = u|_{\partial B}$, and

$$Dw_j = Du_j \llcorner B_k + Du \llcorner (B \setminus \overline{B_k}) + (u_j - u) \otimes n_{B_k} \mathscr{H}^{d-1} \llcorner \partial B_k.$$

From the additivity of $\mathscr{F}[u; \bullet]$ for disjoint sets and the linear growth of f with growth constant $M > 0$ we deduce that

$$\begin{aligned} \mathscr{J}[u; B] &\leq \mathscr{F}[w_j; B] \\ &= \mathscr{F}[u_j; B_k] + \mathscr{F}[u; B \setminus \overline{B}_k] \\ &\quad + \int_{\partial B_k} f^\infty\biggl(\frac{u_j - u}{|u_j - u|} \otimes n_{B_k}\biggr)|u_j - u| \,\mathrm{d}\mathscr{H}^{d-1} \end{aligned}$$

$$\le \mathscr{F}[u_j; B] + M\big(2\mathscr{L}^d + |Du| + |Du_j|\big)(B \setminus \overline{B_k}) + M \int_{\partial B_k} |u_j - u| \, \mathrm{d}\mathscr{H}^{d-1}.$$

Taking the lower limit as $j \to \infty$ (and keeping k fixed), we deduce from $\Lambda(\partial B_k) = 0$ and $|Du| \le \Lambda$ that

$$\mathscr{J}[u; B] \le \liminf_{j\to\infty} \mathscr{F}[u_j; B] + 2M(\mathscr{L}^d + \Lambda)(\overline{B} \setminus B_k).$$

Now let $k \to \infty$. The claim of the lemma follows for our subsequence of u_j's since $\Lambda(\overline{B} \setminus B_k) \to \Lambda(\partial B) = 0$.

It remains to prove the result for our original sequence (u_j). For this, select a subsequence $j(l)$ such that

$$\lim_{l\to\infty} \mathscr{F}[u_{j(l)}; B] = \liminf_{j\to\infty} \mathscr{F}[u_j; B].$$

Then, by the proof above, we get for a further subsequence (still denoted as $j(l)$) that

$$\mathscr{J}[u; B] \le \liminf_{l\to\infty} \mathscr{F}[u_{j(l)}; B] = \liminf_{j\to\infty} \mathscr{F}[u_j; B].$$

This finishes the proof of the lemma. □

Next, we investigate the fine structure of $\mathscr{F}$ around regular and singular points. We will show below that for $(\mathscr{L}^d + |Du|)$-almost every $x_0 \in \Omega$ and for every $\varepsilon > 0$ there exists an $r_0(x_0) > 0$ such that

$$\mathscr{F}[u; U(x_0, r)] \le \mathscr{J}[u, U(x_0, r)] + \varepsilon(\mathscr{L}^d + |Du|)(U(x_0, r)) \tag{11.12}$$

for almost all $r \in (0, r_0(x_0))$, where $U(x_0, r) \subset \Omega$ is an open convex set that contains x_0 and satisfies

$$(\mathscr{L}^d + |Du| + \Lambda)(\partial U(x_0, r)) = 0, \qquad B(x_0, r) \subset U(x_0, r) \subset B(x_0, \kappa r)$$

for a fixed constant $\kappa \ge 1$ and Λ from above. In fact, $U(x_0, r)$ will either be the open ball $B(x_0, r)$ or a cube with side length $2r$; hence we may choose $\kappa = \sqrt{d}$.

Lemma 11.13. *The estimate (11.12) holds for $\mathscr{L}^d$-almost every point $x_0 \in \Omega$ with $U(x_0, r) = B(x_0, r)$.*

Proof. Let $x_0 \in \Omega$ be such that

(a) u is approximately differentiable at x_0 (see Section 10.3);
(b) $\displaystyle\lim_{r\downarrow 0} \frac{Du(B(x_0, r))}{\omega_d r^d} = \frac{\mathrm{d}Du}{\mathrm{d}\mathscr{L}^d}(x_0) = \nabla u(x_0)$;
(c) $\displaystyle\lim_{r\downarrow 0} \frac{|Du|(B(x_0, r))}{\omega_d r^d} = \frac{\mathrm{d}|Du|}{\mathrm{d}\mathscr{L}^d}(x_0) = |\nabla u(x_0)|$;
(d) x_0 is an $\mathscr{L}^d$-Lebesgue point of ∇u.

By the results of Section 10.3, $\mathscr{L}^d$-almost every $x_0 \in \Omega$ has these properties.

The maps

$$u_r(y) := \frac{u(x_0 + ry) - \tilde{u}(x_0)}{r}, \qquad y \in B(0,1),$$

where we denote the precise representative of u by $\tilde{u}$ (see Section 10.3), satisfy

$$u_r \to u_0 \quad \text{in } \mathrm{L}^1, \qquad Du_r \to Du_0 \quad \text{strictly} \qquad \text{as } r \downarrow 0$$

with $u_0(y) := \nabla u(x_0)y$. Indeed, the L^1-strong convergence follows from the approximate differentiability after a change of variables. The strict convergence of Du_r to Du_0 can be seen by using (10.8) and calculating

$$\lim_{r\downarrow 0} |Du_r|(B(0,1)) = \omega_d \lim_{r\downarrow 0} \frac{|Du|(B(x_0,r))}{|B(x_0,r)|} = \omega_d |\nabla u(x_0)| = |Du_0|(B(0,1)).$$

From Lemma 5.6 we know that the integrand f is (globally) Lipschitz continuous in the second argument, say with Lipschitz constant $L > 0$. Since x_0 is a Lebesgue point of ∇u and since $r^{-d}|D^s u|(B(x_0,r)) \to 0$ by the approximate differentiability of u in x_0 (see (10.5)), we get

$$\begin{aligned}
&\big|\mathscr{F}[u; B(x_0,r)] - \mathscr{F}[u_0; B(x_0,r)]\big| \\
&\quad \le L \int_{B(x_0,r)} |\nabla u(y) - \nabla u(x_0)| \,\mathrm{d}y + M|D^s u|(B(x_0,r)) \\
&\quad \le \frac{\varepsilon}{2}|B(x_0,r)|
\end{aligned}$$

for $r > 0$ sufficiently small. Lemma 11.10 (i.e., quasiconvexity) thus implies for all $w \in \mathrm{BV}(B(x_0,r);\mathbb{R}^m)$ with $w = u_0$ on $\partial B(x_0,r)$ that

$$\begin{aligned}
\mathscr{F}[u; B(x_0,r)] &\le \int_{B(x_0,r)} f(\nabla u_0) \,\mathrm{d}x + \frac{\varepsilon}{2}|B(x_0,r)| \\
&\le \mathscr{F}[w; B(x_0,r)] + \frac{\varepsilon}{2}|B(x_0,r)|.
\end{aligned}$$

Taking the infimum over all such w, we arrive at

$$\mathscr{F}[u; B(x_0,r)] \le \mathscr{J}[u_0; B(x_0,r)] + \frac{\varepsilon}{2}|B(x_0,r)|. \tag{11.13}$$

On the other hand,

$$\mathscr{J}[u_0; B(x_0,r)] = r^d \mathscr{J}[u_0; B(0,1)], \qquad \mathscr{J}[u; B(x_0,r)] = r^d \mathscr{J}[u_r; B(0,1)],$$

see (11.10). We thus infer from Lemma 11.11 together with the strict convergence of the Du_r to Du_0 and the strict continuity of the trace operator (recalled in Section 10.3) that

$$\left|\mathscr{J}[u_0; B(x_0, r)] - \mathscr{J}[u; B(x_0, r)]\right| \leq \frac{\varepsilon}{2}|B(x_0, r)|$$

for r sufficiently small. Together with (11.13), the estimate (11.12) follows with $U(x_0, r) := B(x_0, r)$ after selecting $r > 0$ such that $(\mathscr{L}^d + |Du| + \Lambda)(\partial B(x_0, r)) = 0$. □

Lemma 11.14. *The estimate* (11.12) *holds for* $|D^s u|$*-almost every point* $x_0 \in \Omega$ *with* $U(x_0, r)$ *a (rotated) cube.*

Proof. We claim that (11.12) holds for all $x_0 \in \Omega$ such that

(a) $\dfrac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x_0) = a \otimes n$ for some $a \in \mathbb{S}^{m-1}$, $n \in \mathbb{S}^{d-1}$;
(b) $\alpha_r := r^{-d}|Du|(Q_n(x_0, 2r)) \to \infty$ as $r \downarrow 0$, where $Q_n(x_0, 2r)$ is a (henceforth fixed) open cube with midpoint $x_0 \in \Omega$, side-length $2r$, and two faces orthogonal to n.

By virtue of Alberti's Rank-One Theorem 10.7, the Besicovitch Differentiation Theorem A.23, and (10.6), these properties hold for $|D^s u|$-almost every $x_0 \in \Omega$. During the course of the proof we will require further properties of x_0, but every time the exceptional set will be $|D^s u|$-negligible.

Step 1. Let

$$u_r(y) := \frac{1}{\alpha_r} \cdot \frac{u(x_0 + ry) - [u]_{B(x_0,r)}}{r}, \qquad y \in Q_n(0, 2),\ \ r > 0,$$

whereby $[u_r] = \fint_{B(0,1)} u_r \,\mathrm{d}x = 0$. One calculates, using (10.8),

$$Du_r(B) = \frac{Du(x_0 + rB)}{r^d \alpha_r} = \frac{Du(x_0 + rB)}{|Du|(Q_n(x_0, 2r))}$$

for any Borel set $B \subset Q_n(0, 2)$. Hence, $|Du_r|(Q_n(0, 2)) = 1$. By Corollary 10.8 (a consequence of Alberti's Rank-One Theorem 10.7) and the Poincaré inequality in BV, $u_r \overset{*}{\rightharpoonup} u_0$ in $\mathrm{BV}(Q_n(0, 2); \mathbb{R}^m)$ for

$$u_0(y) = a\psi(y \cdot n) \tag{11.14}$$

with a bounded and increasing function $\psi\colon (-1, 1) \to \mathbb{R}$, also see Problem 10.4. Now apply Lemma 10.6 on blow-ups without loss of mass to the measure $|Du|$ to see that we may furthermore assume that

(c) Du_r converges strictly to Du_0 on $Q_n(0, 2)$ and $|Du_0|(Q_n(0, 2)) = 1$.

Then, $Du_0(Q_n(0, 2)) = a \otimes n$.

Define the functions $f_r : \mathbb{R}^{m\times d} \to \mathbb{R}$, $r > 0$, as

$$f_r(A) := \frac{f(\alpha_r A)}{\alpha_r}, \qquad A \in \mathbb{R}^{m\times d},$$

which satisfy $|f_r(A)| \leq M(1+|A|)$ for r small enough. The f_r are quasiconvex (see Problem 11.7) and

$$f_r(A) \to f^\infty(A) \qquad \text{as } r \downarrow 0 \qquad \text{whenever } \operatorname{rank}(A) \leq 1$$

by the rank-one convexity of f.

Next, define the auxiliary functionals $\mathscr{F}_r$, $\mathscr{J}_r$ for Lipschitz subdomains $D \subset \Omega$ and $v \in \mathrm{BV}(D;\mathbb{R}^m)$ via

$$\mathscr{F}_r[v;D] := \int_D f_r(\nabla v)\,\mathrm{d}x + \int_D f_r^\infty\biggl(\frac{\mathrm{d}D^s v}{\mathrm{d}|D^s v|}\biggr)\,\mathrm{d}|D^s v|,$$
$$\mathscr{J}_r[v;D] := \inf\Bigl\{\,\mathscr{F}_r[w;D] \;:\; w \in \mathrm{BV}(D;\mathbb{R}^m) \text{ with } w|_{\partial D} = v|_{\partial D}\,\Bigr\}.$$

We observe that Lemma 11.11 (for $\mathscr{J}_r$) implies

$$\bigl|\mathscr{J}_r[u_r;Q_n(0,2)] - \mathscr{J}_r[u_0;Q_n(0,2)]\bigr| \leq M\int_{\partial Q_n(0,2)} |u_r - u_0|\,\mathrm{d}\mathscr{H}^{d-1}$$
$$\to 0, \tag{11.15}$$

where the convergence follows by the strict convergence of Du_r to Du_0 and the strict continuity of the trace operator in BV.

Step 2. We will show next that

$$\mathscr{J}_r[u_0;Q_n(0,2)] \geq f_r\biggl(\frac{Du_0(Q_n(0,2))}{|Q_n(0,2)|}\biggr)|Q_n(0,2)|. \tag{11.16}$$

With ψ from (11.14) we have

$$Du_0 = (a\otimes n)|Du_0| = (a\otimes n)D\psi,$$

whereby, using (c) above, $1 = |Du_0|(Q_n(0,2)) = 2^{d-1}|D\psi|(-1,1)$. Thus,

$$|D\psi|(-1,1) = \psi(+1^-) - \psi(-1^+) = 2^{1-d},$$

where the values of ψ on the right-hand side are to be understood in the sense of left and right limits, respectively.

Define the staircase function

$$v(y) := a\psi\biggl(y\cdot n - 2\biggl\lfloor\frac{y\cdot n+1}{2}\biggr\rfloor\biggr) + 2^{1-d}a\biggl\lfloor\frac{y\cdot n+1}{2}\biggr\rfloor, \qquad y \in \mathbb{R}^d.$$

Furthermore, set $w_k(y) := v(ky)/k$ for $y \in Q_n(0,2)$ and $k \in \mathbb{N}$. We have that $w_k \to w$ uniformly in $\overline{Q_n(0,2)}$ as $k \to \infty$, where

$$w(y) := 2^{-d}(a \otimes n)y, \qquad y \in \mathbb{R}^d,$$

because ψ is bounded. Moreover, the trace of w_k on $\partial Q_n(0,2)$ converges to the trace of w and hence Lemma 11.11 implies

$$\left| \mathscr{I}_r[w_k; Q_n(0,2)] - \mathscr{I}_r[w; Q_n(0,2)] \right| \le M \int_{\partial Q_n(0,2)} |w_k - w| \, \mathrm{d}\mathscr{H}^{d-1} \to 0. \tag{11.17}$$

Now disjointly split

$$Q_n(0,2) = Z \cup \bigcup_{l=1}^{k^d} Q_l^{(k)}, \qquad |Z| = 0,$$

in the canonical way into a grid of k^d open cubes with two faces orthogonal to n and with side length $2/k$. From (11.10) we infer that

$$\begin{aligned}
\mathscr{I}_r[u_0; Q_n(0,2)] &= \mathscr{I}_r[v; Q_n(0,2)] \\
&= k^d \mathscr{I}_r[w_k; (-1/k, 1/k)^d] \\
&= \sum_{l=1}^{k^d} \mathscr{I}_r[w_k; Q_l^{(k)}] \\
&\ge \mathscr{I}_r[w_k; Q_n(0,2)]
\end{aligned}$$

for all $k \in \mathbb{N}$. The last inequality follows since we may combine admissible functions in the definition of $\mathscr{I}_r[w_k; Q_l^{(k)}]$ into an admissible function in the definition of $\mathscr{I}_r[w_k; Q_n(0,2)]$. Now let $k \to \infty$ and employ (11.17) together with Lemma 11.10 to see that

$$\mathscr{I}_r[u_0; Q_n(0,2)] \ge \mathscr{I}_r[w; Q_n(0,2)] \ge |Q_n(0,2)| f_r(\nabla w).$$

The claim (11.16) now follows from $\nabla w = 2^{-d}(a \otimes n) = Du_0(Q_n(0,2))/|Q_n(0,2)|$ since $Du_0(Q_n(0,2)) = a \otimes n$.

Step 3. From (11.16) and (c) we infer that

$$\begin{aligned}
\mathscr{I}_r[u_0; Q_n(0,2)] &\ge f_r\left(\frac{Du_0(Q_n(0,2))}{|Q_n(0,2)|} \right) |Q_n(0,2)| \\
&= f_r\left(\frac{a \otimes n}{|Q_n(0,2)|} \right) |Q_n(0,2)| \\
&\to f^\infty(a \otimes n) \qquad \text{as } r \downarrow 0.
\end{aligned}$$

Combining this with (11.15), we arrive at

$$\liminf_{r\downarrow 0} \mathscr{J}_r[u_r;\, Q_n(0,2)] \geq f^\infty(a\otimes n).$$

We will show below that

$$\frac{\mathscr{F}[u;\, Q_n(x_0,2r)]}{|Du|(Q_n(x_0,2r))} \to f^\infty(a\otimes n) \qquad \text{as } r\downarrow 0 \tag{11.18}$$

for $|D^s u|$-a.e. $x_0 \in \Omega$. Then, for r sufficiently small,

$$\begin{aligned}\frac{\mathscr{J}[u;\, Q_n(x_0,2r)]}{|Du|(Q_n(x_0,2r))} &= \mathscr{J}_r[u_r;\, Q_n(0,2)]\\ &\geq f^\infty(a\otimes n) - \frac{\varepsilon}{2}\\ &\geq \frac{\mathscr{F}[u;\, Q_n(x_0,2r)]}{|Du|(Q_n(x_0,2r))} - \varepsilon.\end{aligned}$$

Thus, (11.12) follows with $U(x_0,r) := Q_n(x_0,2r)$.

Step 4. To finish the proof of the lemma, it remains to show (11.18) at $|D^s u|$-almost every point $x_0 \in \Omega$. We now additionally assume that x_0 satisfies

(d) $\displaystyle \lim_{r\downarrow 0} \frac{|D^a u|(Q_n(x_0,2r))}{|Du|(Q_n(x_0,2r))} = \frac{\mathrm{d}|D^a u|}{\mathrm{d}|Du|}(x_0) = 0;$
(e) x_0 is a $|D^s u|$-Lebesgue point of $\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}$.

This is no restriction because

$$\frac{\mathrm{d}|D^a u|}{\mathrm{d}|Du|}(x_0) \leq \frac{\mathrm{d}|D^a u|}{\mathrm{d}|D^s u|}(x_0) = 0,$$

which determines the limit in (d) for $|D^s u|$-almost every $x_0 \in \Omega$ by the Besicovitch Differentiation Theorem A.23. So,

$$\frac{\mathrm{d}|D^s u|}{\mathrm{d}|Du|} = \frac{\mathrm{d}|D^a u|}{\mathrm{d}|Du|} + \frac{\mathrm{d}|D^s u|}{\mathrm{d}|Du|} = \frac{\mathrm{d}|Du|}{\mathrm{d}|Du|} = 1 \qquad |Du|\text{ -a.e.}$$

By the Lebesgue point property (e) of x_0,

$$\begin{aligned}&\frac{1}{|D^s u|(Q_n(x_0,2r)} \int_{Q_n(x_0,2r))} \left| f^\infty\!\left(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}\right) - f^\infty(a\otimes n)\right| \,\mathrm{d}|D^s u|\\ &\quad\leq \frac{L}{|D^s u|(Q_n(x_0,2r)} \int_{Q_n(x_0,2r))} \left|\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(y) - \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x_0)\right| \,\mathrm{d}|D^s u|(y)\\ &\quad\to 0 \qquad \text{as } r\downarrow 0,\end{aligned}$$

where $L > 0$ is the Lipschitz constant of f^∞ (which is the same as the Lipschitz constant of f). Furthermore, $\frac{\mathrm{d}|D^s u|}{\mathrm{d}|Du|}(x_0) = 1$ and hence we can replace the denominator in the leading fraction by $|Du|(Q(x_0, r))$. Finally,

$$\begin{aligned} &\frac{1}{|Du|(Q_n(x_0, 2r))} \int_{Q_n(x_0,2r)} |f(\nabla u)| \,\mathrm{d}x \\ &\quad \le \frac{M}{|Du|(Q_n(x_0, 2r))} \int_{Q_n(x_0,2r)} 1 + |\nabla u| \,\mathrm{d}x \\ &\quad \to 0 \qquad \text{as } r \downarrow 0 \end{aligned}$$

since $\alpha_r \to \infty$ and $\frac{\mathrm{d}|D^a u|}{\mathrm{d}|Du|}(x_0) = 0$. Together, these assertions yield (11.18). □

Combining the last two lemmas, we get:

Lemma 11.15. *Let $\Lambda \in \mathscr{M}^+(\Omega)$ and $\varepsilon > 0$. Then, there exist countably many disjoint convex open sets $\{U_k\}_{k\in\mathbb{N}}$ (balls or cubes) with $\Lambda(\partial U_k) = 0$ that cover Ω up to a $(\mathscr{L}^d + |Du|)$-negligible set such that*

$$\mathscr{F}[u; \Omega] \le \sum_{k=1}^{\infty} \mathscr{J}[u; U_k] + \varepsilon.$$

Proof. We have shown in the last two lemmas that $(\mathscr{L}^d + |D^s u|)$-almost every $x_0 \in \Omega$ satisfies (11.12) for sufficiently small radii $r > 0$ and some convex open set $U(x_0, r)$. More precisely, there is an $\mathscr{L}^d$-negligible Borel set $N_1 \subset \Omega$ and a $|D^s u|$-negligible Borel set N_2 such that (11.12) holds at all $x_0 \in (\Omega \setminus N_1) \cup (\Omega \setminus N_2) = \Omega \setminus (N_1 \cap N_2)$. Thus, at such x_0 the assumptions of the following covering theorem hold: □

Theorem 11.16 (Morse covering theorem). *Let $B \subset \mathbb{R}^d$ be a bounded Borel set, $\mu \in \mathscr{M}^+(\mathbb{R}^d)$, $\kappa \ge 1$, and let*

$$\mathscr{C} \subset \Big\{ x + K \;:\; x \in B,\, K \subset \mathbb{R}^d \text{ convex and compact } \Big\}$$

be a family of sets such that for all $x \in B \setminus N$, where $N \subset B$ is a Borel set with $\mu(N) = 0$, and for all $r \in (0, r_0(x_0))$ ($r_0(x_0) > 0$ given for every $x_0 \in B$), there exists an $x + K \in \mathscr{C}$ with

$$B(x, r) \subset x + K \subset B(x, \kappa r).$$

Then, there exists a disjoint countable family $\mathscr{C}' \subset \mathscr{C}$ with

$$\mu\Big(B \setminus \bigcup \mathscr{C}'\Big) = 0.$$

Via this theorem, which is proved in Theorem 1.147 of [122], we can cover Ω up to a $(\mathscr{L}^d + |Du|)$-negligible set by countably many mutually disjoint sets

U_k ($k \in \mathbb{N}$) satisfying (11.12). Note that the original result only holds for the *closures* $\overline{U_k}$, but as $(\mathscr{L}^d + |Du|)(\partial U_k) = 0$, this is equivalent.

Clearly, as an integral functional, $\mathscr{F}[u; \bullet]$ is countably additive and vanishes on $(\mathscr{L}^d + |Du|)$-negligible sets. Thus, (11.12) gives

$$\mathscr{F}[u; \Omega] \le \sum_{k=1}^{\infty} \mathscr{J}[u; U_k] + \varepsilon(\mathscr{L}^d + |Du|)(\Omega).$$

This immediately yields the claim after adjusting ε. □

We can now complete the proof of the main result of this section.

Proof of Theorem 11.7. Let $(u_j) \subset \mathrm{BV}(\Omega; \mathbb{R}^m)$ with $u_j \overset{*}{\rightharpoonup} u$ in $\mathrm{BV}(\Omega; \mathbb{R}^m)$. By Lemma 11.1 in conjunction with Theorem 11.2 we may assume that in fact $u_j \in (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(\Omega; \mathbb{R}^m)$ and that (after selecting a not explicitly labeled subsequence)

$$f(\nabla u_j)\, \mathscr{L}^d \llcorner \Omega \overset{*}{\rightharpoonup} \lambda, \quad |Du_j| \overset{*}{\rightharpoonup} \Lambda \qquad \text{in } \mathscr{M}^+(\overline{\Omega}).$$

By the linear growth assumption (12.2), we have $0 \le \lambda \le M(\mathscr{L}^d \llcorner \Omega + \Lambda)$. Furthermore, f is Lipschitz continuous by Lemma 5.6 with Lipschitz constant $L > 0$, say. We also assume that $|f(0)| \le L$.

Via Lemma 11.15 we construct a countable family of open disjoint sets U_k ($k \in \mathbb{N}$) with $\Lambda(\partial U_k) = 0$ for all $k \in \mathbb{N}$, $(\mathscr{L}^d + |Du|)(\Omega \setminus \bigcup_{k \in \mathbb{N}} U_k) = 0$, and such that

$$\mathscr{F}[u; \Omega] \le \sum_{k=1}^{\infty} \mathscr{J}[u; U_k] + \varepsilon \le \sum_{k=1}^{\infty} \liminf_{j \to \infty} \mathscr{F}[u_j; U_k] + \varepsilon,$$

where we used Lemma 11.12 for each of the U_k's. Since $\Lambda(\partial U_k) = 0$, it holds that

$$\lim_{j \to \infty} \mathscr{F}[u_j; U_k] = \lambda(U_k).$$

Thus,

$$\mathscr{F}[u; \Omega] \le \lambda\biggl(\bigcup_{k \in \mathbb{N}} U_k\biggr) + \varepsilon = \lambda(\Omega) + \varepsilon.$$

By standard results in measure theory (see Lemma A.19), we infer that (note $f \ge 0$)

$$\lambda(\Omega) \le \liminf_{j \to \infty} \mathscr{F}[u_j; \Omega],$$

whereby

$$\mathscr{F}[u; \Omega] \le \liminf_{j \to \infty} \mathscr{F}[u_j; \Omega] + \varepsilon.$$

Thus, as $\varepsilon > 0$ was arbitrary, we have proved that $\mathscr{F} = \mathscr{F}[\bullet; \Omega]$ is lower semicontinuous. □

Extending all functions $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ to $\tilde{u} \in \mathrm{BV}(\widetilde{\Omega};\mathbb{R}^m)$ on a larger domain $\widetilde{\Omega} \Supset \Omega$ by setting $\tilde{u}|_{\widetilde{\Omega}\setminus\overline{\Omega}} := w$ for some $w \in \mathrm{BV}(\widetilde{\Omega}\setminus\overline{\Omega};\mathbb{R}^m)$, and applying Theorem 11.7 in $\widetilde{\Omega}$, we also immediately get the following weak* lower semicontinuity result:

Corollary 11.17. *Assume that* $f\colon \mathbb{R}^{m\times d} \to [0,\infty)$ *is a quasiconvex integrand with linear growth and let* $g \in \mathrm{L}^1(\partial\Omega;\mathbb{R}^m)$*. Then, the functional*

$$\begin{aligned}\mathscr{F}_{\mathrm{ext}}[u] &:= \int_\Omega f(\nabla u(x))\,\mathrm{d}x + \int_\Omega f^{\#}\Bigl(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\Bigr)\,\mathrm{d}|D^s u|(x)\\ &\quad + \int_{\partial\Omega} f^{\#}\bigl((u(x)-g(x))\otimes n_\Omega(x)\bigr)\,\mathrm{d}\mathscr{H}^{d-1}(x), \qquad u \in \mathrm{BV}(\Omega;\mathbb{R}^m),\end{aligned}$$

is weakly lower semicontinuous.*

Remark 11.18. We stated Theorem 11.7 and Corollary 11.17 also in the case when the strong recession function f^∞ does not exist, but only proved them under this additional existence assumption. In the general case (with $f^{\#}$ in the definition of $\mathscr{F}$), the proof of Theorem 11.7 and Corollary 11.17 is the same except that the use of Theorem 11.2 has to be replaced by Remark 11.3 (i.e., the solution to Problem 11.5).

As an existence theorem for minimizers we then get by the Direct Method:

Theorem 11.19. *Assume that* $f\colon \mathbb{R}^{m\times d} \to [0,\infty)$ *is a quasiconvex integrand that satisfies the coercivity and linear growth estimate*

$$\mu|A| \le f(A) \le M(1+|A|) \qquad A \in \mathbb{R}^{m\times d},$$

for some $\mu, M > 0$ *and let* $g \in \mathrm{L}^1(\partial\Omega;\mathbb{R}^m)$*. Then, the functional*

$$\begin{aligned}\mathscr{F}_{\mathrm{ext}}[u] &:= \int_\Omega f(\nabla u(x))\,\mathrm{d}x + \int_\Omega f^{\#}\Bigl(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\Bigr)\,\mathrm{d}|D^s u|(x)\\ &\quad + \int_{\partial\Omega} f^{\#}\bigl((u(x)-g(x))\otimes n_\Omega(x)\bigr)\,\mathrm{d}\mathscr{H}^{d-1}(x), \qquad u \in \mathrm{BV}(\Omega;\mathbb{R}^m),\end{aligned}$$

has a minimizer over the space $\mathrm{BV}(\Omega;\mathbb{R}^m)$.

Note that since the trace operator on $\mathrm{BV}(\Omega;\mathbb{R}^m)$ is not weakly* continuous, we cannot expect that boundary values are preserved along a minimizing sequence.

Example 11.20. The isoperimetric problem from Section 1.2 leads to the following minimization problem:

$$\begin{cases}\text{Minimize} \quad \mathscr{F}[u] := \displaystyle\int_0^1 \sqrt{1+(u(s)')^2}\,\mathrm{d}s + |u(0)-\alpha| + |u(1)-\beta| \\[2ex] \text{over all} \qquad u \in \mathrm{BV}(0,1) \quad \text{with} \quad \displaystyle\int_0^1 u(s)\,\mathrm{d}s = A,\end{cases}$$

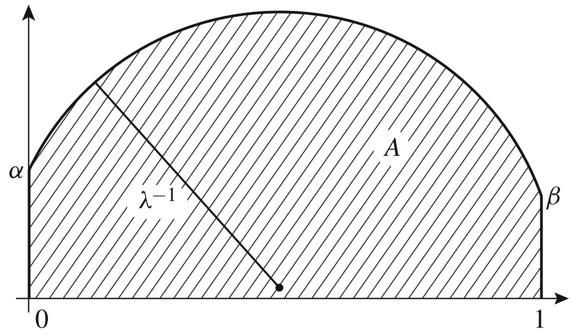


Fig. 11.1 The solution to the isoperimetric problem

where $\alpha, \beta, A > 0$ are given. Note that we have already translated the strict boundary conditions into the *penalty terms* $|u(0) - \alpha|$ and $|u(1) - \beta|$ (observe that the strong recession function of the integrand $f(a) := \sqrt{1+a^2}$ is $f^\infty(a) = |a|$). By the preceding theorem, there exists a solution to this problem (the side constraint can be incorporated like in Section 2.5).

Let us also identify this solution under the assumption that it is of class $\mathrm{W}^{2,1}$ inside the domain $(0, 1)$. Then, we can use Theorem 3.2.1 on Lagrange multipliers to see that u must solve the differential equation

$$\left(\frac{u'}{\sqrt{1+(u')^2}}\right)' = \lambda \qquad \text{in } (0, 1)$$

for some $\lambda \in \mathbb{R}$. The term on the left is the inverse curvature radius of the curve $\gamma(s) := (s, u(s))^T$, which is hence constant. From geometric reasoning we must therefore have that u is part of a circle that is open from below, see Figure 11.1.

In the special case when $\alpha = \beta$ we get that u is biggest when the radius of the circle is $1/2$ and u is a semicircle. Then, the area under the graph is

$$A_{\max} = \alpha + \frac{\pi}{8}.$$

Consequently, if the prescribed area A is larger than $\pi/8$, we must have a jump in u in the left and right endpoints. Of course, this is not directly expressible in the space $\mathrm{BV}(0, 1)$, whose elements are maps defined on the open interval $(0, 1)$. We can, however, as above extend u to all of $\mathbb{R}$ by α and work in the set $\{u \in \mathrm{BV}(\mathbb{R}) : Du \llcorner (\mathbb{R} \setminus [0, 1]) = 0\}$ instead.

11.3 Relaxation

In analogy to Chapter 7 we now consider the situation where the integrand f of the functional $\mathscr{F}$ is not quasiconvex. Then, $\mathscr{F}$ cannot be weakly* lower semicontinuous. For $f : \mathbb{R}^{m\times d} \to [0,\infty)$ a continuous integrand and any Lipschitz subdomain $D \subset \Omega$ we let the (restricted) functional $\mathscr{F}[\bullet; D] : \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m) \to \mathbb{R}$ be given as

$$\mathscr{F}[u; D] := \int_D f(\nabla u(x))\,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m).$$

Then, for $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ we define the **relaxation** $\mathscr{F}_*[\bullet; D]$ of $\mathscr{F}[\bullet; D]$ as

$$\mathscr{F}_*[u; D] := \inf\left\{ \liminf_{j\to\infty} \mathscr{F}[u_j; D] \;:\; (u_j) \subset \mathrm{W}^{1,1}(D;\mathbb{R}^m) \text{ with } u_j \overset{*}{\rightharpoonup} u \text{ in BV} \right\}$$

and we also set $\mathscr{F}_*[u] := \mathscr{F}_*[u;\Omega]$. Note that this definition does not agree with the abstract definition of the relaxation in Chapter 7. Indeed, there we identified the relaxation with the (weakly*) lower semicontinuous envelope, which here is

$$\widetilde{\mathscr{F}}_*[u] := \sup\left\{ \mathscr{H}[u] \;:\; \mathscr{H} \le \mathscr{F} \text{ and } \mathscr{H} \text{ is weakly* lower semicontinuous } \right\}.$$

However, we will show in Theorem 11.21 below that $\mathscr{F}_*$ as defined above is weakly* lower semicontinuous. Then, for all $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ it holds that

$$\begin{aligned}
\widetilde{\mathscr{F}}_*[u] &\le \inf\left\{ \liminf_{j\to\infty} \widetilde{\mathscr{F}}_*[u_j] \;:\; (u_j) \subset \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m) \text{ with } u_j \overset{*}{\rightharpoonup} u \text{ in BV} \right\} \\
&\le \inf\left\{ \liminf_{j\to\infty} \mathscr{F}[u_j] \;:\; (u_j) \subset \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m) \text{ with } u_j \overset{*}{\rightharpoonup} u \text{ in BV} \right\} \\
&= \mathscr{F}_*[u] \\
&\le \widetilde{\mathscr{F}}_*[u]
\end{aligned}$$

and a posteriori we conclude that in fact $\widetilde{\mathscr{F}}_* = \mathscr{F}_*$. Thus, we may work with the more convenient definition of the relaxation given in $\mathscr{F}_*$. In this context also see Problem 7.5.

The main theorem of this section is the following.

Theorem 11.21. *Assume that $f : \mathbb{R}^{m\times d} \to [0,\infty)$ is a continuous integrand with*

$$\mu|A| \le f(A) \le M(1+|A|), \qquad A \in \mathbb{R}^{m\times d},$$

for some $\mu, M > 0$. We denote the quasiconvex envelope of f by Qf. Then, the integral representation

$$\mathscr{F}_*[u] = \int_\Omega Qf(\nabla u(x))\,\mathrm{d}x + \int_\Omega (Qf)^\#\left(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\right)\mathrm{d}|D^s u|(x)$$

holds for all $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$. In particular, the functional $\mathscr{F}_*$ is weakly lower semicontinuous on $\mathrm{BV}(\Omega;\mathbb{R}^m)$.*

Remark 11.22. The result continues to hold if instead of the lower bound $\mu|A| \leq f(A)$ we only assume that $Q(f - \delta|\bullet|) > -\infty$ for some $\delta > 0$, see Problem 11.9.

Proof. We know from Lemma 7.1 that Qf is finite, quasiconvex, and has linear growth, say also with growth constant $M > 0$. We have $Qf(\nabla u) \leq M(1 + |\nabla u|)$ and $(Qf)^{\#}\big(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}\big) \leq M$. Denote by $\mathscr{Q}\mathscr{F}[u]$ the functional on the right-hand side above. The inequality

$$\mathscr{Q}\mathscr{F} \leq \mathscr{F}_* \tag{11.19}$$

follows immediately from the Ambrosio–Dal Maso–Fonseca–Müller Theorem 11.7, which entails that $\mathscr{Q}\mathscr{F}$ is weakly* lower semicontinuous. It remains to show the reverse inequality to (11.19).

Step 1. We claim that

$$\mathscr{F}_*[u] = \inf \left\{ \liminf_{j\to\infty} \int_\Omega f(\nabla u_j(x)) \,\mathrm{d}x \;:\; (u_j) \subset \mathrm{W}^{1,1}(\Omega; \mathbb{R}^m) \right.$$
$$\left. \text{with } u_j \to u \text{ in } \mathrm{L}^1 \right\}.$$

Indeed, if this were false, we could find $(u_j) \subset \mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$ with $u_j \to u$ in L^1 and

$$\mathscr{F}_*[u] > \lim_{j\to\infty} \int_\Omega f(\nabla u_j(x)) \,\mathrm{d}x \geq \mu\|\nabla u_j\|_{\mathrm{L}^1}.$$

So, the ∇u_j are uniformly L^1-bounded and $u_j \overset{*}{\rightharpoonup} u$ in BV, whereby we get the contradiction $\mathscr{F}_*[u] > \mathscr{F}_*[u]$.

Step 2. Using Lemma 11.1 choose a sequence $(u_j) \subset \mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$ with $u_j|_{\partial\Omega} = u|_{\partial\Omega}$ and $u_j \to u$ area-strictly. Furthermore, let $\Omega_0 \Subset \Omega$ be a Lipschitz subdomain with $(\mathscr{L}^d + |Du|)(\partial\Omega_0) = 0$. Since countably piecewise affine functions are dense in $\mathrm{W}^{1,1}(\Omega_0; \mathbb{R}^m)$ under given boundary values (see Theorem A.29), via Theorem 11.2 we may assume that u_j is countably piecewise affine in Ω_0, say $\nabla u_j = A_i^{(j)}$ almost everywhere in $\Omega_i^{(j)} \subset \Omega_0$ ($A_i^{(j)} \in \mathbb{R}^{m\times d}$, $i \in \mathbb{N}$). Here, the $\Omega_i^{(j)}$ are open and disjoint and for every $j \in \mathbb{N}$ it holds that $\Omega_0 = Z^{(j)} \cup \bigcup_i \Omega_i^{(j)}$ for some $\mathscr{L}^d$-negligible set $Z^{(j)}$.

Employing the formula (7.1) for the quasiconvex envelope, we can pick maps $\psi_i^{(j)} \in W_0^{1,\infty}(\Omega_i^{(j)}; \mathbb{R}^m)$ with

$$\int_{\Omega_i^{(j)}} |\psi_i^{(j)}(x)| \,\mathrm{d}x \leq \frac{|\Omega_i^{(j)}|}{j}$$

and

$$\int_{\Omega_i^{(j)}} f(A_i^{(j)} + \nabla\psi_i^{(j)}(x)) \,\mathrm{d}x < \left(Qf(A_i^{(j)}) + \frac{1}{j}\right)|\Omega_i^{(j)}|.$$

Let $v_j \in \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ be defined as

$$v_j := \begin{cases} u_j + \psi_i^{(j)} & \text{in } \Omega_i^{(j)} \ (i \in \mathbb{N}), \\ u_j & \text{in } \Omega \setminus \overline{\Omega_0}, \end{cases}$$

for which $v_j|_{\partial\Omega} = u|_{\partial\Omega}$ and $v_j \to u$ in L^1. Thus,

$$\mathscr{F}_*[u] \le \liminf_{j\to\infty} \int_\Omega f(\nabla v_j(x))\,\mathrm{d}x.$$

Since $v_j = u_j$ on $\Omega \setminus \overline{\Omega_0}$ we may estimate

$$\int_\Omega f(\nabla v_j(x))\,\mathrm{d}x \le \int_{\Omega_0} Qf(\nabla u_j(x))\,\mathrm{d}x + \frac{|\Omega|}{j} + M\int_{\Omega\setminus\overline{\Omega_0}} 1 + |\nabla u_j(x)|\,\mathrm{d}x.$$

Using Theorem 11.2 and Remark 11.3 we can pass to the limit as $j \to \infty$ to get

$$\begin{aligned} \mathscr{F}_*[u] \le \int_{\Omega_0} Qf(\nabla u(x))\,\mathrm{d}x &+ \int_{\Omega_0} (Qf)^\#\biggl(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\biggr)\,\mathrm{d}|D^s u|(x) \\ &+ M(\mathscr{L}^d + |Du|)(\Omega \setminus \Omega_0). \end{aligned}$$

Now let $\Omega_0 \uparrow \Omega$ (in the sense that $\sup_{x\in\Omega_0} \operatorname{dist}(x,\partial\Omega) \to 0$) to see that

$$\mathscr{F}_*[u] \le \int_\Omega Qf(\nabla u(x))\,\mathrm{d}x + \int_\Omega (Qf)^\#\biggl(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\biggr)\,\mathrm{d}|D^s u|(x) = \mathscr{Q}\mathscr{F}[u].$$

Together with (11.19) this concludes the proof of the theorem. □

Notes and Historical Remarks

The proof of Lemma 11.1 is from the appendix of [168] (the case for Lipschitz domains can be found in Lemma B.1 of [44]). Theorem 11.2 in the extended form of Remark 11.3 is from [169]. Results in this direction already appear in [44].

The blow-up method in the proof of the Ambrosio–Dal Maso–Fonseca–Müller Theorem 11.7 was first introduced in [123], also see [48] for a systematic approach to the idea of proving lower semicontinuity and relaxation theorems for integral functionals via the auxiliary functional $\mathscr{J}[u;U]$ from (11.9). In fact, the idea of Lemma 11.15 dates back to [86] where it was used in the context of Sobolev spaces. In the BV-context it seems to have been used for the first time in [47] and [48].

We note that the work by Fonseca & Müller [124] also considered u-dependent integrands. A more general approach to this problem using *liftings* can be found in [230].

The reader is pointed to Problem 11.3 and also to [284] for the construction of a non-convex quasiconvex function with linear growth and to [200] for an example of a non-convex quasiconvex function that is even positively 1-homogeneous. Theorem 8.1 of [164] shows, in a non-constructive fashion, that "many" quasiconvex functions with linear growth must exist.

The notation for recession functions is not consistent in the literature. In many works, the upper weak recession function $f^\#$ is written as f^∞ and simply called the "recession function". We refer to [22], in particular Section 2.5, for a more systematic approach to recession functions and their associated cones.

Problems

11.1. Find a simpler proof of Lemma 11.1 in the case when Ω is a bounded Lipschitz domain as follows: For $u \in \mathrm{BV}(\Omega; \mathbb{R}^m)$ define

$$u_\delta(x) := \int \eta(y)u(x - \delta\rho(x)y)\,\mathrm{d}y, \qquad \delta > 0,$$

where $\eta \in \mathrm{C}_c^\infty(B(0,1))$ is a standard mollifying kernel and ρ is a regularized distance to the boundary of Ω, i.e., $\rho \in \mathrm{C}^\infty(\Omega)$, $C^{-1}\mathrm{dist}(x, \partial\Omega) \le \rho(x) \le C\mathrm{dist}(x, \partial\Omega)$ ($C > 0$), and $|\nabla\rho(x)| \le C$ for all $x \in \Omega$. See, for instance, [246], p. 171, for the construction of such a ρ.

11.2. Show that for a continuous convex function $f\colon \mathbb{R}^N \to \mathbb{R}$ with linear growth it holds that $f^\# = f^\infty$.

11.3. Prove that there exists a non-convex quasiconvex integrand $f\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ with linear growth. Also prove that the strong recession function f^∞ exists. *Hint:* Extend Lemma 7.3 to the case $p = 1$.

11.4. Let $f\colon \Omega \times \mathbb{R}^{m\times d} \to \mathbb{R}$ be a Carathéodory integrand satisfying the lower bound $f(x, A) \ge -C(1 + |A|)$ for all $(x, A) \in \Omega \times \mathbb{R}^{m\times d}$ and some constant $C > 0$. Then, show that there exists a sequence $(f_k)_k$ of continuous integrands for which the strong recession functions f_k^∞ exists and such that

$$\sup_{k\in\mathbb{N}} f_k(x, A) = f(x, A) \qquad \text{and} \qquad \sup_{k\in\mathbb{N}} f_k^\infty(x, A) = f_\#(x, A).$$

11.5. Prove the following strengthened version of Theorem 11.2: Let $f\colon \Omega \times \mathbb{R}^{m\times d} \to [0, \infty)$ be continuous with linear growth such that $A \mapsto f(x, A)$ is rank-one convex (or rank-one concave) for almost every $x \in \Omega$ and moreover

$$f^{\#}(x, A) = f_{\#}(x, A) = (f(x, \bullet))^{\#}(A) = \limsup_{\substack{A' \to A \\ t \to \infty}} \frac{f(x, tA')}{t}$$

for all $(x, A) \in \Omega \times \mathbb{R}^{m \times d}$ such that $\operatorname{rank} A \leq 1$. Then, the functional

$$\mathscr{F}[u] := \int_\Omega f(x, \nabla u(x)) \, dx + \int_\Omega f^{\#}\left(x, \frac{dD^s u}{d|D^s u|}(x)\right) d|D^s u|(x), \qquad u \in \mathrm{BV}(\Omega; \mathbb{R}^m),$$

is continuous with respect to the area-strict convergence on $\mathrm{BV}(\Omega, \mathbb{R}^m)$. *Hint:* Use Problem 11.4 and also Alberti's Rank-One Theorem 10.7.

11.6. Show that if $f : \Omega \times \mathbb{R}^{m \times d} \to \mathbb{R}$ is rank-one convex and the function

$$(f(x, \bullet))^{\#}(A) = \limsup_{\substack{A' \to A \\ t \to \infty}} \frac{f(x, tA')}{t}$$

is continuous in $x \in \Omega$ for fixed A, then

$$f^{\#}(x, A) = f_{\#}(x, A) = (f(x, \bullet))^{\#}(A)$$

for all $(x, A) \in \Omega \times \mathbb{R}^{m \times d}$ such that $\operatorname{rank} A \leq 1$. *Hint:* Use Dini's Theorem, which asserts that if a monotone sequence of continuous functions converges pointwise on a compact space and if the limit function is also continuous, then the convergence is uniform.

11.7. Show that for a quasiconvex $h : \mathbb{R}^{m \times d} \to \mathbb{R}$ with linear growth the rescaled functions

$$h_r(A) := \frac{h(rA)}{r}, \qquad A \in \mathbb{R}^{m \times d},$$

are quasiconvex for all $r > 0$. Conclude that the upper weak recession function

$$h^{\#}(A) := \limsup_{t \to \infty} \frac{h(tA)}{t}$$

is quasiconvex. Show also that if $\operatorname{rank} A \leq 1$, then this upper limit is in fact a proper limit.

11.8. Show that Theorem 11.7 cannot be extended to integrands f taking negative values without restricting the class of admissible BV-sequences.

11.9. Show that Theorem 11.21 continues to hold if instead of the lower bound $\mu|A| \leq f(A)$ we only assume that $Q(f - \delta|\bullet|) > -\infty$ for some $\delta > 0$. *Hint:* Use the Kirchheim–Kristensen Theorem 10.13.

11.10. Prove the statements in Theorem 8.3 for the case $p = 1$. *Hint:* Use the weak-type estimates for Fourier multipliers from Theorem A.35.

Chapter 12
Generalized Young Measures

In this chapter we continue the study of the integral functional

$$\mathscr{F}[u] := \int_\Omega f(x, \nabla u(x))\,\mathrm{d}x + \int_\Omega f^\#\biggl(x, \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\biggr), \qquad u \in \mathrm{BV}(\Omega;\mathbb{R}^m),$$

for a Carathéodory integrand $f\colon \Omega \times \mathbb{R}^{m\times d} \to \mathbb{R}$ with linear growth. In contrast to the preceding chapter, however, here we proceed in a more abstract way: We first introduce the theory of *generalized Young measures*, which extends the standard theory of Young measures developed in Chapter 4. Besides quantifying oscillations (like classical Young measures), this theory crucially allows one to quantify *concentrations* as well, thus providing a rich toolbox for investigating linear-growth functionals. While the (generalized) Young measure approach requires a fair bit of abstract theory, the initial effort is rewarded with a robust general framework that has become a core tool in the calculus of variations with applications way beyond the lower semicontinuity theory of integral functionals.

To get a feel for this tool, let us outline a few basic ideas that play a prominent role in this chapter. Recall that if $p \in (1,\infty]$, then for a standard $\mathrm{W}^{1,p}$-gradient Young measure $\nu = (\nu_x)_{x\in\Omega} \in \mathbf{GY}^p(\Omega;\mathbb{R}^{m\times d})$ we can always find a norm-bounded sequence $(u_j) \subset \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ with $\nabla u_j \overset{\mathbf{Y}}{\to} \nu$ and such that additionally $\{\nabla u_j\}_j$ is L^p-equiintegrable (if $p < \infty$), see Lemma 4.13 for $p < \infty$ and Zhang's Lemma 7.18 for $p = \infty$. Thus, for $p > 1$, we do not need to worry about (L^p-)concentrations in generating sequences of gradients. However, as has already become apparent in Chapters 10 and 11, the linear-growth case $p = 1$ is very special. Here, concentration effects really have to be taken into account and classical Young measures cannot be used to this effect.

The main new, yet fairly simple, idea allowing one to pass from classical to generalized Young measures is to employ a *compactification*. For this, we transform maps $u\colon \Omega \to \mathbb{R}^N$ into maps $\tilde{u}$ taking values in $\mathbb{B}^N$, the (open) unit ball in $\mathbb{R}^N$.

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8_12

This can be achieved by choosing a homeomorphism $\varphi\colon \mathbb{R}^N \to \mathbb{B}^N$ and setting $\tilde{u} := u \circ \varphi^{-1}$. Then, the behavior of u "at infinity" can be understood by studying the behavior of $\tilde{u}$ near $\partial\mathbb{B}^N$. Likewise, when considering a (for simplicity) homogeneous Young measure $\nu \in \mathscr{M}^1(\mathbb{R}^N)$, we can instead study the push-forward $\tilde{\nu} := \varphi_\# \nu \in \mathscr{M}^1(\mathbb{B}^N)$. Generalized Young measures can then be understood as classical Young measures on the compactified space $\overline{\mathbb{B}^N}$. In this way we can treat situations where some mass in a sequence $(\nu_j) \subset \mathscr{M}^1(\mathbb{R}^N)$ escapes to infinity. The precise way this is formulated is slightly different, though, to make the theory more user-friendly.

After developing the functional analysis setup of generalized Young measures, including the introduction of a suitable set of "test integrands", we turn to the class of generalized Young measures that are generated by BV-sequences. In particular, localization ("blow-up") principles will play a prominent role, just like they did for classical Young measures. Finally, we will consider how these tools can be used to establish Jensen-type inequalities, which will yield lower semicontinuity results for integral functionals.

12.1 Functional Analysis Setup

We first define the space $\mathbf{E}(\Omega; \mathbb{R}^N)$, whose elements are the "test integrands" for generalized Young measures. In order to do so, we introduce for $f \in \mathrm{C}(\overline{\Omega} \times \mathbb{R}^N)$ and $g \in \mathrm{C}(\overline{\Omega} \times \mathbb{B}^N)$, where by $\mathbb{B}^N$ we denote the open unit ball in $\mathbb{R}^N$, the following linear transformations:

$$(Sf)(x, \hat{A}) := (1 - |\hat{A}|)\, f\left(x, \frac{\hat{A}}{1 - |\hat{A}|}\right), \qquad x \in \overline{\Omega},\ \hat{A} \in \mathbb{B}^N, \tag{12.1}$$

$$(S^{-1}g)(x, A) := (1 + |A|)\, g\left(x, \frac{A}{1 + |A|}\right), \qquad (x, A) \in \overline{\Omega} \times \mathbb{R}^N.$$

Clearly, $S^{-1} \circ S$ and $S \circ S^{-1}$ are the identities on $\mathrm{C}(\overline{\Omega} \times \mathbb{R}^N)$ and $\mathrm{C}(\overline{\Omega} \times \mathbb{B}^N)$, respectively. Then we set

$$\mathbf{E}(\Omega; \mathbb{R}^N) := \left\{ f \in \mathrm{C}(\overline{\Omega} \times \mathbb{R}^N) \ :\ Sf \in \mathrm{C}(\overline{\Omega \times \mathbb{B}^N}) \right\}.$$

Here, the condition "$Sf \in \mathrm{C}(\overline{\Omega \times \mathbb{B}^N})$" is to be understood in the way that $Sf \in \mathrm{C}(\overline{\Omega} \times \mathbb{B}^N)$ has a (necessarily unique) continuous extension to $\overline{\Omega \times \mathbb{B}^N}$, which is also denoted by Sf. As the norm on $\mathbf{E}(\Omega; \mathbb{R}^N)$ we use the natural choice

$$\|f\|_{\mathbf{E}(\Omega;\mathbb{R}^N)} := \|Sf\|_{\mathrm{C}(\overline{\Omega \times \mathbb{B}^N})} = \sup_{(x,\hat{A}) \in \overline{\Omega \times \mathbb{B}^N}} |Sf(x, \hat{A})|,$$

under which $\mathbf{E}(\Omega; \mathbb{R}^N)$ becomes a Banach space and the operator $S\colon \mathbf{E}(\Omega; \mathbb{R}^N) \to \mathrm{C}(\overline{\Omega \times \mathbb{B}^N})$ is an isometric isomorphism. In particular, $\mathbf{E}(\Omega; \mathbb{R}^N)$ is separable.

Since $|f(x, A)| = (1+|A|)|Sf(x, (1+|A|)^{-1}A)|$, all $f \in \mathbf{E}(\Omega; \mathbb{R}^N)$ have linear growth, i.e., there exists a constant $M > 0$ with

$$|f(x, A)| \le M(1 + |A|), \qquad (x, A) \in \overline{\Omega} \times \mathbb{R}^N. \tag{12.2}$$

Moreover, our definition of $\mathbf{E}(\Omega; \mathbb{R}^N)$ is designed so that the (strong) recession function $f^\infty \colon \overline{\Omega} \times \mathbb{R}^N \to \mathbb{R}$ exists, which was defined in (11.8) as

$$f^\infty(x, A) := \lim_{\substack{x' \to x \\ A' \to A \\ t \to \infty}} \frac{f(x', tA')}{t}, \qquad (x, A) \in \overline{\Omega} \times \mathbb{R}^N.$$

We also recall that f^∞ is positively 1-homogeneous. Note that f^∞ agrees with Sf on $\overline{\Omega} \times \mathbb{S}^{N-1}$, as can be seen by substituting $t = s/(1 - s)$, $s \in (0, 1)$, and letting $s \to 1$.

A **generalized Young measure** on the bounded open set $\Omega \subset \mathbb{R}^d$ with values in $\mathbb{R}^N$ is a triple $\nu = (\nu_x, \lambda_\nu, \nu_x^\infty)$ consisting of

(i) a parametrized family of probability measures $(\nu_x)_{x \in \Omega} \subset \mathscr{M}^1(\mathbb{R}^N)$, called the **oscillation measure**;
(ii) a positive finite measure $\lambda_\nu \in \mathscr{M}^+(\overline{\Omega})$, called the **concentration measure**;
(iii) a parametrized family of probability measures $(\nu_x^\infty)_{x \in \overline{\Omega}} \subset \mathscr{M}^1(\mathbb{S}^{N-1})$, called the **concentration-direction measure**;

and satisfying the conditions

(iv) the map $x \mapsto \nu_x$ is weakly* measurable with respect to $\mathscr{L}^d$, i.e., the function $x \mapsto \langle f(x, \bullet), \nu_x \rangle$ is $\mathscr{L}^d$-measurable for all bounded Borel functions $f \colon \Omega \times \mathbb{R}^N \to \mathbb{R}$;
(v) the map $x \mapsto \nu_x^\infty$ is weakly* measurable with respect to λ_ν, i.e., the function $x \mapsto \langle f^\infty(x, \bullet), \nu_x^\infty \rangle$ is λ_ν-measurable for all bounded Borel functions $f \colon \Omega \times \mathbb{S}^{N-1} \to \mathbb{R}$;
(vi) $x \mapsto \langle |\bullet|, \nu_x \rangle \in \mathrm{L}^1(\Omega)$.

We identify generalized Young measures μ, ν if $\mu_x = \nu_x$ for $\mathscr{L}^d$-almost every $x \in \Omega$, $\lambda_\mu = \lambda_\nu$, and $\mu_x^\infty = \nu_x^\infty$ for λ_μ-almost every $x \in \overline{\Omega}$. All these (equivalence classes of) generalized Young measures are collected in the set

$$\mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^N).$$

If for the target dimension we have $N = 1$, then we simply write $\mathbf{Y}^{\mathscr{M}}(\Omega)$ instead of $\mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R})$. In all of the following we usually refer to generalized Young measures simply as "Young measures".

The **duality pairing** between $f \in \mathbf{E}(\Omega;\mathbb{R}^N)$ and $\nu \in \mathbf{Y}^{\mathscr{M}}(\Omega;\mathbb{R}^N)$ is defined as

$$\begin{aligned}\langle\!\langle f, \nu \rangle\!\rangle &:= \int_\Omega \langle f(x,\bullet), \nu_x\rangle \,\mathrm{d}x + \int_{\overline{\Omega}} \langle f^\infty(x,\bullet), \nu_x^\infty\rangle \,\mathrm{d}\lambda_\nu(x) \\ &= \int_\Omega \int_{\mathbb{R}^N} f(x,A)\,\mathrm{d}\nu_x(A)\,\mathrm{d}x + \int_{\overline{\Omega}} \int_{\mathbb{S}^{N-1}} f^\infty(x,A)\,\mathrm{d}\nu_x^\infty(A)\,\mathrm{d}\lambda_\nu(x).\end{aligned}$$

In this way, the space $\mathbf{Y}^{\mathscr{M}}(\Omega;\mathbb{R}^N)$ can be considered a part of the dual space of $\mathbf{E}(\Omega;\mathbb{R}^N)$. A sequence of Young measures $(\nu_j) \subset \mathbf{Y}^{\mathscr{M}}(\Omega;\mathbb{R}^N)$ **converges weakly*** to $\nu \in \mathbf{Y}^{\mathscr{M}}(\Omega;\mathbb{R}^N)$, written as "$\nu_j \overset{*}{\rightharpoonup} \nu$", if for every $f \in \mathbf{E}(\Omega;\mathbb{R}^N)$ it holds that $\langle\!\langle f, \nu_j \rangle\!\rangle \to \langle\!\langle f, \nu \rangle\!\rangle$.

The **barycenter** of a Young measure $\nu \in \mathbf{Y}^{\mathscr{M}}(\Omega;\mathbb{R}^N)$ is the measure $[\nu] \in \mathscr{M}(\overline{\Omega};\mathbb{R}^N)$ given by

$$[\nu] := [\nu_x]\,\mathscr{L}^d_x \llcorner \Omega + [\nu_x^\infty]\,\lambda_\nu(\mathrm{d}x), \tag{12.3}$$

where $[\mu] := \int A \,\mathrm{d}\mu(A)$ and we wrote $\mathscr{L}^d_x$, $\lambda_\nu(\mathrm{d}x)$ to emphasize that the measures $\mathscr{L}^d$, λ_ν act with respect to the x-variable. It is not hard to see that if $\nu_j \overset{*}{\rightharpoonup} \nu$ in $\mathbf{Y}^{\mathscr{M}}(\Omega;\mathbb{R}^N)$, then $[\nu_j] \overset{*}{\rightharpoonup} [\nu]$ in $\mathscr{M}(\overline{\Omega};\mathbb{R}^N)$.

In the following we study further the space $\mathbf{Y}^{\mathscr{M}}(\Omega;\mathbb{R}^N)$ and establish a fundamental compactness result for weak* convergence. Notice that since $\mathbf{Y}^{\mathscr{M}}(\Omega;\mathbb{R}^N) \subset \mathbf{E}(\Omega;\mathbb{R}^N)^*$, a sequence of Young measures that is suitably bounded (to be detailed below) has a weakly*-converging subsequence in $\mathbf{E}(\Omega;\mathbb{R}^N)^*$. However, it is not a priori clear that the limit is also a Young measure. Here and in the following we always identify $\mathrm{C}(\overline{\Omega \times \mathbb{B}^N})^*$ with $\mathscr{M}(\overline{\Omega \times \mathbb{B}^N})$ via the Riesz Representation Theorem A.21.

We first observe that the linear transformation $S \colon \mathbf{E}(\Omega;\mathbb{R}^N) \to \mathrm{C}(\overline{\Omega \times \mathbb{B}^N})$ defined in (12.1) is an isomorphism, hence the operator

$$S^{-*} := (S^{-1})^* \colon \mathbf{E}(\Omega;\mathbb{R}^N)^* \to \mathscr{M}(\overline{\Omega \times \mathbb{B}^N}),$$

that is, the dual operator of the inverse of S, acts on a Young measure $\nu \in \mathbf{Y}^{\mathscr{M}}(\Omega;\mathbb{R}^N)$ as

$$\begin{aligned}\langle \Phi, S^{-*}\nu\rangle &= \langle\!\langle S^{-1}\Phi, \nu\rangle\!\rangle \\ &= \int_\Omega \langle S^{-1}\Phi(x,\bullet), \nu_x\rangle \,\mathrm{d}x + \int_{\overline{\Omega}} \langle \Phi(x,\bullet), \nu_x^\infty\rangle \,\mathrm{d}\lambda_\nu(x)\end{aligned} \tag{12.4}$$

for any $\Phi \in \mathrm{C}(\overline{\Omega \times \mathbb{B}^N})$ (notice that $(S^{-1}\Phi)^\infty|_{\overline{\Omega}\times\mathbb{S}^{N-1}} = \Phi|_{\overline{\Omega}\times\mathbb{S}^{N-1}}$). In particular,

$$\langle Sf, S^{-*}\nu\rangle = \langle\!\langle f, \nu\rangle\!\rangle \qquad \text{for all } f \in \mathbf{E}(\Omega;\mathbb{R}^N).$$

See Figure 12.1 for a diagram of the duality relationships.

$$
\begin{array}{ccc}
\mathbf{E}(\Omega;\mathbb{R}^N) \ni f & \xrightarrow{\;S\;} & Sf \in \mathrm{C}(\overline{\Omega\times\mathbb{B}^N}) \\
\downarrow & & \downarrow \\
\langle\!\langle \bullet,\bullet \rangle\!\rangle & & \langle \bullet,\bullet \rangle \\
\uparrow & & \uparrow \\
\mathbf{Y}^{\mathscr{M}}(\Omega;\mathbb{R}^N) \ni \nu & \xrightarrow{\;S^{-*}\;} & S^{-*}\nu \in \mathscr{M}^+(\overline{\Omega\times\mathbb{B}^N}) \text{ with (12.5)}
\end{array}
$$

Fig. 12.1 Duality relationships.

Lemma 12.1. *The set* $S^{-*}(\mathbf{Y}^{\mathscr{M}}(\Omega;\mathbb{R}^N)) \subset \mathscr{M}(\overline{\Omega\times\mathbb{B}^N})$ *consists of all the positive measures* $\mu \in \mathscr{M}^+(\overline{\Omega\times\mathbb{B}^N})$ *that satisfy*

$$\int_{\overline{\Omega\times\mathbb{B}^N}} \varphi(x)(1-|A|)\,\mathrm{d}\mu(x,A) = \int_\Omega \varphi(x)\,\mathrm{d}x \qquad \textit{for all } \varphi \in C(\overline{\Omega}). \tag{12.5}$$

Proof. Step 1. For $\nu \in \mathbf{Y}^{\mathscr{M}}(\Omega;\mathbb{R}^N)$ and $\Phi = \mathbb{1}$, (12.4) gives

$$(S^{-*}\nu)(\overline{\Omega\times\mathbb{B}^N}) = \big\langle \mathbb{1}, S^{-*}\nu \big\rangle = \int_\Omega \big\langle 1+|\bullet|, \nu_x \big\rangle\,\mathrm{d}x + \lambda_\nu(\overline{\Omega}) < \infty$$

by the assumptions on ν. Thus, $S^{-*}\nu$ is a finite measure on $\overline{\Omega\times\mathbb{B}^N}$. Moreover, for $\varphi \in \mathrm{C}(\overline{\Omega})$ set $\Phi(x,A) := \varphi(x)(1-|A|)$, which gives

$$\int_{\overline{\Omega\times\mathbb{B}^N}} \varphi(x)(1-|A|)\,\mathrm{d}(S^{-*}\nu)(x,A) = \big\langle \Phi, S^{-*}\nu \big\rangle = \big\langle\!\big\langle \varphi\otimes\mathbb{1}, \nu \big\rangle\!\big\rangle = \int_\Omega \varphi(x)\,\mathrm{d}x.$$

This implies (12.5). The positivity of $S^{-*}\nu$ follows from the fact that for $\Phi \geq 0$ it holds that

$$\big\langle \Phi, S^{-*}\nu \big\rangle = \big\langle\!\big\langle S^{-1}\Phi, \nu \big\rangle\!\big\rangle \geq 0.$$

Step 2. To prove the converse, let $\mu \in \mathscr{M}^+(\overline{\Omega\times\mathbb{B}^N})$ be such that (12.5) holds. We need to construct a Young measure $\nu \in \mathbf{Y}^{\mathscr{M}}(\Omega;\mathbb{R}^N)$ with $\mu = S^{-*}\nu$, that is,

$$\big\langle Sf, \mu \big\rangle = \big\langle Sf, S^{-*}\nu \big\rangle = \big\langle\!\big\langle f, \nu \big\rangle\!\big\rangle \qquad \text{for all } f \in \mathbf{E}(\Omega;\mathbb{R}^N).$$

By the Disintegration Theorem 4.4 we infer the existence of a measure $\kappa \in \mathscr{M}^+(\overline{\Omega})$ and a weakly* κ-measurable family $(\eta_x)_{x\in\overline{\Omega}} \subset \mathscr{M}^1(\overline{\mathbb{B}^N})$ such that

$$\mu = \kappa(\mathrm{d}x) \otimes \eta_x.$$

For ease of notation in the following we will suppress all mention of the integration variable x for κ and $\mathscr{L}^d \llcorner \Omega$. Let

$$\kappa = g\mathscr{L}^d \llcorner \Omega + \kappa^s, \qquad g \in \mathrm{L}^1(\Omega),\ \kappa^s \text{ singular to } \mathscr{L}^d,$$

be the Lebesgue–Radon–Nikodým decomposition of κ. From (12.5) we get that

$$\langle 1-|\bullet|, \eta_x\rangle\big[g\mathscr{L}^d \llcorner \Omega + \kappa^s\big] = \mathscr{L}^d \llcorner \Omega.$$

Thus, $\langle 1-|\bullet|, \eta_x\rangle\,\kappa^s = 0$, whereby η_x is concentrated in $\mathbb{S}^{N-1}$ for κ^s-almost every $x \in \overline{\Omega}$. Consequently, for $f \in \mathbf{E}(\Omega;\mathbb{R}^N)$ (recall $Sf = f^\infty$ on $\overline{\Omega}\times\mathbb{S}^{N-1}$),

$$\begin{aligned}\big\langle Sf(x,\bullet), \eta_x\big\rangle \kappa &= \left(g(x)\int_{\mathbb{B}^N} Sf(x,\bullet)\,\mathrm{d}\eta_x\right)\mathscr{L}^d \llcorner \Omega \\ &\quad + \left(\eta_x(\mathbb{S}^{N-1})g(x) \,⨍_{\partial\mathbb{B}^N} f^\infty(x,\bullet)\,\mathrm{d}\eta_x\right)\mathscr{L}^d \llcorner \Omega \\ &\quad + \left(\eta_x(\mathbb{S}^{N-1}) \,⨍_{\partial\mathbb{B}^N} f^\infty(x,\bullet)\,\mathrm{d}\eta_x\right)\kappa^s. \end{aligned} \tag{12.6}$$

Define the measures $\nu_x \in \mathscr{M}^+(\mathbb{R}^N)$, $x \in \Omega$, via

$$\big\langle h, \nu_x\big\rangle := g(x)\int_{\mathbb{B}^N} Sh\,\mathrm{d}\eta_x, \qquad h \in \mathrm{C}_0(\mathbb{R}^N);$$

the measures $\nu_x^\infty \in \mathscr{M}^+(\mathbb{S}^{N-1})$, $x \in \overline{\Omega}$, via

$$\big\langle h^\infty, \nu_x^\infty\big\rangle := ⨍_{\mathbb{S}^{N-1}} h^\infty\,\mathrm{d}\eta_x, \qquad h^\infty \in \mathrm{C}(\mathbb{S}^{N-1});$$

and

$$\lambda_\nu := \eta_x(\mathbb{S}^{N-1})\,\kappa = \eta_x(\mathbb{S}^{N-1})\big[g\mathscr{L}^d \llcorner \Omega + \kappa^s\big] \in \mathscr{M}^+(\overline{\Omega}).$$

Then, (12.6) becomes

$$\big\langle Sf(x,\bullet), \eta_x\big\rangle \kappa = \big\langle f(x,\bullet), \nu_x\big\rangle \mathscr{L}^d \llcorner \Omega + \big\langle f^\infty(x,\bullet), \nu_x^\infty\big\rangle \lambda_\nu. \tag{12.7}$$

Step 3. Next, we will show that ν_x and ν_x^∞ are indeed probability measures. For ν_x^∞, which is defined through an averaged integral, this is obvious. For ν_x, at $\mathscr{L}^d$-almost every $x \in \Omega$ observe that by (12.7) with $f(x,A) := \varphi(x)$ and (12.5) we infer that

$$\begin{aligned}\int_\Omega \varphi(x)\big\langle \mathbb{1}, \nu_x\big\rangle\,\mathrm{d}x &= \int_\Omega \varphi(x)\big\langle 1-|\bullet|, \eta_x\big\rangle\,\mathrm{d}\kappa(x) \\ &= \int_{\overline{\Omega\times\mathbb{B}^N}} \varphi(x)(1-|A|)\,\mathrm{d}\mu(x,A) \\ &= \int_\Omega \varphi(x)\,\mathrm{d}x\end{aligned}$$

for every $\varphi \in \mathrm{C}(\overline{\Omega})$. Thus, $\langle \mathbb{1}, \nu_x \rangle = 1$ for $\mathscr{L}^d$-almost every $x \in \Omega$.

Finally, using $f(x, A) := 1 + |A|$ in (12.7),

$$\int_\Omega \big\langle 1 + |\bullet|, \nu_x \big\rangle \,\mathrm{d}x + \lambda_\nu(\overline{\Omega}) = \big\langle \mathbb{1}, \mu \big\rangle = \mu(\overline{\Omega \times \mathbb{B}^N}) < \infty$$

since $Sf(x, A) = \mathbb{1}$. Therefore, $x \mapsto \langle |\bullet|, \nu_x \rangle \in \mathrm{L}^1(\Omega)$ and $\lambda_\nu(\overline{\Omega}) < \infty$. □

Corollary 12.2. *The set* $\mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^N)$ *is weakly* closed (as a subset of* $\mathbf{E}(\Omega; \mathbb{R}^N)^*$*).*

Proof. It suffices to observe that condition (12.5) is weak*-continuous (in the topological sense). Indeed, since $\varphi(x)(1 - |A|)$ for $\varphi \in \mathrm{C}(\overline{\Omega})$ is an admissible test function for the weak* topology on $\mathscr{M}(\overline{\Omega \times \mathbb{B}^N}) \cong \mathrm{C}(\overline{\Omega \times \mathbb{B}^N})^*$, the map

$$\mu \in \mathscr{M}(\overline{\Omega \times \mathbb{B}^N}) \;\mapsto\; \int_{\overline{\Omega \times \mathbb{B}^N}} \varphi(x)(1 - |A|) \,\mathrm{d}\mu(x, A)$$

is continuous in the (locally convex) weak* topology and thus $S^{-*}(\mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^N)) \subset \mathscr{M}(\overline{\Omega \times \mathbb{B}^N})$ is weakly* closed in $\mathrm{C}(\overline{\Omega \times \mathbb{B}^N})^*$. Via the isomorphism S^* the weak* closedness is transported to the set $\mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^N)$. □

The following is the basic *compactness principle* for generalized Young measures.

Corollary 12.3. *Let* $(\nu_j) \subset \mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^N)$ *be a sequence of Young measures such that*

$$\sup_{j \in \mathbb{N}} \big\langle\!\big\langle \mathbb{1} \otimes |\bullet|, \nu_j \big\rangle\!\big\rangle < \infty$$

or, equivalently,

(i) the functions $x \mapsto \langle |\bullet|, (\nu_j)_x \rangle$ *are uniformly bounded in* $\mathrm{L}^1(\Omega)$ *and*
(ii) the sequence $(\lambda_{\nu_j}(\overline{\Omega}))_j$ *is uniformly bounded.*

Then, there exists a subsequence (not explicitly labeled) such that $\nu_j \overset{*}{\rightharpoonup} \nu$ *for a Young measure* $\nu \in \mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^N)$.

Proof. Let $\Phi \in \mathrm{C}(\overline{\Omega \times \mathbb{B}^N})$ with $\|\Phi\|_\infty \le 1$. Then,

$$\begin{aligned}
\big|\big\langle \Phi, S^{-*}\nu_j \big\rangle\big| &= \big|\big\langle\!\big\langle S^{-1}\Phi, \nu_j \big\rangle\!\big\rangle\big| \\
&\le \int_\Omega \int_{\mathbb{R}^N} (1 + |A|) \left| \Phi\left(x, \frac{A}{1 + |A|}\right) \right| \,\mathrm{d}(\nu_j)_x(A) \,\mathrm{d}x \\
&\qquad + \int_{\overline{\Omega}} \int_{\mathbb{S}^{N-1}} |\Phi(x, A)| \,\mathrm{d}(\nu_j)_x^\infty(A) \,\mathrm{d}\lambda_\nu(x) \\
&\le \sup_{j \in \mathbb{N}} \left(\int_\Omega \big\langle 1 + |\bullet|, (\nu_j)_x \big\rangle \,\mathrm{d}x + \lambda_{\nu_j}(\overline{\Omega}) \right) \\
&< \infty.
\end{aligned}$$

Thus, the sequence $(S^{-*}\nu_j)$ is uniformly norm-bounded in $\mathcal{M}^+(\overline{\Omega \times \mathbb{B}^N})$ and hence there exists a weakly* converging subsequence, say $S^{-*}\nu_j \overset{*}{\rightharpoondown} \mu$ in $\mathcal{M}^+(\overline{\Omega \times \mathbb{B}^N})$. By Corollary 12.2, the limit μ is again the transformation under S^{-*} of a Young measure. So, $\mu = S^{-*}\nu$ for a Young measure $\nu \in \mathbf{Y}^{\mathcal{M}}(\Omega; \mathbb{R}^N)$. □

12.2 Generation and Examples

Having built the foundation of the theory of generalized Young measures, we now proceed to the question of *generation* of the said Young measures by sequences of L^1-bounded sequences of maps or, more generally, sequences of (vector) Radon measures that have uniformly bounded mass.

Let $\gamma \in \mathcal{M}(\overline{\Omega}; \mathbb{R}^N)$ be a (finite) Radon measure with Lebesgue–Radon–Nikodým decomposition

$$\gamma = g\mathcal{L}^d \llcorner \Omega + \gamma^s, \qquad \text{where} \qquad g \in \mathrm{L}^1(\Omega; \mathbb{R}^N),\ \gamma^s \text{ singular to } \mathcal{L}^d.$$

To γ we associate an **elementary Young measure** $\delta[\gamma] \in \mathbf{Y}^{\mathcal{M}}(\Omega; \mathbb{R}^N)$ via

$$\delta[\gamma]_x := \delta_{g(x)} \quad \mathcal{L}^d\text{-a.e.}, \qquad \lambda_{\delta[\gamma]} := |\gamma^s|, \qquad \delta[\gamma]_x^\infty := \delta_{P(x)} \quad |\gamma^s|\text{-a.e.},$$

where

$$P := \frac{\mathrm{d}\gamma^s}{\mathrm{d}|\gamma^s|} \in \mathrm{L}^1(\overline{\Omega}, |\gamma^s|; \mathbb{S}^{N-1}).$$

We will see momentarily why this definition is chosen as such. We say that a sequence $(\gamma_j) \subset \mathcal{M}(\overline{\Omega}; \mathbb{R}^N)$ with $\sup_j |\gamma_j|(\overline{\Omega}) < \infty$ **generates** the Young measure $\nu = (\nu_x, \lambda_\nu, \nu_x^\infty) \in \mathbf{Y}^{\mathcal{M}}(\Omega; \mathbb{R}^N)$, in symbols "$\gamma_j \overset{\mathbf{Y}}{\to} \nu$", if $\delta[\gamma_j] \overset{*}{\rightharpoondown} \nu$ in $\mathbf{Y}^{\mathcal{M}}(\Omega; \mathbb{R}^N)$, that is,

$$\langle\!\langle f, \delta[\gamma_j] \rangle\!\rangle \to \langle\!\langle f, \nu \rangle\!\rangle \qquad \text{for all } f \in \mathbf{E}(\Omega; \mathbb{R}^N).$$

Unwinding the definitions further, $\gamma_j \overset{\mathbf{Y}}{\to} \nu$ means that for all $f \in \mathbf{E}(\Omega; \mathbb{R}^N)$,

$$\begin{aligned} &f\biggl(x, \frac{\mathrm{d}\gamma_j}{\mathrm{d}\mathcal{L}^d}(x)\biggr) \mathcal{L}_x^d \llcorner \Omega \ + \ f^\infty\biggl(x, \frac{\mathrm{d}\gamma_j^s}{\mathrm{d}|\gamma_j^s|}(x)\biggr) \, |\gamma_j^s|(\mathrm{d}x) \\ &\quad \overset{*}{\rightharpoondown} \ \bigl\langle f(x, \bullet), \nu_x \bigr\rangle \mathcal{L}_x^d \llcorner \Omega \ + \ \bigl\langle f^\infty(x, \bullet), \nu_x^\infty \bigr\rangle \lambda_\nu(\mathrm{d}x) \quad \text{in } \mathcal{M}(\overline{\Omega}). \end{aligned}$$

If $v_j \mathcal{L}^d \llcorner \Omega \overset{\mathbf{Y}}{\to} \nu$ for a sequence of uniformly L^1-bounded maps $(v_j) \subset \mathrm{L}^1(\Omega; \mathbb{R}^N)$, we simply write $v_j \overset{\mathbf{Y}}{\to} \nu$.

The following result justifies the definition of elementary Young measures.

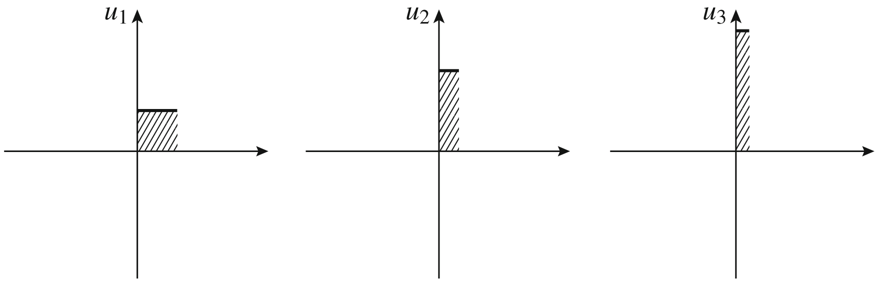


Fig. 12.2 A concentrating sequence

Proposition 12.4. *Let* $(\gamma_j) \subset \mathscr{M}(\overline{\Omega}; \mathbb{R}^N)$. *Then,* $\gamma_j \to \gamma$ *area-strictly if and only if* $\gamma_j \xrightarrow{\mathbf{Y}} \delta[\gamma]$.

Proof. If $\gamma_j \to \gamma$ area-strictly in $\mathscr{M}(\overline{\Omega}; \mathbb{R}^N)$, then $\gamma_j \xrightarrow{\mathbf{Y}} \delta[\gamma]$ follows by the same strategy as in the proof of Theorem 11.2, which is based on Reshetnyak's Continuity Theorem 10.3. For the converse test the weak* convergence $\delta[\gamma_j] \overset{*}{\rightharpoonup} \delta[\gamma]$ with the integrand $f(x, A) := \sqrt{1 + |A|^2}$, which lies in $\mathbf{E}(\Omega; \mathbb{R}^N)$. □

The following is customarily called the *Fundamental Theorem* of the generalized Young measure theory.

Theorem 12.5. *Let* $(\gamma_j) \subset \mathscr{M}(\overline{\Omega}; \mathbb{R}^N)$ *be a sequence of Radon measures such that*

$$\sup_{j \in \mathbb{N}} |\gamma_j|(\overline{\Omega}) < \infty.$$

Then, there exists a subsequence (not explicitly labeled) with $\gamma_j \xrightarrow{\mathbf{Y}} \nu$ *for some* $\nu \in \mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^N)$.

Proof. Set $\nu_j := \delta[\gamma_j]$, the elementary Young measure associated with γ_j and apply the compactness result from Corollary 12.3. The assumptions of that corollary are satisfied since the quantities $\langle\!\langle \mathbb{1} \otimes |\cdot|, \delta[\gamma_j] \rangle\!\rangle = |\gamma_j|(\overline{\Omega})$ are uniformly bounded. □

We have already seen many examples of oscillation effects, most directly in the examples of classical Young measures from Sections 4.2, 4.4. All these examples carry over to the present theory of generalized Young measures (with zero concentration measure). The following examples will illustrate model cases of concentration effects.

Example 12.6. Take $\Omega := (0, 1)$ and set $u_j := j\mathbb{1}_{(0,1/j)}$, see Figure 12.2. Then, $u_j \xrightarrow{\mathbf{Y}} \nu \in \mathbf{Y}^{\mathscr{M}}((0, 1))$ with

$$\nu_x = \delta_0 \quad \text{a.e.}, \qquad \lambda_\nu = \delta_0, \qquad \nu_0^\infty = \delta_{+1}.$$

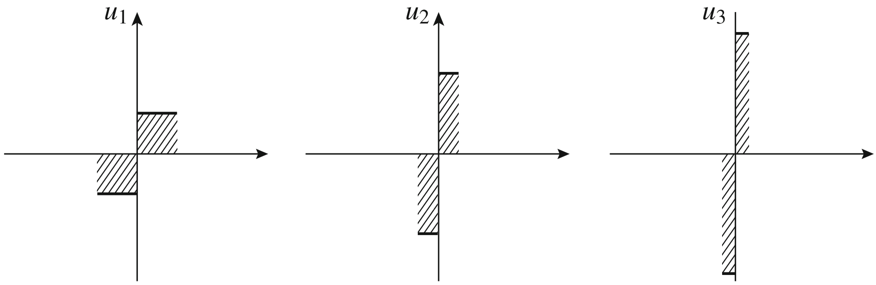


Fig. 12.3 Another concentrating sequence

Example 12.7. On $\Omega := (-1, 1)$ define $u_j := j(\mathbb{1}_{(0,1/j)} - \mathbb{1}_{(-1/j,0)})$, see Figure 12.3. Then, $u_j \xrightarrow{\mathbf{Y}} \nu \in \mathbf{Y}^{\mathscr{M}}((0,1))$ with

$$\nu_x = \delta_0 \quad \text{a.e.}, \qquad \lambda_\nu = 2\delta_0, \qquad \nu_0^\infty = \frac{1}{2}\delta_{+1} + \frac{1}{2}\delta_{-1}.$$

Example 12.8 (Diffuse concentration). On $\Omega := (0, 1)$ define the uniformly L^1-bounded sequence

$$u_j := \sum_{k=0}^{j-1} j\mathbb{1}_{\left(\frac{k}{j},\frac{k}{j}+\frac{1}{j^2}\right)}, \qquad j \in \mathbb{N},$$

see Figure 12.4. The u_j converge to zero almost everywhere and in the **biting sense**, i.e., there exists an increasing sequence of subsets $\Omega_k \subset \Omega$ with $|\Omega_k| \uparrow |\Omega|$ as $k \to \infty$ and $u_j \rightharpoonup 0$ in $\mathrm{L}^1(\Omega_k)$ for all $k \in \mathbb{N}$ (see Section 6.4 in [222] for more on this biting convergence). On the other hand, $u_j \overset{*}{\rightharpoonup} \mathbb{1}$ in the sense of measures. Consequently, the (u_j) cannot be equiintegrable; otherwise by Vitali's Convergence Theorem A.11 the limits would agree. We can also verify this directly:

$$\lim_{h\to\infty} \sup_{j\in\mathbb{N}} \int_{\{|u_j|\geq h\}} |u_j| \,\mathrm{d}x = 1.$$

Furthermore, (u_j) has no L^1-weakly converging subsequence: Since L^1-weak convergence implies weak* convergence in the sense of measures, the L^1-weak limit would have to be $\mathbb{1}$, contradicting the biting limit 0. It is the task of Problem 12.3 to show that $u_j \xrightarrow{\mathbf{Y}} \nu \in \mathbf{Y}^{\mathscr{M}}((0,1))$ with

$$\nu_x = \delta_0 \quad \text{a.e.}, \qquad \lambda_\nu = \mathscr{L}^1 \llcorner (0,1), \qquad \nu_x^\infty = \delta_{+1}.$$

Example 12.9. The last example can be modified (now with $\mathbb{R}^2$ as target space) to

$$v_j(x) := \sum_{k=0}^{j-1} j\mathbb{1}_{\left(\frac{k}{j},\frac{k}{j}+\frac{1}{j^2}\right)}(x) \begin{bmatrix} \cos(2\pi j^2 x) \\ \sin(2\pi j^2 x) \end{bmatrix}.$$

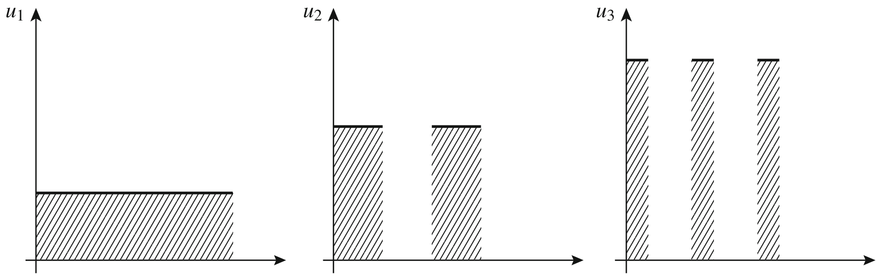


Fig. 12.4 A diffuse concentration

One can compute that $\nu_j \xrightarrow{\mathbf{Y}} \nu$ for $\nu \in \mathbf{Y}^{\mathscr{M}}((0,1);\mathbb{R}^2)$ given as

$$\nu_x = \delta_0 \quad \mathscr{L}^d\text{-a.e.}, \qquad \lambda_\nu = \mathscr{L}^1 \llcorner (0,1), \qquad \nu_x^\infty = \frac{1}{2\pi}\mathscr{H}^1 \llcorner \mathbb{S}^1.$$

Indeed, the ν_j concentrate in all directions uniformly, hence ν_x^∞ must be the uniform probability measure on $\mathbb{S}^1$.

We finish this section with a useful density result.

Lemma 12.10. *There exists a countable set* $\{f_k\}_{k\in\mathbb{N}} = \{\varphi_k \otimes h_k\}_{k\in\mathbb{N}} \subset \mathbf{E}(\Omega;\mathbb{R}^N)$ *with* $\varphi_k \in C(\overline{\Omega})$ *and* $h_k \in C(\mathbb{R}^N)$ *such that for* $\nu_j, \nu \in \mathbf{Y}^{\mathscr{M}}(\Omega;\mathbb{R}^N)$ *the condition*

$$\langle\!\langle f_k, \nu_j \rangle\!\rangle \to \langle\!\langle f_k, \nu \rangle\!\rangle \quad \textit{for all } k \in \mathbb{N}$$

implies $\nu_j \overset{*}{\rightharpoonup} \nu$. *Moreover, all the* h_k *can be chosen to be Lipschitz continuous.*

Proof. First, assuming that $f := \mathbb{1} \otimes |\bullet|$ is in the collection, we may always suppose that the sequence (ν_j) converges weakly*. It remains to identify the limit.

Take countable sets $\mathscr{A} \subset \mathrm{C}(\overline{\Omega})$, $\mathscr{B} \subset \mathrm{C}_c^1(\mathbb{R}^N)$ that are dense in $\mathrm{C}(\overline{\Omega})$ and $\mathrm{C}_0(\mathbb{R}^N)$, respectively, in the $\|\bullet\|_\infty$-norm. Furthermore, let $\mathscr{C} \subset \mathrm{C}^1(\mathbb{S}^{N-1})$ be countable and dense in $\mathrm{C}(\mathbb{S}^{N-1})$ with $\mathbb{1} \in \mathscr{C}$. Let $(Gh^\infty)(A) := |A|h^\infty(A/|A|)$ $(A \in \mathbb{R}^N)$ for $h^\infty \in \mathscr{C}$ and define

$$\{\varphi_k \otimes h_k\}_k := (\mathscr{A} \otimes \mathscr{B}) \cup (\mathscr{A} \otimes G(\mathscr{C})) \subset \mathbf{E}(\Omega;\mathbb{R}^N),$$

where the tensor product is understood to act elementwise. By standard results of measure theory, the values of

$$\langle\!\langle \varphi \otimes h, \nu \rangle\!\rangle = \int_\Omega \varphi(x)\langle h, \nu_x \rangle \,\mathrm{d}x \qquad \text{for all } \varphi \in \mathscr{A}, h \in \mathscr{B}$$

determine the $\mathrm{L}^1(\Omega)$-function $x \mapsto \langle h, \nu_x \rangle$ and then in turn the measures ν_x $\mathscr{L}^d$-almost everywhere. Testing with $\mathscr{A} \otimes (G\mathbb{1})$ gives for all $\varphi \in \mathrm{C}(\overline{\Omega})$ that

$$\langle\!\langle \varphi \otimes G\mathbb{1}, \nu \rangle\!\rangle = \int_\Omega \varphi(x)\langle G\mathbb{1}, \nu_x\rangle \,\mathrm{d}x + \int_{\overline{\Omega}} \varphi(x)\langle \mathbb{1}, \nu_x^\infty\rangle \,\mathrm{d}\lambda_\nu(x).$$

Since the first integral is already identified and the second integral reduces to $\int_{\overline{\Omega}} \varphi \,\mathrm{d}\lambda_\nu$, the measure λ_ν is also uniquely determined. The identification of ν_x^∞ (up to a
λ_ν-negligible set) is similar to that of ν_x.

By construction, every $h_k \in \mathscr{B}$ is Lipschitz continuous. If $h_k = G(h^\infty)$ with $h^\infty \in \mathscr{C}$, then there is a constant $C > 0$ with

$$\left| h^\infty\left(\frac{A}{|A|}\right) - h^\infty\left(\frac{B}{|B|}\right)\right| \le C\left|\frac{A}{|A|} - \frac{B}{|B|}\right| \qquad \text{for all } A, B \in \mathbb{R}^N \setminus \{0\}.$$

We may estimate

$$\begin{aligned} \big|h_k(A) - h_k(B)\big| &\le \left| h^\infty\left(\frac{A}{|A|}\right) - h^\infty\left(\frac{B}{|B|}\right)\right||B| + \left| h^\infty\left(\frac{A}{|A|}\right)\right||A - B| \\ &\le C\left|\frac{A}{|A|}|B| - B\right| + \Big(\max_{\mathbb{S}^{N-1}} |h^\infty|\Big)|A - B| \\ &\le \Big(2C + \max_{\mathbb{S}^{N-1}} |h^\infty|\Big)|A - B|, \end{aligned}$$

hence those h_k are also Lipschitz continuous. □

12.3 Extended Representation

We now extend the representation of limits via Young measures to a larger class of integrands than $\mathbf{E}(\Omega; \mathbb{R}^N)$, called the **representation integrands** and defined as follows:

$$\mathbf{R}(\Omega; \mathbb{R}^N) := \Big\{ f : \overline{\Omega} \times \mathbb{R}^N \to \mathbb{R} \;:\; f \text{ Carathéodory with linear growth and } f^\infty \in \mathrm{C}(\overline{\Omega} \times \mathbb{R}^N) \text{ exists as in (11.8)} \Big\}.$$

Note that we do not identify integrands that are equal almost everywhere.

Proposition 12.11. *Let* $\nu_j \overset{*}{\rightharpoonup} \nu$ *in* $\mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^N)$ *and assume either*

(i) $f \in \mathbf{R}(\Omega; \mathbb{R}^N)$ *or*
(ii) $f(x, A) = \mathbb{1}_B(x) g(x, A)$, *where* $g \in \mathbf{E}(\Omega; \mathbb{R}^N)$ *and* $B \subset \overline{\Omega}$ *is a Borel set with* $(\mathscr{L}^d + \lambda_\nu)(\partial B) = 0$.

Then, $\langle\!\langle f, \nu_j \rangle\!\rangle \to \langle\!\langle f, \nu \rangle\!\rangle$.

Before we come to the proof, we note that for $f \in \mathbf{R}(\Omega; \mathbb{R}^N)$ the expression

$$\langle\!\langle f, \nu \rangle\!\rangle = \int_\Omega \int_{\mathbb{R}^N} f(x, A) \, \mathrm{d}\nu_x(A) \, \mathrm{d}x + \int_{\overline{\Omega}} \int_{\mathbb{S}^{N-1}} f^\infty(x, A) \, \mathrm{d}\nu_x^\infty(A) \, \mathrm{d}\lambda_\nu(x)$$

is well-defined by the weak* measurability of $(\nu_x)_x$ with respect to $\mathscr{L}^d$ and the weak* measurability of $(\nu_x^\infty)_x$ with respect to λ_ν. In case (ii), where f^∞ is not continuous, the discontinuity set is Borel-measurable and $(\mathscr{L}^d + \lambda_\nu)$-negligible, so the above expression is still well-defined.

Proof. We will show the representation for a Carathéodory integrand $f : \overline{\Omega} \times \mathbb{R}^N \to \mathbb{R}$ that possesses a jointly continuous recession function $f^\infty : (\overline{\Omega} \setminus Z) \times \mathbb{R}^N \to \mathbb{R}$ in the sense of (11.8), where $Z \subset \overline{\Omega}$ is a Borel set with $(\mathscr{L}^d + \lambda_\nu)(Z) = 0$. This implies both (i) and (ii).

Step 1. First, we assume that $f(x, \bullet)$ has uniformly bounded support, that is, $\operatorname{supp} f(x, \bullet) \Subset B(0, R)$ for all $x \in \overline{\Omega}$ and a fixed $R > 0$. Since $f^\infty \equiv 0$, we need to show

$$\int_\Omega \big\langle f(x, \bullet), (\nu_j)_x \big\rangle \, \mathrm{d}x \to \int_\Omega \big\langle f(x, \bullet), \nu_x \big\rangle \, \mathrm{d}x.$$

Let $\{\psi_k\}_k \subset \mathrm{C}_0(B(0, R/(1+R)))$ be a countable and dense family and fix $\varepsilon > 0$. Set

$$E_k := \Big\{ x \in \overline{\Omega} \; : \; \|Sf(x, \bullet) - \psi_k\|_\infty \le \varepsilon \Big\}, \qquad k \in \mathbb{N},$$

which is a measurable set since, by the continuity of $f(x, \bullet)$ for fixed $x \in \overline{\Omega}$,

$$E_k = \bigcap_{A \in \mathbb{Q}^N} \Big\{ x \in \overline{\Omega} \; : \; |Sf(x, A) - \psi_k(A)| \le \varepsilon \Big\}.$$

We have that $\bigcup_{k \in \mathbb{N}} E_k = \overline{\Omega}$ because for any fixed $x \in \overline{\Omega}$ it holds that $Sf(x, \bullet) \in \mathrm{C}_0(B(0, R/(1+R)))$ and thus $x \in E_k$ for at least one $k \in \mathbb{N}$. Consequently, the sets $F_k := E_k \setminus \bigcup_{i=1}^{k-1} F_i$ form a measurable disjoint partition of $\overline{\Omega}$. Define

$$g_\varepsilon(x, A) := \sum_{k=1}^\infty \mathbb{1}_{F_k}(x) S^{-1}\psi_k(A), \qquad (x, A) \in \overline{\Omega} \times \mathbb{R}^N.$$

Then, $\|Sf - Sg_\varepsilon\|_\infty \le \varepsilon$, whereby in particular

$$\|g_\varepsilon\|_\infty \le (1 + R)(\|Sf\|_\infty + \varepsilon).$$

For each fixed $k \in \mathbb{N}$ we infer from $\nu_j \overset{*}{\rightharpoonup} \nu$ that

$$\int \varphi(x)\big\langle S^{-1}\psi_k, (\nu_j)_x\big\rangle \,\mathrm{d}x \to \int \varphi(x)\big\langle S^{-1}\psi_k, \nu_x\big\rangle \,\mathrm{d}x, \qquad \varphi \in \mathrm{C}^\infty(\overline{\Omega}).$$

Thus, for all $k \in \mathbb{N}$,

$$\int_{F_k} \big\langle S^{-1}\psi_k, (\nu_j)_x\big\rangle \,\mathrm{d}x \to \int_{F_k} \big\langle S^{-1}\psi_k, \nu_x\big\rangle \,\mathrm{d}x \tag{12.8}$$

because we can L^1-approximate $\mathbb{1}_{F_k}$ by smooth functions. Since

$$\left|\int_{F_k} \big\langle S^{-1}\psi_k, (\nu_j)_x\big\rangle \,\mathrm{d}x\right| \le (1+R)(\|Sf\|_\infty + \varepsilon)|F_k|,$$

we can then use the dominated convergence theorem for sums and (12.8) to compute

$$\begin{aligned}
\lim_{j\to\infty} \int_\Omega \big\langle g_\varepsilon(x, \bullet), (\nu_j)_x\big\rangle \,\mathrm{d}x &= \lim_{j\to\infty} \sum_{k=1}^\infty \int_{F_k} \big\langle S^{-1}\psi_k, (\nu_j)_x\big\rangle \,\mathrm{d}x \\
&= \sum_{k=1}^\infty \lim_{j\to\infty} \int_{F_k} \big\langle S^{-1}\psi_k, (\nu_j)_x\big\rangle \,\mathrm{d}x \\
&= \sum_{k=1}^\infty \int_{F_k} \big\langle S^{-1}\psi_k, \nu_x\big\rangle \,\mathrm{d}x \\
&= \int_\Omega \big\langle g_\varepsilon(x, \bullet), \nu_x\big\rangle \,\mathrm{d}x.
\end{aligned}$$

Moreover, $\|Sf - Sg_\varepsilon\|_\infty \le \varepsilon$, and so,

$$\left|\int_\Omega \big\langle g_\varepsilon(x, \bullet), (\nu_j)_x\big\rangle \,\mathrm{d}x - \int_\Omega \big\langle f(x, \bullet), (\nu_j)_x\big\rangle \,\mathrm{d}x\right| \le \varepsilon \int_\Omega \big\langle 1 + |\bullet|, (\nu_j)_x\big\rangle \,\mathrm{d}x \le \varepsilon C$$

for some j-independent constant $C > 0$; the same holds with ν in place of ν_j. Combining these arguments,

$$\left|\lim_{j\to\infty} \langle\!\langle f, \nu_j\rangle\!\rangle - \langle\!\langle f, \nu\rangle\!\rangle\right| \le \left|\lim_{j\to\infty} \langle\!\langle g_\varepsilon, \nu_j\rangle\!\rangle - \langle\!\langle g_\varepsilon, \nu\rangle\!\rangle\right| + 2\varepsilon C = 2\varepsilon C.$$

As $\varepsilon > 0$ was arbitrary, we have proved the assertion in the case when f has uniformly bounded support.

Step 2. Next, we extend the representation to Carathéodory integrands $f \colon \overline{\Omega} \times \mathbb{R}^N \to \mathbb{R}$ with $f^\infty \equiv 0$. Fix $\varepsilon > 0$ and let $r \in (0, 1)$ be so large that $|Sf(x, A)| \le \varepsilon$ for all $x \in \overline{\Omega}$ and $A \in \mathbb{B}^N$ with $|A| \ge r$. We can see that such an $r > 0$ must exist since otherwise we could find sequences $(x_n) \subset \overline{\Omega}$ and $(A_n) \subset \mathbb{B}^N$ with $|A_n| \geq 1 - n^{-1}$ and $Sf(x_n, A_n) \geq \varepsilon$. Without loss of generality we may assume $x_n \to x \in \overline{\Omega}$, $A_n \to A \in \mathbb{S}^{N-1}$ and so f^∞ cannot be zero everywhere.

Select a cut-off function $\rho \in C_c^\infty(\mathbb{R}^N; [0, 1])$ with $\rho \equiv 1$ on $B(0, r/(1-r))$. Then, writing $f\rho$ for the function $(x, A) \mapsto f(x, A)\rho(A)$, where $x \in \overline{\Omega}$, $A \in \mathbb{R}^N$, we have

$$\left|\left\langle\!\left\langle f\rho - f, \nu_j \right\rangle\!\right\rangle\right| \leq \varepsilon \left[\int_\Omega \left\langle 1 + |\bullet|, (\nu_j)_x \right\rangle \, \mathrm{d}x + \lambda_{\nu_j}(\overline{\Omega}) \right] \leq \varepsilon C$$

for some constant $C > 0$ and all $j \in \mathbb{N}$; the same holds with ν in place of ν_j. By the previous step,

$$\left\langle\!\left\langle f\rho, \nu_j \right\rangle\!\right\rangle \to \left\langle\!\left\langle f\rho, \nu \right\rangle\!\right\rangle$$

and so, combining with the previous estimate, the conclusion follows in the case when $f^\infty \equiv 0$.

Step 3. Finally, when f^∞ is not identically zero, we write

$$f = g + f^\infty \qquad \text{with} \qquad g^\infty \equiv 0.$$

The last step applies to g and we get

$$\left\langle\!\left\langle g, \nu_j \right\rangle\!\right\rangle \to \left\langle\!\left\langle g, \nu \right\rangle\!\right\rangle. \tag{12.9}$$

To investigate the convergence for the positively 1-homogeneous functions f^∞, we define the measures $\mu_j := S^{-*}\nu_j \in \mathscr{M}^+(\overline{\Omega \times \mathbb{B}^N})$. Then, $\mu_j \overset{*}{\rightharpoonup} \mu := S^{-*}\nu$ and

$$\left\langle \Phi, \mu_j \right\rangle = \int_\Omega \left\langle S^{-1}\Phi(x, \bullet), (\nu_j)_x \right\rangle \, \mathrm{d}x + \int_{\overline{\Omega}} \left\langle \Phi(x, \bullet), (\nu_j)_x^\infty \right\rangle \, \mathrm{d}\lambda_{\nu_j}(x)$$

for all $\Phi \in C(\overline{\Omega \times \mathbb{B}^N})$. Testing this with all Φ such that $\|\Phi\|_\infty \leq 1$, we get

$$|\mu_j| \leq \langle 1 + |\bullet|, (\nu_j)_x \rangle \, \mathscr{L}^d \llcorner \Omega + \lambda_{\nu_j} \overset{*}{\rightharpoonup} \langle 1 + |\bullet|, \nu_x \rangle \, \mathscr{L}^d \llcorner \Omega + \lambda_\nu.$$

Denote by Λ the weak* limit of the $|\mu_j|$ in $\mathscr{M}^+(\overline{\Omega})$. Then, $(\mathscr{L}^d + \lambda_\nu)(Z) = 0$ implies $\Lambda(Z \times \overline{\mathbb{B}^N}) = 0$. By standard results in measure theory (see Lemma A.22), for any bounded Borel function $\Psi \colon \overline{\Omega \times \mathbb{B}^N} \to \mathbb{R}$ with a Λ-negligible set of discontinuity points we have

$$\left\langle \Psi, \mu_j \right\rangle \to \left\langle \Psi, \mu \right\rangle.$$

For $\Psi := Sf^\infty$, which by assumption is continuous outside $Z \times \overline{\mathbb{B}^N}$, this gives

$$\left\langle\!\left\langle f^\infty, \nu_j \right\rangle\!\right\rangle = \left\langle Sf^\infty, \mu_j \right\rangle \to \left\langle Sf^\infty, \mu \right\rangle = \left\langle\!\left\langle f^\infty, \nu \right\rangle\!\right\rangle.$$

Combining this with (12.9), we arrive at

$$\langle\langle f, \nu_j \rangle\rangle = \langle\langle g, \nu_j \rangle\rangle + \langle\langle f^\infty, \nu_j \rangle\rangle \to \langle\langle g, \nu \rangle\rangle + \langle\langle f^\infty, \nu \rangle\rangle = \langle\langle f, \nu \rangle\rangle.$$

This shows the assertion at the beginning of the proof and thus (i) and (ii). □

Example 12.12. On $\Omega := (-1, 1)$ let $f(x, A) := \mathbb{1}_{(0,1)}(x)|A|$, for which $f^\infty(x, A) = \mathbb{1}_{(0,1)}(x)|A|$. Then, for $\nu_j \in \mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^N)$ given as

$$(\nu_j)_x := \delta_0 \quad \text{a.e.,} \qquad \lambda_{\nu_j} := \delta_{1/j}, \qquad (\nu_j)^\infty_{1/j} := \delta_{+1}$$

we have $\nu_j \overset{*}{\rightharpoonup} \nu$ for

$$\nu_x = \delta_0 \quad \text{a.e.,} \qquad \lambda_\nu = \delta_0, \qquad \nu^\infty_0 = \delta_{+1}.$$

On the other hand,

$$\lim_{j\to\infty} \langle\langle f, \nu_j \rangle\rangle = \lim_{j\to\infty} \int_{(0,1)} \mathbb{1}_{(0,1)}(x)\, \mathrm{d}\delta_{1/j}(x) = 1 \neq 0 = \langle\langle f, \nu \rangle\rangle.$$

Here, the discontinuity set $\{0, 1\}$ of f is not negligible with respect to $\mathscr{L}^d + \delta_0$. This example therefore shows that in Proposition 12.11 (ii) the assumption $(\mathscr{L}^d + \lambda_\nu)(\partial B) = 0$ is necessary.

12.4 Strong Precompactness of Sequences

By Vitali's Convergence Theorem A.11, the absence of oscillations and concentrations implies strong precompactness of an L^1-bounded sequence. In this section we study in more detail how compactness properties are reflected in the generated Young measure.

We start with oscillations:

Lemma 12.13. *Let $(V_j) \subset \mathrm{L}^1(\Omega; \mathbb{R}^N)$ with $V_j \overset{\mathbf{Y}}{\to} \nu \in \mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^N)$. Then, the sequence (V_j) converges in measure to $V \in \mathrm{L}^1(\Omega; \mathbb{R}^N)$ if and only if $\nu_x = \delta_{V(x)}$ almost everywhere.*

Proof. The proof is very similar to that of Lemma 4.12 once we observe that the integrand

$$f(x, A) := \frac{|A - V(x)|}{1 + |A - V(x)|}, \qquad (x, A) \in \Omega \times \mathbb{R}^N,$$

from that proof lies in $\mathbf{R}(\Omega; \mathbb{R}^N)$ and $f^\infty \equiv 0$. □

Next, we show how the concentration parts λ_ν, $(\nu^\infty_x)_x$ of a Young measure $\nu \in \mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^N)$ reflect the equiintegrability properties of the generating sequence.

Lemma 12.14. *Let* $(V_j) \subset \mathrm{L}^1(\Omega;\mathbb{R}^N)$ *with* $V_j \overset{\mathbf{Y}}{\to} \nu \in \mathbf{Y}^{\mathscr{M}}(\Omega;\mathbb{R}^N)$.

(i) The sequence (V_j) *is equiintegrable if and only if* $\lambda_\nu = 0$.

(ii) For $f \in \mathbf{E}(\Omega;\mathbb{R}^N)$ *let* $F_j(x) := f(x, V_j(x))$ $(x \in \Omega)$*. Then, the sequence* (F_j) *is equiintegrable if and only if* $\langle |f^\infty(x,\bullet)|, \nu_x^\infty\rangle = 0$ *for* λ_ν*-almost every* $x \in \overline{\Omega}$.

Proof. We only need to show (ii), for (i) take $f(x,A) := |A|$. So, let $f \in \mathbf{E}(\Omega;\mathbb{R}^N)$ and $R > 0$. Set

$$\eta_R := \limsup_{j\to\infty} \int_{\{|F_j|\geq R\}} |F_j| \,\mathrm{d}x, \qquad \eta_\infty := \lim_{R\uparrow\infty} \eta_R.$$

The family (F_j) is equiintegrable if and only if $\eta_\infty = 0$ (see Appendix A.3). In the following we will prove the formula

$$\eta_\infty = \int_{\overline{\Omega}} \bigl\langle |f^\infty(x,\bullet)|, \nu_x^\infty \bigr\rangle \,\mathrm{d}\lambda_\nu(x), \tag{12.10}$$

which implies (ii).

Let, for $t \geq 0$,

$$h(t) := \begin{cases} 0 & \text{if } 0 < t < \frac{1}{2}, \\ 2t-1 & \text{if } \frac{1}{2} \leq t \leq 1, \\ t & \text{if } t > 1, \end{cases} \qquad h_R(t) := Rh\Bigl(\frac{t}{R}\Bigr).$$

We have

$$h_{2R}(t) \leq t\mathbb{1}_{[R,\infty)}(t) \leq h_R(t) \quad \text{for all } t \geq 0, \qquad (h_R \circ |f|)^\infty = |f^\infty|.$$

This allows us to estimate

$$\int_\Omega h_{2R}(|F_j(x)|) \,\mathrm{d}x \leq \int_{\{|F_j|\geq R\}} |F_j(x)| \,\mathrm{d}x \leq \int_\Omega h_R(|F_j(x)|) \,\mathrm{d}x.$$

Letting $j \to \infty$, we arrive at

$$\begin{aligned} &\int_\Omega \bigl\langle h_{2R} \circ |f(x,\bullet)|, \nu_x \bigr\rangle \,\mathrm{d}x + \int_{\overline{\Omega}} \bigl\langle |f^\infty(x,\bullet)|, \nu_x^\infty \bigr\rangle \,\mathrm{d}\lambda_\nu(x) \\ &\quad \leq \eta_R \leq \int_\Omega \bigl\langle h_R \circ |f(x,\bullet)|, \nu_x \bigr\rangle \,\mathrm{d}x + \int_{\overline{\Omega}} \bigl\langle |f^\infty(x,\bullet)|, \nu_x^\infty \bigr\rangle \,\mathrm{d}\lambda_\nu(x). \end{aligned}$$

For $R \to \infty$ the first integral in both the first and the last expression vanishes and (12.10) follows. □

We can now combine the last two lemmas via Vitali's Convergence Theorem A.11 to decide whether a L^1-bounded sequences is strongly precompact from the generated Young measure:

Corollary 12.15. *Let* $(V_j) \subset \mathrm{L}^1(\Omega; \mathbb{R}^N)$ *with* $V_j \overset{\mathbf{Y}}{\to} \nu \in \mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^N)$. *Then,* $V_j \to V$ *in* L^1 *if and only if* $\nu_x = \delta_{V(x)}$ *almost everywhere and* $\lambda_\nu = 0$.

12.5 BV-Young Measures

Like for classical Young measures, the most important subclass of generalized Young measures are those generated by gradients. Here, we consider gradients of $\mathrm{W}^{1,1}$-maps or, more generally, BV-derivatives. We thus specialize the framework defined in the preceding sections to $\mathbb{R}^N = \mathbb{R}^{m\times d}$ and we denote by $\mathbb{B}^{m\times d}$ and $\partial\mathbb{B}^{m\times d}$ the unit ball and the unit sphere in $\mathbb{R}^{m\times d}$, respectively.

For $u \in \mathrm{BV}(\Omega; \mathbb{R}^m)$ we associate with $Du \in \mathscr{M}(\Omega; \mathbb{R}^{m\times d})$ the elementary Young measure $\delta[Du] \in \mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^{m\times d})$ as before, that is,

$$\delta[Du]_x := \delta_{\nabla u(x)} \quad \mathscr{L}^d\text{-a.e.}, \qquad \lambda_{\delta[Du]} := |D^s u|, \qquad \delta[Du]_x^\infty := \delta_{P(x)} \quad |D^s u|\text{-a.e.},$$

where

$$P := \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|} \in \mathrm{L}^1(\Omega, |D^s u|; \partial\mathbb{B}^{m\times d}).$$

For $(u_j) \subset \mathrm{BV}(\Omega; \mathbb{R}^m)$ the sequence (Du_j) generates the Young measure $\nu \in \mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^{m\times d})$, in symbols "$Du_j \overset{\mathbf{Y}}{\to} \nu$", if

$$\delta[Du_j] \overset{*}{\rightharpoonup} \nu \quad \text{in } \mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^{m\times d}).$$

We collect all these **BV-Young measures** ν in the set

$$\mathbf{BVY}(\Omega; \mathbb{R}^{m\times d}) \subset \mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^{m\times d}),$$

where we again identify equivalent Young measures as we did for $\mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^{m\times d})$.

The Fundamental Theorem 12.5 adapts to our BV-context as follows:

Theorem 12.16. *Let* $(u_j) \subset \mathrm{BV}(\Omega; \mathbb{R}^m)$ *be a uniformly norm-bounded sequence. Then, there exists a subsequence (not explicitly labeled) such that* $Du_j \overset{\mathbf{Y}}{\to} \nu$ *for some* $\nu \in \mathbf{BVY}(\Omega; \mathbb{R}^{m\times d})$.

All the examples from Section 12.2 are in fact BV-Young measures since the generating sequences are defined on a one-dimensional domain and thus are trivially derivatives.

The restriction of the barycenter $[\nu]$ to Ω is a BV-derivative since for a sequence $u_j \overset{*}{\rightharpoonup} u$ in $\mathrm{BV}(\Omega; \mathbb{R}^m)$ and the integrand $f(x, A) := A$, it follows from (12.3) that

$$Du_j \overset{*}{\rightharpoonup} [\nu] = [\nu_x]\, \mathscr{L}_x^d \llcorner \Omega + [\nu_x^\infty]\, \lambda_\nu(\mathrm{d}x) \quad \text{in } \mathscr{M}(\overline{\Omega}; \mathbb{R}^{m\times d}).$$

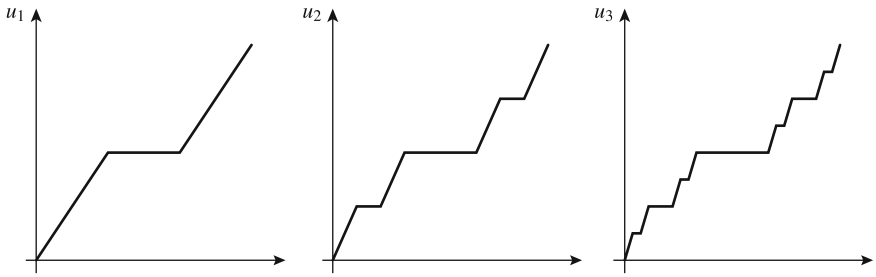


Fig. 12.5 The construction of the standard $1/3$-Cantor function

On the other hand, $Du_j \overset{*}{\rightharpoonup} Du$ in $\mathcal{M}(\Omega; \mathbb{R}^{m\times d})$, and so

$$Du = [\nu] \llcorner \Omega = [\nu_x]\, \mathcal{L}^d_x \llcorner \Omega + [\nu_x^\infty]\, (\lambda_\nu \llcorner \Omega)(\mathrm{d}x).$$

Any $u \in \mathrm{BV}(\Omega; \mathbb{R}^m)$ with $Du = [\nu] \llcorner \Omega$ is called an **underlying deformation** of ν.

If $\lambda_\nu^s \llcorner \Omega = g\,|D^s u| + \lambda_\nu^*$ is the Lebesgue–Radon–Nikodým decomposition of $\lambda_\nu^s \llcorner \Omega$ with respect to $|D^s u|$, then

$$\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x) = [\nu_x^\infty]\, g(x) \qquad \text{for } |D^s u|\text{-a.e. } x \in \Omega$$

and

$$[\nu_x^\infty] = 0 \qquad \text{for } \lambda_\nu^*\text{-a.e. } x \in \Omega.$$

Then, by Alberti's Rank-One Theorem 10.7, the matrix $[\nu_x^\infty]$ has rank one for $|D^s u|$-almost every $x \in \Omega$ (note that $g(x) \neq 0$ for $|D^s u|$-almost every $x \in \Omega$). Thus,

$$\operatorname{rank}\, [\nu_x^\infty] \leq 1 \qquad \text{for } \lambda_\nu^s\text{-a.e. } x \in \Omega. \tag{12.11}$$

Before we proceed with the abstract theory, we give another example of a BV-Young measure.

Example 12.17 (Cantor functions). Fix $\delta \in (0, 1/2)$ and let C_j be the j'th set in the construction of the δ-Cantor set, i.e., $C_0 := (0, 1)$ and we obtain C_{j+1} from C_j by removing from each interval in C_j the centered open interval of length $\delta^j(1 - 2\delta)$. The usual Cantor set is obtained for $\delta = 1/3$. Then, the δ-Cantor set is

$$C := \bigcap_{j\in\mathbb{N}} C_j.$$

The maps (see Figure 12.5)

$$u_j(x) := \frac{1}{(2\delta)^j} \int_0^x \mathbb{1}_{C_j} \, \mathrm{d}y$$

converge area-strictly in $\mathrm{BV}(0,1)$ to the **δ-Cantor function** $\chi_\delta \in \mathrm{C}((0,1))$, where

$$\chi_\delta(x) = \alpha^{-1}\mathscr{H}^\gamma(C \cap [0,x]), \qquad D\chi_\delta = \alpha^{-1}\mathscr{H}^\gamma \llcorner C,$$

for

$$\gamma = \frac{\ln 2}{\ln(1/\delta)}, \qquad \alpha = \mathscr{H}^\gamma(C) = 2^{-\gamma}\frac{\pi^{\gamma/2}}{\Gamma(1+\gamma/2)},$$

where Γ is Euler's Γ-function, see p. 60 ff. in [183] for the details. By Proposition 12.4 the Young measure $\nu \in \mathbf{BVY}((0,1))$ generated by (Du_j) is

$$\nu_x = \delta_0 \quad \mathscr{L}^d\text{-a.e.}, \qquad \lambda_\nu = \alpha^{-1}\,\mathscr{H}^\gamma \llcorner C, \qquad \nu_x^\infty = \delta_{+1} \quad \mathscr{H}^\gamma - \text{a.e.}$$

The following result on good generating sequences is often useful.

Proposition 12.18. *Let $\nu \in \mathbf{BVY}(\Omega;\mathbb{R}^{m\times d})$.*

(i) There exists a sequence $(u_j) \subset (W^{1,1} \cap C^\infty)(\Omega;\mathbb{R}^m)$ with $\nabla u_j \overset{\mathbf{Y}}{\to} \nu$.
(ii) If $\lambda_\nu(\partial\Omega) = 0$, then the u_j from (i) can be chosen to satisfy $u_j|_{\partial\Omega} = u|_{\partial\Omega}$ for any underlying deformation $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ of ν.

See Problem 12.5 for the necessity of the assumption $\lambda_\nu(\partial\Omega) = 0$ in (ii).

Proof. Ad (i). Let $(v_j) \subset \mathrm{BV}(\Omega;\mathbb{R}^m)$ be a generating sequence for ν, that is, $Dv_j \overset{\mathbf{Y}}{\to} \nu$. We also take a countable collection $\{f_k\} \subset \mathbf{E}(\Omega;\mathbb{R}^{m\times d})$ that determines the Young measure convergence as in Lemma 12.10. Then, from the area-strict density of smooth functions in $\mathrm{BV}(\Omega;\mathbb{R}^m)$, see Lemma 11.1, in conjunction with Proposition 12.4 we construct for each $j \in \mathbb{N}$ a map $u_j \in (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(\Omega;\mathbb{R}^m)$ with

$$\left| \int_\Omega f_k(x,\nabla v_j(x))\,\mathrm{d}x + \int_\Omega f_k^\infty\left(x, \frac{\mathrm{d}D^s v_j}{\mathrm{d}|D^s v_j|}(x)\right)\mathrm{d}|D^s v_j|(x) - \int_\Omega f_k(x,\nabla u_j(x))\,\mathrm{d}x \right| \le \frac{1}{j}$$

whenever $k \le j$. Then,

$$\int_\Omega f_k(x,\nabla u_j(x))\,\mathrm{d}x \to \langle\!\langle f_k,\nu\rangle\!\rangle \qquad \text{as } j\to\infty \qquad \text{for all } k\in\mathbb{N},$$

and so, Lemma 12.10 implies $\nabla u_j \overset{\mathbf{Y}}{\to} \nu$.

Ad (ii). From (i) we know that there exist a sequence $(v_j) \subset (\mathrm{W}^{1,1}\cap\mathrm{C}^\infty)(\Omega;\mathbb{R}^m)$ with $Dv_j \overset{\mathbf{Y}}{\to} \nu$. We may add constants to the v_j and take another subsequence such that in addition $v_j \to u$ in L^1. Then choose a sequence $(\rho_n) \subset \mathrm{C}_c^\infty(\Omega;[0,1])$ of cut-off functions with $\rho_n \uparrow \mathbb{1}_\Omega$ pointwise as $n\to\infty$ and such that for

$$K_n := \left\{ x\in\Omega \;:\; \rho_n(x) = 1 \right\}$$

it holds that

$$(\mathscr{L}^d + \lambda_\nu)(\partial K_n) = 0.$$

Now employ Lemma 11.1 to get a sequence $(w_j) \subset (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(\Omega; \mathbb{R}^m)$ such that $w_j|_{\partial\Omega} = u|_{\partial\Omega}$ and $w_j \to u$ area-strictly. Set

$$u_{j,n} := \rho_n v_j + (1 - \rho_n) w_j \in \mathrm{W}^{1,1}(\Omega; \mathbb{R}^m).$$

Then, $u_{j,n}|_{\partial\Omega} = u|_{\partial\Omega}$ and

$$\nabla u_{j,n} = \rho_n \nabla v_j + (1 - \rho_n)\nabla w_j + (v_j - w_j) \otimes \nabla \rho_n$$

and

$$\lim_{n\to\infty} \lim_{j\to\infty} \|u_{j,n} - u\|_{\mathrm{L}^1(\Omega;\mathbb{R}^m)} = 0.$$

Thus, for $f \in \mathbf{E}(\Omega; \mathbb{R}^{m\times d})$ with linear-growth constant $M > 0$ we can estimate

$$\begin{aligned}
&\left|\langle\!\langle f, \delta[Du_{j,n}]\rangle\!\rangle - \langle\!\langle f, \delta[Dv_j]\rangle\!\rangle\right| \\
&\quad\le \int_{\Omega\setminus K_n} |f(\nabla u_{j,n})| + |f(\nabla v_j)| \,\mathrm{d}x \\
&\quad\le M \int_{\Omega\setminus K_n} 2 + 2|\nabla v_j| + |\nabla w_j| + |v_j - w_j||\nabla \rho_n| \,\mathrm{d}x.
\end{aligned}$$

Hence, since $(\mathscr{L}^d + \lambda_\nu)(\partial K_n) = 0$,

$$\begin{aligned}
&\limsup_{n\to\infty} \limsup_{j\to\infty} \left|\langle\!\langle f, \delta[Du_{j,n}]\rangle\!\rangle - \langle\!\langle f, \delta[Dv_j]\rangle\!\rangle\right| \\
&\quad\le \limsup_{n\to\infty} 2M\left(\int_{\overline{\Omega}\setminus K_n} \left\langle 1 + |\bullet|, \nu_x \right\rangle \mathrm{d}x + \lambda_\nu(\overline{\Omega} \setminus K_n) + |Du|(\Omega \setminus K_n)\right) \\
&\quad= 2M\lambda_\nu(\partial\Omega) \\
&\quad= 0.
\end{aligned}$$

We can then select a diagonal sequence $u_j := u_{j,n(j)}$ with the desired properties. □

Like for classical Young measures, the question arises whether one can characterize the subclass of BV-Young measures $\mathbf{BVY}(\Omega; \mathbb{R}^{m\times d})$ in $\mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^{m\times d})$. The analogue of the Kinderlehrer–Pedregal Theorem 7.15 is the following:

Theorem 12.19 (Kristensen–Rindler 2010 [168]). *Let $\nu \in \mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^{m\times d})$ with*

$$\lambda_\nu(\partial\Omega) = 0. \tag{12.12}$$

*Then, $\nu \in \mathbf{BVY}(\Omega; \mathbb{R}^{m\times d})$ if and only if there exists a map $u \in \mathrm{BV}(\Omega; \mathbb{R}^m)$ with $[\nu] = Du$ and for all quasiconvex $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ with linear growth the **regular Jensen-type inequality***

$$h\left([\nu_x] + [\nu_x^\infty]\frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x)\right) \leq \langle h, \nu_x\rangle + \langle h^\#, \nu_x^\infty\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x)$$

holds at almost every $x \in \Omega$.

Remark 12.20. It turns out that in the situation of the preceding theorem for all quasiconvex $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ with linear growth the **singular Jensen-type inequality**

$$h^\#([\nu_x^\infty]) \leq \langle h^\#, \nu_x^\infty\rangle$$

also holds at λ_ν^s-almost every $x \in \Omega$, see Proposition 12.27.

The proof of Theorem 12.19 is quite long and involved, so we omit it here. The interested reader is referred to [229] for a direct argument and also to [162] for a refinement. An extension of the above theorem without the assumption (12.12) is in [23].

12.6 Localization

In Proposition 5.14 we saw how we can "blow-up" or "localize" a classical gradient Young measure $\nu = (\nu_x)_x \in \mathbf{GY}^p(\Omega; \mathbb{R}^{m\times d})$, $p \in [1, \infty)$. As a result, for almost every $x_0 \in \Omega$ the probability measure ν_{x_0} is a homogeneous Young measure in its own right. The discussion in this section parallels these developments for generalized Young measures. As the barycenter is now a measure instead of a function, however, we need to employ the theory of tangent measures from Section 10.2 and we will also distinguish between *regular* and *singular* blow-ups.

Before we come to the localization principles, we need to introduce local versions of generalized Young measures. Define $\mathbf{Y}^{\mathscr{M}}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$ like $\mathbf{Y}^{\mathscr{M}}(\mathbb{R}^d; \mathbb{R}^N)$ but with λ_ν only in $\mathscr{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ and $x \mapsto \langle|\bullet|, \nu_x\rangle \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$. Now, $\mathbf{Y}^{\mathscr{M}}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$ can be seen as part of the dual space to $\mathbf{E}_c(\mathbb{R}^d; \mathbb{R}^N)$, which is defined like $\mathbf{E}(\mathbb{R}^d; \mathbb{R}^N)$, but requiring in addition that for $f \in \mathbf{E}_c(\mathbb{R}^d; \mathbb{R}^N) \subset \mathrm{C}(\mathbb{R}^d \times \mathbb{R}^N)$ there exists a compact set $K \subset \mathbb{R}^d$ with $\operatorname{supp} f(\bullet, A) \subset K$ for all $A \in \mathbb{R}^N$. Likewise, the **(local) weak* convergence** $\nu_j \overset{*}{\rightharpoonup} \nu$ in $\mathbf{Y}^{\mathscr{M}}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$ is defined with respect to $\mathbf{E}_c(\mathbb{R}^d; \mathbb{R}^N)$, that is, $\nu_j \overset{*}{\rightharpoonup} \nu$ in $\mathbf{Y}^{\mathscr{M}}_{\mathrm{loc}}$ if $\langle\!\langle f, \nu_j\rangle\!\rangle \to \langle\!\langle f, \nu\rangle\!\rangle$ for all $f \in \mathbf{E}_c(\mathbb{R}^d; \mathbb{R}^N)$. All of the results from the preceding sections also hold *mutatis mutandis* in $\mathbf{Y}^{\mathscr{M}}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$. In particular, we have the following compactness result.

Corollary 12.21. *Let* $(\nu_j) \subset \mathbf{Y}^{\mathscr{M}}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$ *satisfy*

$$\sup_{j \in \mathbb{N}} \langle\!\langle \varphi \otimes |\bullet|, \nu_j \rangle\!\rangle < \infty \qquad \text{for all } \varphi \in C_c(\mathbb{R}^d).$$

Then, there exists a subsequence (not explicitly labeled) such that $\nu_j \overset{*}{\rightharpoonup} \nu$ *for a Young measure* $\nu \in \mathbf{Y}^{\mathscr{M}}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$.

Finally, define $\mathbf{BVY}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^{m \times d})$ as the space of all those local (generalized) Young measures that are generated by derivatives of sequences in $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^m)$.

We are now in a position to state and prove the first localization principle.

Proposition 12.22. *Let* $\nu \in \mathbf{BVY}(\Omega; \mathbb{R}^{m \times d})$. *Then, for* $\mathscr{L}^d$*-almost every* $x_0 \in \Omega$ *there exists a* ***regular tangent Young measure*** $\sigma \in \mathbf{BVY}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^{m \times d})$, *that is,*

$$[\sigma] \in \mathrm{Tan}([\nu], x_0), \qquad\qquad \sigma_y = \nu_{x_0} \quad \mathscr{L}^d\text{-a.e.,} \tag{12.13}$$

$$\lambda_\sigma = \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x_0)\, \mathscr{L}^d \in \mathrm{Tan}(\lambda_\nu, x_0), \qquad \sigma_y^\infty = \nu_{x_0}^\infty \quad \lambda_\sigma\text{-a.e.} \tag{12.14}$$

Proof. Step 1. Let the family $\{\varphi_k \otimes h_k\}_{k \in \mathbb{N}} \subset \mathbf{E}_c(\mathbb{R}^d; \mathbb{R}^{m \times d})$ determine the (local) weak* Young measure convergence; the construction of this family is analogous to the proof of Lemma 12.10.

Choose $x_0 \in \Omega$ with the following properties:

(a) There exists a sequence $r_n \downarrow 0$ such that with $P_0 := \frac{\mathrm{d}[\nu]}{\mathrm{d}\mathscr{L}^d}(x_0)$ it holds that

$$\gamma_n := r_n^{-d} T_\#^{(x_0, r_n)}[\nu] \overset{*}{\rightharpoonup} P_0\, \mathscr{L}^d \in \mathrm{Tan}([\nu], x_0),$$

where $T_\#^{(x_0,r_n)}[\nu]$ denotes the push-forward of the measure $[\nu]$ under the rescaling map $T^{(x_0,r_n)}(x) := (x - x_0)/r_n$ as in Section 10.2;

(b) $\lim_{r \downarrow 0} r^{-d} \lambda_\nu^s(B(x_0, r)) = 0$, where λ_ν^s is the Lebesgue-singular part of λ_ν, and

$$\frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x_0)\, \mathscr{L}^d \in \mathrm{Tan}(\lambda_\nu, x_0);$$

(c) the point x_0 is an $\mathscr{L}^d$-Lebesgue point for the functions

$$x \mapsto \langle h_k, \nu_x \rangle + \langle h_k^\infty, \nu_x^\infty \rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x), \qquad k \in \mathbb{N}.$$

Indeed, at $\mathscr{L}^d$-almost every $x_0 \in \Omega$ condition (a) follows from Theorem A.20 and Lemma 10.4; condition (b) is a consequence of the Besicovitch Differentiation Theorem A.23 and Proposition 10.5; finally, condition (c) is again implied by Theorem A.20.

Step 2. By virtue of Proposition 12.18 take a sequence $(u_j) \subset (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(\Omega;\mathbb{R}^m)$ with $\nabla u_j \overset{\mathbf{Y}}{\to} \nu$ and let $\tilde{u}_j \in \mathrm{BV}(\mathbb{R}^d;\mathbb{R}^m)$ be the extension of u_j by zero to all of $\mathbb{R}^d$. Define

$$v_j^{(n)}(y) := \frac{\tilde{u}_j(x_0 + r_n y)}{r_n}, \qquad y \in \mathbb{R}^d.$$

We have the following transformation rules for $T_\#^{(x_0,r_n)}\mu$, where $\mu \in \mathscr{M}(\Omega;\mathbb{R}^N)$:

$$\frac{\mathrm{d}T_\#^{(x_0,r_n)}\mu}{\mathrm{d}\mathscr{L}^d} = r_n^d \frac{\mathrm{d}\mu}{\mathrm{d}\mathscr{L}^d}(x_0 + r_n \bullet), \qquad \frac{\mathrm{d}T_\#^{(x_0,r_n)}\mu}{\mathrm{d}|T_\#^{(x_0,r_n)}\mu|} = \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x_0 + r_n \bullet).$$

We may then compute

$$\begin{aligned} Dv_j^{(n)} &= r_n^{-d} T_\#^{(x_0,r_n)} D\tilde{u}_j \\ &= \nabla u_j(x_0 + r_n \bullet)\, \mathscr{L}^d + r_n^{-1}\big(u_j(x_0 + r_n \bullet)|_{\partial\Omega_n} \otimes n_{\Omega_n}\big)\, \mathscr{H}^{d-1} \llcorner \partial\Omega_n. \end{aligned}$$

Here, $\Omega_n := r_n^{-1}(\Omega - x_0)$, $n_{\Omega_n} \colon \partial\Omega_n \to \mathbb{S}^{d-1}$ is the unit inner normal to $\partial\Omega_n$, and $u_j(x_0 + r_n \bullet)|_{\partial\Omega_n}$ is the (inner) trace of $y \mapsto u_j(x_0 + r_n y)$ on $\partial\Omega_n$. Then, also using the BV-Poincaré inequality (10.7) and the boundedness of the BV-trace operator, see Section 10.3,

$$\begin{aligned} \|v_j^{(n)}\|_{\mathrm{BV}(\mathbb{R}^d;\mathbb{R}^m)} &\le C(n)|Dv_j^{(n)}|(\mathbb{R}^d) \\ &= C(n)|D\tilde{u}_j|(\mathbb{R}^d) \\ &\le C(n)\|u_j\|_{\mathrm{BV}(\Omega;\mathbb{R}^m)}, \end{aligned} \tag{12.15}$$

where we have absorbed all n-dependent constants (including r_n^{-d}) into $C(n) > 0$. If we hold n fixed, this expression is j-uniformly bounded. Consequently, we may select an n-dependent subsequence of the j's (not explicitly labeled) such that $Dv_j^{(n)} \overset{\mathbf{Y}}{\to} \sigma^{(n)}$ for some $\sigma^{(n)} \in \mathbf{BVY}(\mathbb{R}^d;\mathbb{R}^{m\times d})$.

Step 3. Fix $\varphi_k \otimes h_k \in \mathbf{E}_c(\mathbb{R}^d;\mathbb{R}^{m\times d})$ from the Young measure-determining collection above and choose $n \in \mathbb{N}$ so large that $\operatorname{supp}\varphi_k \Subset \Omega_n$. Then, the boundary measure in $Dv_j^{(n)}$ (the part supported on $\partial\Omega_n$) can be neglected in the following calculation:

$$\begin{aligned} \big\langle\!\big\langle \varphi_k \otimes h_k, \sigma^{(n)} \big\rangle\!\big\rangle &= \lim_{j\to\infty} \int \varphi_k(y) h_k\big(\nabla v_j^{(n)}(y)\big)\,\mathrm{d}y \\ &= \lim_{j\to\infty} \int \varphi_k(y) h_k\big(\nabla u_j(x_0 + r_n y)\big)\,\mathrm{d}y \\ &= \lim_{j\to\infty} \frac{1}{r_n^d} \int \varphi_k\Big(\frac{x - x_0}{r_n}\Big) h_k(\nabla u_j(x))\,\mathrm{d}x \\ &= \frac{1}{r_n^d} \Big\langle\!\Big\langle \varphi_k\Big(\frac{\bullet - x_0}{r_n}\Big) \otimes h_k, \nu \Big\rangle\!\Big\rangle \end{aligned}$$

$$= \frac{1}{r_n^d} \int \varphi_k\Big(\frac{x-x_0}{r_n}\Big)\Big[\big\langle h_k, \nu_x\big\rangle + \big\langle h_k^\infty, \nu_x^\infty\big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x)\Big] \,\mathrm{d}x$$
$$+ \frac{1}{r_n^d} \int \varphi_k\Big(\frac{x-x_0}{r_n}\Big)\big\langle h_k^\infty, \nu_x^\infty\big\rangle \,\mathrm{d}\lambda_\nu^s(x).$$

We call the last two integrals the regular and the singular parts respectively.

For the regular part we have

$$\frac{1}{r_n^d} \int \varphi_k\Big(\frac{x-x_0}{r_n}\Big)\Big[\big\langle h_k, \nu_x\big\rangle + \big\langle h_k^\infty, \nu_x^\infty\big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x)\Big] \,\mathrm{d}x$$
$$= \int \varphi_k(y)\Big[\big\langle h_k, \nu_{x_0+r_n y}\big\rangle + \big\langle h^\infty, \nu_{x_0+r_n y}^\infty\big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x_0+r_n y)\Big] \,\mathrm{d}y$$
$$\to \int \varphi_k(y)\Big[\big\langle h_k, \nu_{x_0}\big\rangle + \big\langle h_k^\infty, \nu_{x_0}^\infty\big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x_0)\Big] \,\mathrm{d}y$$

as $n \to \infty$ ($r_n \downarrow 0$) by the Lebesgue point property (c) above.

Turning to the singular part, let $N \in \mathbb{N}$ be so large that $\operatorname{supp} \varphi_k \subset B(0, N)$. By assumption (b) on x_0,

$$\left|\frac{1}{r_n^d} \int \varphi_k\Big(\frac{x-x_0}{r_n}\Big)\big\langle h_k^\infty, \nu_x^\infty\big\rangle \,\mathrm{d}\lambda_\nu^s(x)\right| \le M\|\varphi_k\|_\infty \cdot \frac{\lambda_\nu^s(B(x_0, Nr_n))}{r_n^d} \to 0$$

as $n \to \infty$, where $M := \sup\{\, |h^\infty(A)| \;:\; A \in \partial\mathbb{B}^{m\times d} \,\}$.

Step 4. We may assume that the integrands $\varphi \otimes |\bullet|$ for a dense set of $\varphi \in \mathrm{C}_c(\mathbb{R}^d)$ (in the $\|\bullet\|_\infty$-norm) are contained in the collection $\{\varphi_k \otimes h_k\}_k$. Then, the above arguments imply

$$\sup_{n\in\mathbb{N}} \big|\langle\!\langle \varphi \otimes |\bullet|, \sigma^{(n)} \rangle\!\rangle\big| < \infty \qquad \text{for all } \varphi \in \mathrm{C}_c(\mathbb{R}^d).$$

Hence, the compactness result from Corollary 12.21 applies and we may select a subsequence (not relabeled) with

$$\sigma^{(n)} \overset{*}{\rightharpoonup} \sigma \in \mathbf{BVY}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^{m\times d}).$$

Here we also used that $\mathbf{BVY}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^{m\times d})$ is sequentially weakly* closed, see Problem 12.9. Since $[\sigma^{(n)}] = \gamma_n$ plus a jump part that moves out to infinity in the limit, we furthermore get that $[\sigma] \in \operatorname{Tan}([\nu], x_0)$. This implies the first assertion in (12.13).

It also follows from the preceding arguments that

$$\langle\!\langle \varphi_k \otimes h, \sigma \rangle\!\rangle = \int \varphi_k(y)\Big[\big\langle h_k, \nu_{x_0}\big\rangle + \big\langle h_k^\infty, \nu_{x_0}^\infty\big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x_0)\Big] \,\mathrm{d}y$$

for all $\varphi_k \otimes h_k$ from the family exhibited at the beginning of the proof. Since the $\varphi_k \otimes h_k$ are dense in $\mathbf{E}_c(\mathbb{R}^d; \mathbb{R}^{m\times d})$ (in the sense that they determine Young measures), we have $\sigma_y = \nu_{x_0}$ and $\sigma_y^\infty = \nu_{x_0}^\infty$ for $\mathscr{L}^d$-almost every $y \in \mathbb{R}^d$, that is, the second assertion of (12.13) and the second assertion of (12.14) hold. Finally, the first assertion of (12.14) follows since

$$\lambda_\sigma = \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x_0)\,\mathscr{L}^d \in \mathrm{Tan}(\lambda_\nu, x_0)$$

by (b). □

Next, we investigate localization at singular points (we remark that this result is not actually needed anywhere in the sequel).

Proposition 12.23. *Let* $\nu \in \mathbf{BVY}(\Omega; \mathbb{R}^{m\times d})$*. Then, for* λ_ν^s*-almost every* $x_0 \in \Omega$ *there exists a* ***singular tangent Young measure*** $\sigma \in \mathbf{BVY}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^{m\times d})$*, that is,*

$$[\sigma] \in \mathrm{Tan}([\nu], x_0), \qquad \sigma_y = \delta_0 \quad \mathscr{L}^d\text{-a.e.,} \tag{12.16}$$

$$\lambda_\sigma \in \mathrm{Tan}(\lambda_\nu^s, x_0) \setminus \{0\}, \qquad \sigma_y^\infty = \nu_{x_0}^\infty \quad \lambda_\sigma\text{-a.e.} \tag{12.17}$$

Proof. Step 1. Let $\{g_k\} \subset \mathrm{C}(\mathbb{R}^{m\times d})$ be a countable set of positively 1-homogeneous functions whose restrictions to $\partial\mathbb{B}^{m\times d}$ are dense in $\mathrm{C}(\partial\mathbb{B}^{m\times d})$. Let $x_0 \in \Omega$ be such that the following properties hold:

(a) There exist sequences $r_n \downarrow 0$, $c_n > 0$ and $\lambda_0 \in \mathrm{Tan}(\lambda_\nu^s, x_0) \setminus \{0\}$ such that

$$c_n T_\#^{(x_0,r_n)} \lambda_\nu^s \overset{*}{\rightharpoonup} \lambda_0; \tag{12.18}$$

(b) it holds that

$$\lim_{r\downarrow 0} \frac{1}{\lambda_\nu^s(B(x_0,r))} \int_{B(x_0,r)} 1 + \big\langle |\bullet|, \nu_x \big\rangle + \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x)\;\mathrm{d}x = 0; \tag{12.19}$$

(c) the point x_0 is a λ_ν^s-Lebesgue point for the functions

$$x \mapsto [\nu_x^\infty] \qquad \text{and} \qquad x \mapsto \big\langle g_k, \nu_x^\infty \big\rangle, \quad k \in \mathbb{N}.$$

By Proposition 10.5 condition (a) holds at λ_ν^s-almost every $x_0 \in \Omega$, condition (b) follows from the Besicovitch Differentiation Theorem A.23, and condition (c) is a direct consequence of Theorem A.20. Further, by (10.3), the constants c_n in (12.18) can be chosen to be

$$c_n = c\big[\lambda_\nu^s(\overline{B(x_0, Rr_n)})\big]^{-1}$$

for any (henceforth fixed) $R > 0$ and corresponding $c > 0$ with the property that $\lambda_0(B(0,R)) > 0$. We also choose R such that $\lambda_\nu^s(\partial B(x_0, Rr_n)) = 0$ for all n, which is always possible by the finiteness of the measure λ_ν^s (see Problem 10.1). Then, we get from (12.18) that there exist constants $\beta_N > 0$, $N = 1, 2, \ldots$, such that

$$\limsup_{n\to\infty} c_n \lambda_\nu^s(B(x_0, Nr_n)) = \limsup_{n\to\infty} c \cdot \frac{\lambda_\nu^s(B(x_0, Nr_n))}{\lambda_\nu^s(B(x_0, Rr_n))} \le \beta_N. \tag{12.20}$$

Thus, with (12.19), we compute

$$
\begin{aligned}
&\limsup_{n\to\infty}\, c_n\langle\!\langle \mathbb{1}_{B(x_0,Nr_n)} \otimes |\bullet|, \nu\rangle\!\rangle \\
&\quad= \limsup_{n\to\infty}\Bigg[\frac{c}{\lambda_\nu^s(B(x_0,Rr_n))}\int_{B(x_0,Nr_n)}\langle|\bullet|,\nu_x\rangle + \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x)\,\mathrm{d}x \\
&\qquad\qquad + c\cdot\frac{\lambda_\nu^s(B(x_0,Nr_n))}{\lambda_\nu^s(B(x_0,Rr_n))}\Bigg] \\
&\quad\le 0+\beta_N.
\end{aligned}
$$

Similarly,

$$
\limsup_{n\to\infty}\big(c_n T_\#^{(x_0,r_n)}|[\nu]|\big)(B(0,N)) \le \beta_N \qquad \text{for all } N\in\mathbb{N}.
$$

After selecting a subsequence (not explicitly labeled) of the r_n, we may thus suppose that

$$
c_n T_\#^{(x_0,r_n)}[\nu] \overset{*}{\rightharpoonup} \tau \in \mathrm{Tan}([\nu],x_0). \tag{12.21}
$$

We note that it is possible that $\tau = 0$ (clearly, $\tau \ne 0$ for $[\nu]$-almost every $x_0 \in \operatorname{supp}[\nu]$, but not necessarily for λ_ν^s-almost every $x_0 \in \Omega$).

Step 2. Let $(u_j) \subset (\mathrm{W}^{1,1}\cap\mathrm{C}^\infty)(\Omega;\mathbb{R}^m)$ be such that $\nabla u_j \overset{\mathbf{Y}}{\to} \nu$, see Proposition 12.18. Denote by $\tilde u_j \in \mathrm{BV}(\mathbb{R}^d;\mathbb{R}^m)$ the extension of u_j by zero and define

$$
v_j^{(n)}(y) := r_n^{d-1}c_n\tilde u_j(x_0+r_ny), \qquad y\in\mathbb{R}^d.
$$

We have

$$
\begin{aligned}
Dv_j^{(n)} &= c_n T_\#^{(x_0,r_n)} D\tilde u_j \\
&= r_n^d c_n\nabla u_j(x_0+r_n\bullet)\,\mathscr{L}^d + r_n^{d-1}c_n\big(u_j(x_0+r_n\bullet)|_{\partial\Omega_n}\otimes n_{\Omega_n}\big)\,\mathscr{H}^{d-1}\llcorner\Omega_n,
\end{aligned}
$$

where $\Omega_n := r_n^{-1}(\Omega - x_0)$. Similarly to (12.15), we get

$$
\|v_j^{(n)}\|_{\mathrm{BV}(\mathbb{R}^d;\mathbb{R}^m)} \le C(n)\|u_j\|_{\mathrm{BV}(\Omega;\mathbb{R}^m)}.
$$

Hence, holding n fixed, we may assume, up to taking an n-dependent subsequence of the j's, that $Dv_j^{(n)} \overset{\mathbf{Y}}{\to} \sigma^{(n)}$ for some $\sigma^{(n)} \in \mathbf{BVY}(\mathbb{R}^d;\mathbb{R}^{m\times d})$ as $j\to\infty$.

Step 3. Fix a positively 1-homogeneous $g\in\mathrm{C}(\mathbb{R}^{m\times d})$ and $\varphi\in\mathrm{C}_c(\mathbb{R}^d)$. For all n so large that $\operatorname{supp}\varphi \Subset \Omega_n$,

$$
\begin{aligned}
\langle\!\langle \varphi\otimes g, \sigma^{(n)}\rangle\!\rangle &= \lim_{j\to\infty}\int \varphi(y)g\big(\nabla v_j^{(n)}(y)\big)\,\mathrm{d}y \\
&= \lim_{j\to\infty} r_n^d c_n\int \varphi(y)g\big(\nabla u_j(x_0+r_ny)\big)\,\mathrm{d}y
\end{aligned}
$$

$$
\begin{aligned}
&= \lim_{j\to\infty} c_n \int \varphi\Big(\frac{x-x_0}{r_n}\Big) g(\nabla u_j(x))\,\mathrm{d}y \\
&= c_n \,\Big\langle\!\!\Big\langle\, \varphi\Big(\frac{\bullet - x_0}{r_n}\Big)\otimes g, \nu \,\Big\rangle\!\!\Big\rangle \\
&= c_n \int \varphi\Big(\frac{x-x_0}{r_n}\Big)\Big[\big\langle g,\nu_x\big\rangle + \big\langle g,\nu_x^\infty\big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x)\Big]\,\mathrm{d}x \\
&\qquad + c_n \int \varphi\Big(\frac{x-x_0}{r_n}\Big)\big\langle g,\nu_x^\infty\big\rangle\,\mathrm{d}\lambda_\nu^s(x).
\end{aligned} \tag{12.22}
$$

As before, we call the last two integrals the regular and singular parts, respectively.

The regular part is estimated as follows: Set $M := \sup\{\,|g(A)| \,:\, A\in\partial\mathbb{B}^{m\times d}\,\}$ and pick $N\in\mathbb{N}$ large enough such that $\operatorname{supp}\varphi\subset B(0,N)$. Via (12.20) we assume n to be so large that

$$
c_n\lambda_\nu^s(B(x_0,Nr_n)) \le \beta_N + 1.
$$

Then,

$$
\begin{aligned}
&\left| c_n \int \varphi\Big(\frac{x-x_0}{r_n}\Big)\Big[\big\langle g,\nu_x\big\rangle + \big\langle g,\nu_x^\infty\big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x)\Big]\,\mathrm{d}x \right| \\
&\quad \le c_n M\|\varphi\|_\infty \int_{B(x_0,Nr_n)} \big\langle|\bullet|,\nu_x\big\rangle + \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x)\,\mathrm{d}x \\
&\quad \le \frac{M\|\varphi\|_\infty(\beta_N+1)}{\lambda_\nu^s(B(x_0,Nr_n))} \int_{B(x_0,Nr_n)} \big\langle|\bullet|,\nu_x\big\rangle + \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x)\,\mathrm{d}x \\
&\quad \to 0 \qquad \text{as } n\to\infty.
\end{aligned} \tag{12.23}
$$

Here, the convergence in the last line follows from (12.19). Plugging this back into (12.22), we get

$$
\limsup_{n\to\infty} \big\langle\!\!\big\langle \varphi\otimes g, \sigma^{(n)} \big\rangle\!\!\big\rangle = \limsup_{n\to\infty}\, c_n \int \varphi\Big(\frac{x-x_0}{r_n}\Big)\big\langle g,\nu_x^\infty\big\rangle\,\mathrm{d}\lambda_\nu^s(x). \tag{12.24}
$$

For $g = |\bullet|$, also using (12.18), this gives

$$
\begin{aligned}
\limsup_{n\to\infty} \big\langle\!\!\big\langle \varphi\otimes|\bullet|, \sigma^{(n)} \big\rangle\!\!\big\rangle &= \limsup_{n\to\infty}\, c_n \int \varphi\Big(\frac{x-x_0}{r_n}\Big)\,\mathrm{d}\lambda_\nu^s(x) \\
&= \limsup_{n\to\infty} \int \varphi\,\mathrm{d}\big(c_n T_\#^{(x_0,r_n)}\lambda_\nu^s\big) \\
&= \int \varphi\,\mathrm{d}\lambda_0.
\end{aligned}
$$

In particular,

$$\limsup_{n\to\infty} \left\langle\!\left\langle \varphi \otimes |\bullet|, \sigma^{(n)} \right\rangle\!\right\rangle \le \|\varphi\|_\infty \, \lambda_0(\operatorname{supp}\varphi).$$

Thus, the Young measure compactness criterion, see Corollary 12.21, implies that there exists a subsequence of the r_n's (not explicitly labeled) with

$$\sigma^{(n)} \overset{*}{\rightharpoonup} \sigma \in \mathbf{BVY}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^{m\times d}),$$

where we also used that $\mathbf{BVY}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^{m\times d})$ is sequentially weakly* closed, see Problem 12.9. Furthermore, (12.24) implies

$$\left\langle\!\left\langle \varphi \otimes g, \sigma \right\rangle\!\right\rangle = \lim_{n\to\infty} c_n \int \varphi\Big(\frac{x - x_0}{r_n}\Big) \langle g, \nu_x^\infty \rangle \, \mathrm{d}\lambda_\nu^s(x). \tag{12.25}$$

Step 4. We have

$$[\sigma^{(n)}] = c_n T_\#^{(x_0, r_n)}[\nu] + \mu_n,$$

where $\mu_n \in \mathscr{M}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^{m\times d})$ is a measure carried by the set $\partial\Omega_n$, whereby $\mu_n \overset{*}{\rightharpoonup} 0$. From (12.21), we thus infer that $[\sigma^{(n)}] \overset{*}{\rightharpoonup} \tau$ as $n \to \infty$ and

$$[\sigma] = \tau \in \operatorname{Tan}([\nu], x_0),$$

which is the first statement in (12.16).

We now take functions $\varphi \in \mathrm{C}_c(\mathbb{R}^d)$, $\chi \in \mathrm{C}_c(\mathbb{R}^{m\times d})$ and in a similar fashion to (12.22) derive that

$$\begin{aligned}
\left\langle\!\left\langle \varphi \otimes |\bullet|\chi(\bullet), \sigma^{(n)} \right\rangle\!\right\rangle &= \lim_{j\to\infty} c_n \int \varphi\Big(\frac{x - x_0}{r_n}\Big) |\nabla u_j(x)| \chi\big(r_n^d c_n \nabla u_j(x)\big) \, \mathrm{d}x \\
&= c_n \left\langle\!\!\left\langle \varphi\Big(\frac{\bullet - x_0}{r_n}\Big) \otimes |\bullet|\chi(r_n^d c_n \bullet), \nu \right\rangle\!\!\right\rangle \\
&\to 0 \qquad \text{as } n \to \infty.
\end{aligned}$$

Here, the last convergence follows analogously to (12.23) since χ has compact support in $\mathbb{R}^{m\times d}$. Thus,

$$\left\langle\!\left\langle \varphi \otimes |\bullet|\chi(\bullet), \sigma \right\rangle\!\right\rangle = 0$$

for all φ, χ as above. From this, varying φ and χ, we conclude that $\sigma_y = \delta_0$ for $\mathscr{L}^d$-almost every $y \in \mathbb{R}^d$, which is the second statement of (12.16).

Moving on to the first assertion of (12.17), use $g := |\bullet|$ in (12.25) and the previously shown fact $\sigma_y = \delta_0$ almost everywhere to derive

$$\begin{aligned}\int \varphi \, \mathrm{d}\lambda_\sigma &= \langle\!\langle \varphi \otimes |\bullet|, \sigma \rangle\!\rangle \\ &= \lim_{n\to\infty} c_n \int \varphi\Big(\frac{x - x_0}{r_n}\Big) \, \mathrm{d}\lambda_\nu^s(x) \\ &= \lim_{n\to\infty} \int \varphi \, \mathrm{d}\big(c_n T_\#^{(x_0,r_n)} \lambda_\nu^s\big) \\ &= \int \varphi \, \mathrm{d}\lambda_0\end{aligned}$$

for any $\varphi \in \mathrm{C}_c(\mathbb{R}^d)$. Here, the last equality is a consequence of (12.18). Thus, indeed $\lambda_\sigma = \lambda_0 \in \mathrm{Tan}(\lambda_\nu^s, x_0)$.

Finally, to establish the second statement of (12.17), we first prove the following assertion: Let $U \subset \mathbb{R}^d$ be open and bounded with $(\mathscr{L}^d + \lambda_\sigma)(\partial U) = 0$ and let $g \in \mathrm{C}(\mathbb{R}^{m\times d})$ be positively 1-homogeneous. Then,

$$\langle\!\langle \mathbb{1}_U \otimes g, \sigma \rangle\!\rangle = \big\langle g, \nu_{x_0}^\infty \big\rangle \lambda_\sigma(U). \tag{12.26}$$

Once this is established we may vary U and g to see that $\sigma_y^\infty = \nu_{x_0}^\infty$ for λ_σ-almost every $y \in \mathbb{R}^d$, which is the second claim of (12.17).

To prove (12.26), we assume $\lambda_\sigma(U) > 0$ since the case $\lambda_\sigma(U) = 0$ is trivial. Then use $\varphi = \mathbb{1}_U$ in (12.25) via Proposition 12.11 (ii) to get

$$\int_U \big\langle g, \sigma_y^\infty \big\rangle \, \mathrm{d}\lambda_\sigma(y) = \langle\!\langle \mathbb{1}_U \otimes g, \sigma \rangle\!\rangle = \lim_{n\to\infty} c_n \int_{x_0 + r_n U} \big\langle g, \nu_x^\infty \big\rangle \, \mathrm{d}\lambda_\nu^s(x).$$

Since $\lambda_\sigma \in \mathrm{Tan}(\lambda_\nu^s, x_0)$ and $\lambda_\sigma(U) > 0$, (10.3) implies

$$c_n = \frac{\tilde{c}(U)}{\lambda_\nu^s(x_0 + r_n U)}$$

for some $\tilde{c}(U) > 0$. Consequently,

$$\begin{aligned}\lim_{n\to\infty} c_n \int_{x_0+r_nU} \big\langle g, \nu_x^\infty \big\rangle \, \mathrm{d}\lambda_\nu^s(x) &= \tilde{c}(U) \cdot \lim_{n\to\infty} \fint_{x_0+r_nU} \big\langle g, \nu_x^\infty \big\rangle \, \mathrm{d}\lambda_\nu^s(x) \\ &= \tilde{c}(U) \big\langle g, \nu_{x_0}^\infty \big\rangle\end{aligned}$$

by the Lebesgue point property (c) of x_0 (first for $g = g_k$ and then by density for the general case). Thus, we have

$$\langle\!\langle \mathbb{1}_U \otimes g, \sigma \rangle\!\rangle = \int_U \big\langle g, \sigma_y^\infty \big\rangle \, \mathrm{d}\lambda_\sigma(y) = \tilde{c}(U) \big\langle g, \nu_{x_0}^\infty \big\rangle.$$

With $g = |\bullet|$ we see that $\tilde{c}(U) = \lambda_\sigma(U)$ and (12.26) follows. □

12.7 Lower Semicontinuity

We now show how the theory of generalized Young measures can be used to investigate the weak* lower semicontinuity properties of functionals with linear-growth integrands. First, we extend Theorem 11.2 as follows:

Proposition 12.24. *Let $f \in \mathbf{R}(\Omega; \mathbb{R}^{m\times d})$. Then, the functional*

$$\widetilde{\mathscr{F}}[u] := \int_\Omega f(x, \nabla u(x))\,\mathrm{d}x + \int_\Omega f^\infty\Bigl(x, \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\Bigr)\,\mathrm{d}|D^s u|(x) + \int_{\partial\Omega} f^\infty\bigl(x, u(x)\otimes n_\Omega(x)\bigr)\,\mathrm{d}\mathscr{H}^{d-1}(x), \qquad u \in \mathrm{BV}(\Omega; \mathbb{R}^m),$$

where $n_\Omega\colon \partial\Omega \to \mathbb{S}^{d-1}$ is the unit inner normal to $\partial\Omega$ ("inner" with respect to Ω), is the area-strictly continuous extension to the space $\mathrm{BV}(\Omega; \mathbb{R}^m)$ *of the functional*

$$\mathscr{F}[u] := \int_\Omega f(x, \nabla u(x))\,\mathrm{d}x + \int_{\partial\Omega} f^\infty\bigl(x, u(x)\otimes n_\Omega(x)\bigr)\,\mathrm{d}\mathscr{H}^{d-1}(x),$$

where $u \in \mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$.

Proof. Extend u by zero to a larger Lipschitz domain $\Omega' \Supset \Omega$. Then, $\widetilde{\mathscr{F}}$ as defined above is the counterpart of the usual $\mathscr{F}$, but on Ω'. The space $\mathrm{W}^{1,1}(\Omega'; \mathbb{R}^m)$ is area-strictly dense in $\mathrm{BV}(\Omega'; \mathbb{R}^m)$ by Lemma 11.1. The result then follows by Proposition 12.4 together with Proposition 12.11 on extended representation. □

The following is the main lower semicontinuity result of this section. In comparison to the Ambrosio–Dal Maso–Fonseca–Müller Theorem 11.7 we allow the integrand to be x-dependent, but we need to require that the recession function exists in the strong sense (we note that already [124] treated x-dependent integrands, albeit under further technical assumptions).

Theorem 12.25. *Let $f \in \mathbf{R}(\Omega; \mathbb{R}^{m\times d})$ be quasiconvex in the second argument, that is, we suppose that $f\colon \overline{\Omega} \times \mathbb{R}^{m\times d} \to [0, \infty)$ is such that the following assumptions hold:*

- *(i) f is a Carathéodory integrand;*
- *(ii) $|f(x, A)| \le M(1 + |A|)$ for some $M > 0$ and all $x \in \overline{\Omega}$, $A \in \mathbb{R}^{m\times d}$;*
- *(iii) the strong recession function f^∞ exists in the sense of (11.8) and is (jointly) continuous on $\overline{\Omega} \times \mathbb{R}^{m\times d}$;*
- *(iv) $f(x, \cdot)$ is quasiconvex for all $x \in \overline{\Omega}$.*

Then, the functional

$$\mathscr{F}[u] := \int_\Omega f(x, \nabla u(x))\,\mathrm{d}x + \int_\Omega f^\infty\Bigl(x, \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\Bigr)\,\mathrm{d}|D^s u|(x) + \int_{\partial\Omega} f^\infty\bigl(x, u(x)\otimes n_\Omega(x)\bigr)\,\mathrm{d}\mathscr{H}^{d-1}(x), \qquad u \in \mathrm{BV}(\Omega; \mathbb{R}^m),$$

is lower semicontinuous with respect to the weak convergence in* $\mathrm{BV}(\Omega; \mathbb{R}^m)$.

Remark 12.26. The boundary term in $\mathscr{F}$ may be omitted, see Problem 12.10.

The proof proceeds by showing Jensen-type inequalities for BV-Young measures at regular and singular points.

Proposition 12.27. *Let $\nu \in \mathbf{BVY}(\Omega; \mathbb{R}^{m\times d})$. Then, for all quasiconvex and continuous $h \colon \mathbb{R}^{m\times d} \to [0, \infty)$ with linear growth, it holds that*

(i) $h\left([\nu_x] + [\nu_x^\infty]\frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x)\right) \le \langle h, \nu_x\rangle + \langle h^\#, \nu_x^\infty\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x)$ *for $\mathscr{L}^d$-a.e. $x \in \Omega$;*

(ii) $h^\#([\nu_x^\infty]) \le \langle h^\#, \nu_x^\infty\rangle$ *for λ_ν^s-a.e. $x \in \Omega$, where λ_ν^s is the singular part of λ_ν with respect to $\mathscr{L}^d$.*

In the proof of the proposition we will need an approximation lemma, which we state in slightly more generality than is strictly necessary here (we only need the x-independent case) because of the independent interest of the result.

Lemma 12.28. *For every function $f \in C(\overline{\Omega} \times \mathbb{R}^N)$ with linear growth there is a decreasing sequence $(f_n) \subset \mathbf{E}(\Omega; \mathbb{R}^N)$ with*

$$\inf_{n\in\mathbb{N}} f_n = \lim_{n\to\infty} f_n = f, \qquad \inf_{n\in\mathbb{N}} f_n^\infty = \lim_{n\to\infty} f_n^\infty = f^\# \qquad \textit{(pointwise).}$$

Furthermore, if $M \ge 0$ is such that $|f(x, A)| \le M(1 + |A|)$, then the f_n can also be chosen to satisfy $|f_n(x, A)| \le M(1 + |A|)$.

Proof. We denote by $(Sf)^{\mathrm{usc}} \colon \overline{\Omega \times \mathbb{B}^N} \to \mathbb{R}$ the upper semicontinuous extension of $Sf \colon \overline{\Omega} \times \mathbb{B}^N \to \mathbb{R}$ to $\overline{\Omega \times \mathbb{B}^N}$, that is,

$$(Sf)^{\mathrm{usc}}(x, \hat{A}) := \limsup_{\substack{x_n\to x\\ \hat{A}_n\to\hat{A}}} Sf(x_n, \hat{A}_n), \qquad (x, \hat{A}) \in \overline{\Omega \times \mathbb{B}^N}.$$

Since Sf is continuous on $\overline{\Omega} \times \mathbb{B}^N$, we have $(Sf)^{\mathrm{usc}}|_{\overline{\Omega}\times\mathbb{B}^N} = Sf$. For sequences $x_n \to x$ in $\overline{\Omega}$, $(\hat{A}_n) \to \hat{A} \in \mathbb{B}^N$, $\hat{A} \in \partial\mathbb{B}^N$, and $t_n \to \infty$, it holds that

$$\limsup_{n\to\infty} \frac{f(x_n, t_n\hat{A}_n)}{t_n} = \limsup_{n\to\infty} (t_n^{-1} + |\hat{A}_n|) Sf\left(x_n, \frac{\hat{A}_n}{t_n^{-1} + |\hat{A}_n|}\right).$$

Hence, $f^\# = (Sf)^{\mathrm{usc}}|_{\overline{\Omega}\times\mathbb{S}^{N-1}}$. Indeed, "$\le$" holds since $(Sf)^{\mathrm{usc}}$ is upper semicontinuous and "$\ge$" follows by taking a sequence $(x_n, \hat{A}_n) \to (x, \hat{A})$ with $Sf(x_n, \hat{A}_n) \to (Sf)^{\mathrm{usc}}(x, \hat{A})$ and setting $t_n := (1 - |\hat{A}_n|)^{-1}$.

Let now $(g_k) \subset \mathrm{C}(\overline{\Omega \times \mathbb{B}^N})$ be a decreasing sequence with $g_k \downarrow (Sf)^{\mathrm{usc}}$ pointwise in $\overline{\Omega \times \mathbb{B}^N}$ and $|g_k| \le M$. Then set $f_k := S^{-1} g_k$ for which $f^k \downarrow f$, f_k^∞ exists in the sense of (11.8), and $f_k^\infty = g_k|_{\overline{\Omega}\times\mathbb{S}^{N-1}} \downarrow (Sf)^{\mathrm{usc}}|_{\overline{\Omega}\times\mathbb{S}^{N-1}} = f^\#$.

Finally, assuming $|f(x, A)| \le M(1 + |A|)$ for some $M \ge 0$, it holds that $|Sf| \le M$, whereby $|(Sf)^{\mathrm{usc}}| \le M$ as well. □

Proof of Proposition 12.27. *Ad (i).* Let $\sigma \in \mathbf{BVY}(B(0,1); \mathbb{R}^{m\times d})$ be a regular tangent Young measure to ν at a suitable $x_0 \in \Omega$ as in Proposition 12.22, which we consider to be restricted to $B(0,1)$ (see Problem 12.8). Then,

$$[\sigma] = F\,\mathscr{L}^d \llcorner B(0,1), \qquad \text{where} \qquad F = [\nu_{x_0}] + [\nu^\infty_{x_0}]\frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x_0).$$

Via Proposition 12.18 (ii) we obtain a sequence $(v_n) \subset (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(B(0,1); \mathbb{R}^m)$ with $Dv_n \xrightarrow{\mathbf{Y}} \sigma$ and $v_n = Fx$ on $\partial B(0,1)$.

For continuous functions $h\colon \mathbb{R}^{m\times d} \to [0,\infty)$ with linear growth the application of Lemma 12.28 yields a sequence $(\mathbb{1} \otimes h_k) \subset \mathbf{E}(B(0,1); \mathbb{R}^{m\times d})$ such that $h_k \downarrow h$, $h_k^\infty \downarrow h^\#$ pointwise, and all h_k have uniformly bounded linear growth constants. Then, using the quasiconvexity, we get

$$\begin{aligned} h(F) &\le \limsup_{n\to\infty} \fint_{B(0,1)} h(\nabla v_n)\,\mathrm{d}x \\ &\le \lim_{n\to\infty} \fint_{B(0,1)} h_k(\nabla v_n)\,\mathrm{d}x \\ &= \frac{1}{\omega_d}\big\langle\!\big\langle \mathbb{1} \otimes h_k, \sigma \big\rangle\!\big\rangle \\ &= \big\langle h_k, \nu_{x_0}\big\rangle + \big\langle h_k^\infty, \nu^\infty_{x_0}\big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x_0) \end{aligned}$$

for all $k \in \mathbb{N}$, where we used (12.13) and (12.14). We may then invoke the monotone convergence theorem and utilize $h_k \downarrow h$, $h_k^\infty \downarrow h^\#$ to conclude.

Ad (ii). From (12.11) we know

$$\operatorname{rank}\,[\nu^\infty_x] \le 1 \qquad \text{for } \lambda^s_\nu\text{-a.e. } x \in \Omega.$$

By the Kirchheim–Kristensen Theorem 10.13, $h^\#$ is convex at matrices of rank at most one, hence (ii) follows immediately from the Jensen-type inequality stated in (10.17). □

Proof of Theorem 12.25. Given a sequence $u_j \overset{*}{\rightharpoonup} u$ in $\mathrm{BV}(\Omega; \mathbb{R}^m)$, we consider all u_j, u to be extended to the whole space by zero, whereby

$$Du_j = Du_j \llcorner \Omega + (u_j \otimes n_\Omega)\,\mathscr{H}^{d-1} \llcorner \partial\Omega.$$

Select a subsequence $(u_{j(l)})$ such that

$$\liminf_{j\to\infty} \mathscr{F}[u_j] = \lim_{l\to\infty} \mathscr{F}[u_{j(l)}]$$

and

$$Du_{j(l)} \xrightarrow{\mathbf{Y}} \nu \in \mathbf{BVY}(\mathbb{R}^d; \mathbb{R}^{m\times d}),$$

the latter being possible by Theorem 12.16. It suffices to show lower semicontinuity along the sequence $(u_{j(l)})$, which in the following we just denote by (u_j).

We have

$$[\nu] = Du + (u \otimes n_\Omega)\, \mathscr{H}^{d-1} \llcorner \partial\Omega.$$

Let λ_ν^* be the singular part of λ_ν with respect to $|D^s u| + |u|\, \mathscr{H}^{d-1} \llcorner \partial\Omega$; in particular, λ_ν^* is carried by a $(|D^s u| + |u|\, \mathscr{H}^{d-1} \llcorner \partial\Omega)$-negligible set. We compute using (12.11) and the arguments preceding it that

$$[\nu_x] + [\nu_x^\infty] \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x) = \frac{\mathrm{d}[\nu]}{\mathrm{d}\mathscr{L}^d}(x) = \begin{cases} \nabla u(x) & \text{for } \mathscr{L}^d\text{-a.e. } x \in \Omega, \\ 0 & \text{for } \mathscr{L}^d\text{-a.e. } x \in \mathbb{R}^d \setminus \Omega; \end{cases}$$

$$\frac{[\nu_x^\infty]}{|[\nu_x^\infty]|} = \frac{\mathrm{d}[\nu]^s}{\mathrm{d}|[\nu]^s|}(x) = \begin{cases} \dfrac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x) & \text{for } |D^s u|\text{-a.e. } x \in \Omega, \\ \dfrac{u|_{\partial\Omega}(x)}{|u|_{\partial\Omega}(x)|} \otimes n_\Omega(x) & \text{for } |u|\, \mathscr{H}^{d-1}\text{-a.e. } x \in \partial\Omega; \end{cases}$$

$$[\nu_x^\infty] = 0 \qquad \text{for } \lambda_\nu^*\text{-a.e. } x \in \mathbb{R}^d;$$

$$|[\nu_x^\infty]|\, \lambda_\nu^s = |D^s u| + |u|\, \mathscr{H}^{d-1} \llcorner \partial\Omega;$$

$$[\nu_x] = 0 \qquad \text{for all } x \in \mathbb{R}^d \setminus \overline{\Omega};$$

$$\lambda_\nu \llcorner (\mathbb{R}^d \setminus \overline{\Omega}) = 0.$$

Also extend f to $\mathbb{R}^d \times \mathbb{R}^{m\times d}$ as follows: first extend f^∞ restricted to $\overline{\Omega} \times \partial\mathbb{B}^{m\times d}$ continuously to $\mathbb{R}^d \times \partial\mathbb{B}^{m\times d}$ and then set $f(x, A) := f^\infty(x, A)$ for $x \in \mathbb{R}^d \setminus \overline{\Omega}$ and $A \in \mathbb{R}^{m\times d}$. This extended f is still a Carathéodory integrand, f^∞ is jointly continuous and $f(x, 0) = 0$ for all $x \in \mathbb{R}^d \setminus \overline{\Omega}$.

Then, Proposition 12.11 on representation of limits with integrands of class **R** and the Jensen-type inequalities in Proposition 12.27 (applied in a larger bounded Lipschitz domain $\Omega' \Supset \Omega$) give

$$\begin{aligned} \liminf_{j\to\infty} \mathscr{F}[u_j] &= \int_\Omega \langle f(x, \cdot), \nu_x \rangle + \langle f^\infty(x, \cdot), \nu_x^\infty \rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x)\, \mathrm{d}x \\ &\quad + \int_{\overline{\Omega}} \langle f^\infty(x, \cdot), \nu_x^\infty \rangle\, \mathrm{d}\lambda_\nu^s(x) \\ &\geq \int_\Omega f\Big(x, [\nu_x] + [\nu_x^\infty] \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathscr{L}^d}(x)\Big)\, \mathrm{d}x + \int_{\overline{\Omega}} f^\infty\big(x, [\nu_x^\infty]\big)\, \mathrm{d}\lambda_\nu^s(x) \\ &= \mathscr{F}[u]. \end{aligned}$$

This finishes the proof. □

We thus have the following existence theorem.

Theorem 12.29. *Let $f : \overline{\Omega} \times \mathbb{R}^{m\times d} \to [0, \infty)$ be such that the following assumptions hold:*

(i) *f is a Carathéodory integrand;*
(ii) *$|f(x, A)| \le M(1 + |A|)$ for some $M > 0$ and all $x \in \overline{\Omega}$, $A \in \mathbb{R}^{m\times d}$;*
(iii) *the strong recession function f^∞ exists in the sense of (11.8) and is (jointly) continuous on $\overline{\Omega} \times \mathbb{R}^{m\times d}$;*
(iv) *$f(x, \bullet)$ is quasiconvex for all $x \in \overline{\Omega}$;*
(v) *$\mu|A| \le f(x, A)$ for some $\mu > 0$ and all $(x, A) \in \overline{\Omega} \times \mathbb{R}^{m\times d}$.*

Then, the functional

$$\mathscr{F}[u] := \int_\Omega f(x, \nabla u(x))\, \mathrm{d}x + \int_\Omega f^\infty\Bigl(x, \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\Bigr)\, \mathrm{d}|D^s u|(x) \\ + \int_{\partial\Omega} f^\infty\bigl(x, u(x) \otimes n_\Omega(x)\bigr)\, \mathrm{d}\mathscr{H}^{d-1}(x), \qquad u \in \mathrm{BV}(\Omega; \mathbb{R}^m),$$

has a minimizer over the space $\mathrm{BV}(\Omega; \mathbb{R}^m)$.

Notes and Historical Remarks

The theory of generalized Young measures started with the article [99] by DiPerna and Majda, who were interested in turbulence concentrations in fluid dynamics (their measures, however, described L^2-concentrations). The L^1-framework was first introduced by Alibert & Bouchitté [6] and then developed into the form presented here in [168], partly inspired by the approaches to classical Young measures in the papers [39, 258]. Many authors have developed the theory further, a selection of some relevant papers is [121, 155, 171, 172, 258, 265].

In our setting we only work with the sphere compactification of $\mathbb{R}^N$. This is somewhat implicit, but the basic idea is explained in the opening remarks of this chapter, also see Problem 12.2. The theory can be extended to much more general target spaces and compactifications. For instance, $\mathbb{R}^N$ may be replaced by a Banach space X with the analytic Radon–Nikodým property, i.e., the validity of the Radon–Nikodým theorem for X-valued measures, for this see [24]. Further, we may employ another compactification of $\mathbb{R}^N$, for instance, the compactification generated by a separable, complete ring of continuous bounded functions, see Section 4.8 in [119] and also [60], or even the Stone–Čech compactification $\beta\mathbb{R}^N$. Finally, Ω may be replaced by a general finite measure space. Such generalizations are discussed in [6, 24, 57, 99, 172].

A generalized Young measure $\nu \in \mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^N)$ may also be described in the spirit of Berliocchi–Lasry [39] as follows:

$$\nu := (\mathscr{L}^d_x \llcorner \Omega) \otimes \nu_x \qquad \text{and} \qquad \nu^\infty := \lambda_\nu \otimes \nu^\infty_x.$$

Then,

$$\nu(D \times \mathbb{R}^N) = |\Omega \cap D| \quad \text{and} \quad \nu^\infty(D \times \mathbb{S}^{N-1}) = \lambda_\nu(D)$$

for all Borel sets $D \subset \overline{\Omega}$. So, ν can be understood as a classical Young measure with respect to $\mathscr{L}^d \llcorner \Omega$ and target space $\mathbb{R}^N$ and ν^∞ can be understood as a classical Young measure with respect to λ_ν and target space $\mathbb{S}^{N-1}$.

Alternative approaches to quantify concentration effects are varifolds [9–11, 125] and currents [115, 116, 134, 135]. Time-dependent generalized Young measures have also been developed for quasistatic evolution in plasticity theory [83–85]; in this context also see [95].

Lemma 12.14 and Lemma 12.28 are adapted from [6]. The characterization result of Theorem 12.19 also holds for BD-Young measures (i.e., those Young measures generated by symmetric derivatives of functions of bounded deformation), see [93].

The approach to BV-lower semicontinuity through generalized Young measures in Theorem 12.25 was first implemented in this form in [228], the version we present here includes the shortening possible by the Kirchheim–Kristensen Theorem 10.13. We remark that it is also possible to incorporate x-dependence into the strategy we employed for the proof of Theorem 11.7, but the Young measure proof gives this result immediately (if the strong recession function of the integrand exists). We refer to [16] for weak* lower semicontinuity results for functionals defined on PDE-constrained measures.

We finally mention the very recent work [23], where the Souček space is considered as a more natural space of underlying deformations for BV-Young measures with boundary concentrations.

Problems

12.1. Show that $f \in \mathrm{C}(\overline{\Omega} \times \mathbb{R}^N)$ is an element of $\mathbf{E}(\Omega; \mathbb{R}^N)$ if and only if f^∞ exists in the sense of (11.8).

12.2. Show that every element of $\mathbf{Y}^{\mathscr{M}}(\Omega; \mathbb{R}^N)$ can be identified with a classical L^∞-Young measure with values in the *sphere compactification* $\sigma\mathbb{R}^N \cong \mathbb{B}^N$ of $\mathbb{R}^N$ (where $\mathbb{R}^N$ is embedded into $\sigma\mathbb{R}^N$ via the map $v \mapsto (1 + |v|)^{-1}v$).

12.3. In the situation of Example 12.8, show that $u'_j \overset{\mathbf{Y}}{\to} \nu$ with

$$\nu_x = \delta_0 \quad \text{a.e.}, \qquad \lambda_\nu = \mathscr{L}^1 \llcorner (0, 1), \qquad \nu_x^\infty = \delta_{+1} \quad \text{a.e.}$$

12.4. The limit $x' \to x$ in the definition of f^∞ cannot be omitted without breaking important parts of the generalized Young measure theory:

(i) For $\Omega := (-1, 1)$ find a Carathéodory integrand $f : \overline{\Omega} \times \mathbb{R} \to \mathbb{R}$ such that $f^\infty(0, \cdot)$ is not well-defined in the sense of (11.8), but

$$\widetilde{f^\infty}(x,A) := \lim_{\substack{A' \to A \\ t \to \infty}} \frac{f(x, tA')}{t} = 0, \qquad (x,A) \in \Omega \times \mathbb{R}^N.$$

Note the omission of the limit $x' \to x$ in contrast to the definition (11.8) of f^∞.

(ii) Find a uniformly norm-bounded sequence $(v_j) \subset \mathrm{L}^1(-1,1)$ generating the Young measure $\nu \in \mathbf{Y}^{\mathscr{M}}((-1,1))$ with

$$\nu_x = \delta_0 \quad \text{a.e.}, \qquad \lambda_\nu = 2\delta_0, \qquad \nu_0^\infty = \delta_{+1}$$

and

$$\widetilde{\langle\!\langle f, \delta[v_j] \rangle\!\rangle} \not\to \widetilde{\langle\!\langle f, \nu \rangle\!\rangle},$$

where $\widetilde{\langle\!\langle \cdot, \cdot \rangle\!\rangle}$ is defined like the usual duality product $\langle\!\langle \cdot, \cdot \rangle\!\rangle$, but with f^∞ replaced by $\widetilde{f^\infty}$. Conclude that Proposition 12.11 cannot hold for $\widetilde{\langle\!\langle \cdot, \cdot \rangle\!\rangle}$ in place of $\langle\!\langle \cdot, \cdot \rangle\!\rangle$.

12.5. Let $u_j := \mathbb{1}_{(0,1/j)}$ on $\Omega := (0,1)$, for which $Du_j = -\delta_{1/j}$, $u_j \overset{*}{\rightharpoonup} 0$ in $\mathrm{BV}(0,1)$, and $Du_j \overset{\mathbf{Y}}{\to} \nu \in \mathbf{BVY}((0,1))$ with

$$\nu_x = \delta_0 \quad \text{a.e.}, \qquad \lambda_\nu = \delta_0, \qquad \nu_0^\infty = \delta_{-1} \quad \text{a.e.}$$

Prove that no sequence $(v_j) \subset \mathrm{BV}((0,1))$ with $v_j(0) = v_j(1) = 0$ can generate this ν. This shows that in Proposition 12.18 (ii) the assumption $\lambda_\nu(\partial\Omega) = 0$ cannot be omitted.

12.6. Prove that for any $\lambda \in \mathscr{M}^+(\overline{\Omega})$, there exists a BV-Young measure $\nu \in \mathbf{BVY}(\Omega; \mathbb{R}^{m \times d})$ with $[\nu] = 0$ and $\lambda_\nu = \lambda$.

12.7. Use the result of Problem 10.5 to derive the singular Jensen inequality in Proposition 12.27 without the use of the Kirchheim–Kristensen theorem or Alberti's Rank-One Theorem. Deduce from this the Lower Semicontinuity Theorem 12.25. This shows that neither Alberti's rank-one theorem nor the Kirchheim–Kristensen theorem are necessary to prove weak* lower semicontinuity in BV.

12.8. Let $\Omega_0, \Omega \subset \mathbb{R}^d$ be two bounded Lipschitz domains with $\Omega_0 \Subset \Omega$ and let $\nu \in \mathbf{BVY}(\Omega; \mathbb{R}^{m \times d})$. If $\lambda_\nu(\partial\Omega_0) = 0$, then the restriction $\nu \llcorner \Omega_0$, that is,

$$(\nu \llcorner \Omega_0)_x = \nu_x, \qquad (\nu \llcorner \Omega_0)_x^\infty = \nu_x^\infty \quad \text{for all } x \in \overline{\Omega_0}$$

and

$$\lambda_{\nu \llcorner \Omega_0} := \lambda_\nu \llcorner \Omega_0,$$

lies in $\mathbf{BVY}(\Omega; \mathbb{R}^{m \times d})$.

12.9. Show that $\mathbf{BVY}(\mathbb{R}^d; \mathbb{R}^{m \times d})$ and $\mathbf{BVY}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^{m \times d})$ are sequentially weakly* closed.

12.10. Show that in the situation of Theorem 12.25 also the functional (without boundary term)

$$\mathscr{F}[u] := \int_\Omega f(x, \nabla u(x))\, \mathrm{d}x + \int_\Omega f^\#\biggl(x, \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\biggr)\, \mathrm{d}|D^s u|(x), \qquad u \in \mathrm{BV}(\Omega; \mathbb{R}^m),$$

is lower semicontinuous with respect to weak* convergence in the space $\mathrm{BV}(\Omega; \mathbb{R}^m)$.

Chapter 13
Γ-Convergence

Often, a functional of interest depends on the value of a parameter, say a small $\varepsilon > 0$, as we have seen with the functionals $\mathscr{F}_\varepsilon$ from the examples on phase transitions and composite elastic materials in Sections 1.9 and 1.10, respectively. In these cases the goal often lies not in minimizing $\mathscr{F}_\varepsilon$ for one particular value of ε, but in determining the *asymptotic limit* of the minimization problems as $\varepsilon \downarrow 0$. Concretely, we need to identify, if possible, a *limit functional* $\mathscr{F}_0$ such that the minimizers and minimum values of the $\mathscr{F}_\varepsilon$ (if they exist) converge to the minimizers and minimum values of $\mathscr{F}_0$ as $\varepsilon \downarrow 0$. While many different situations can be considered, here we study the following prototypical problems:

- The phase transition example from Section 1.9 leads to a **singularly-perturbed problem**, where the $\mathscr{F}_\varepsilon$ have the form

$$\mathscr{F}_\varepsilon[u] := \int_\Omega f(u(x)) + \varepsilon^2 |\nabla u(x)|^2 \,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,2}(\Omega).$$

  Here, $f : \mathbb{R} \to [0, \infty)$ is a continuous multi-well potential. As the parameter $\varepsilon > 0$ tends to zero, the regularizing term becomes weaker and (approximate) minimizers may develop sharp interfaces, i.e., jumps. It is thus not unrealistic to expect that the limit functional $\mathscr{F}_0$ is of a fundamentally different nature. This is indeed the case, as we will see in the first part of this chapter.
- A higher-order and vectorial version of the previous problem occurs for the functionals

$$\mathscr{F}_\varepsilon[u] := \int_\Omega f(\nabla u(x)) + \varepsilon^2 |\nabla^2 u(x)|^2 \,\mathrm{d}x, \qquad u \in \mathrm{W}^{2,2}(\Omega; \mathbb{R}^m),$$

  where $f : \mathbb{R}^{m \times d} \to [0, \infty)$ is again a continuous multi-well potential. In this problem the limit functional is only known in some special cases.

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8_13

- The homogenization problem for composite materials from Section 1.10 leads to the following **periodic homogenization problem**: Let the integrand $f\colon \mathbb{R}^d \times \mathbb{R}^{m\times d} \to \mathbb{R}$ be 1-periodic in the first argument and consider

$$\mathscr{F}_\varepsilon[u] := \int_\Omega f\Big(\frac{x}{\varepsilon}, \nabla u(x)\Big)\,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m).$$

  This functional describes microscopic oscillations, for example, a fine lamination between different materials. The aim is to derive the macroscopic limit behavior. Naturally, we expect an x-independent integrand in the limit. This will be rigorously established in the second part of this chapter.

The unifying idea behind the above problems is that we believe $\mathscr{F}_\varepsilon$ for some nonzero but small $\varepsilon > 0$ to be the "true" model and $\mathscr{F}_0$ to be the "simplified" model, which, however, incorporates the lingering effects of the true model in the limit $\varepsilon \downarrow 0$. This is a very powerful approach since the study of $\mathscr{F}_0$ is not complicated by lower-order effects and thus the features of minimizers that we are really interested in may be much more apparent.

While the above problems have individual theories of considerable complexity, there is a common framework in which we can analyze sequences of functionals. This is the theory of Γ-convergence, which was introduced by Ennio De Giorgi in the 1970s. Its basic notion is a convergence of functionals that is compatible with the Direct Method. It has all the desired properties mentioned above, namely that the minima (infima) of the functionals $\mathscr{F}_\varepsilon$ converge to the minimum of the limit functional $\mathscr{F}_0$ and that, under an appropriate uniform coercivity hypothesis, the minimizers (or approximate minimizers) of $\mathscr{F}_\varepsilon$ converge in a suitable sense to a minimizer of $\mathscr{F}_0$.

13.1 Abstract Γ-Convergence

Let X be a complete metric space. The functional $\mathscr{F}_\infty\colon X \to \mathbb{R}\cup\{+\infty\}$ is called the **(sequential) Γ-limit** of the functionals $\mathscr{F}_k\colon X \to \mathbb{R}\cup\{+\infty\}$, $k \in \mathbb{N}$, if the following two conditions are satisfied:

(H1) For all sequences $(u_k) \subset X$ with $u_k \to u$ in X, the **lim inf-inequality** holds:

$$\mathscr{F}_\infty[u] \le \liminf_{k\to\infty} \mathscr{F}_k[u_k].$$

(H2) For all $u \in X$ there exists a **recovery sequence** $(u_k) \subset X$, that is, $u_k \to u$ in X and

$$\mathscr{F}_\infty[u] = \lim_{k\to\infty} \mathscr{F}_k[u_k].$$

The Γ-limit of the sequence $(\mathscr{F}_k)_k$, if it exists, is uniquely determined (see Problem 13.1) and denoted by Γ-$\lim_k \mathscr{F}_k$.

We note that assuming the first condition, the second condition in the definition of Γ-convergence can be replaced by the **lim sup-inequality**: For all $u \in X$ there exists a sequence $u_k \to u$ in X such that

$$\mathscr{F}_\infty[u] \geq \limsup_{k\to\infty} \mathscr{F}_k[u_k].$$

If a sequence of functionals $\mathscr{F}_k\colon X \to \mathbb{R} \cup \{+\infty\}$ converges locally uniformly to a functional $\mathscr{F}_\infty\colon X \to \mathbb{R} \cup \{+\infty\}$, i.e., for all $u \in X$ there exists some open neighborhood $U \Subset X$ of u such that $\sup_{u\in U} |\mathscr{F}_k[u] - \mathscr{F}_\infty[u]| \to 0$ as $k \to \infty$, and if $\mathscr{F}_\infty$ is lower semicontinuous, then the $\mathscr{F}_k$ also Γ-converge to $\mathscr{F}_\infty$: The lim inf-inequality holds because for any $u_k \to u$ we have $u_k \in U$ for k sufficiently large and hence

$$\liminf_{k\to\infty} \mathscr{F}_k[u_k] \geq \liminf_{k\to\infty} \mathscr{F}_\infty[u_k] \geq \mathscr{F}_\infty[u].$$

The lim sup-inequality holds for the constant recovery sequence.

Pointwise convergence, however, does not in general imply Γ-convergence:

Example 13.1. In $X = \mathbb{R}$ define for $k \in \mathbb{N}$,

$$\mathscr{F}_k[x] := -\delta_{-1/k}(x) + \delta_{1/k}(x) = \begin{cases} \pm 1 & \text{if } x = \pm 1/k, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\mathscr{F}_\infty[x] := -\delta_0(x) = \begin{cases} -1 & \text{if } x = 0, \\ 0 & \text{otherwise.} \end{cases}$$

We claim that Γ-$\lim_k \mathscr{F}_k = \mathscr{F}_\infty$. Indeed, on each open set $U \Subset \mathbb{R}\setminus\{0\}$, the functions $\mathscr{F}_k$ converge to the zero function uniformly. Together with the lower semicontinuity of $\mathscr{F}_\infty$, this implies Γ-convergence on $\mathbb{R}\setminus\{0\}$. For $x = 0$ we have Γ-$\lim \mathscr{F}_k(0) = -1$ by using the recovery sequence $x_k := -1/k$ (the lim inf-inequality is trivial).

This example shows that the Γ-limit does not necessarily coincide with the pointwise limit (which for our $\mathscr{F}_k$ is the zero function). In fact, even for the constant sequence of functionals $(\mathscr{F}_1)_k$, the Γ-limit is $-\delta_{-1} \neq \mathscr{F}_1$, as can be verified easily. One can also show that Γ-$\lim_k(-\mathscr{F}_k) = \mathscr{F}_\infty$, hence Γ-$\lim_k(-\mathscr{F}_k) \neq -$ Γ-$\lim_k \mathscr{F}_k$ and consequently the Γ-limit is not linear.

The theory of Γ-convergence is very useful and widely used in the calculus of variations. One of the reasons for this is the following fact.

Proposition 13.2. *$\mathscr{F}_\infty =$ Γ-$\lim_k \mathscr{F}_k$ is lower semicontinuous.*

Proof. We will prove $\mathscr{F}_\infty[u] \leq \liminf_{j\to\infty} \mathscr{F}_\infty[u_j]$ for all $u_j \to u$ in X. For every u_j choose a recovery sequence $(u_k^{(j)})_k$, i.e., $u_k^{(j)} \to u_j$ as $k \to \infty$ and $\mathscr{F}_\infty[u_j] =$

$\lim_{k\to\infty} \mathscr{F}_k[u_k^{(j)}]$. Next, denoting the metric in X by d, we choose a strictly increasing sequence of indices $(k(j))_j$ such that

$$d\big(u_{k(j)}^{(j)}, u_j\big) \le \frac{1}{j} \qquad \text{and} \qquad |\mathscr{F}_{k(j)}(u_{k(j)}^{(j)}) - \mathscr{F}_\infty[u_j]| \le \frac{1}{j}.$$

Then set

$$\tilde{u}_l := \begin{cases} u_{k(j)}^{(j)}, & \text{if } l = k(j) \text{ for some } j \in \mathbb{N}, \\ u, & \text{otherwise.} \end{cases}$$

We have that $\tilde{u}_k \to u$ and the lim inf-inequality implies

$$\mathscr{F}_\infty[u] \le \liminf_{k\to\infty} \mathscr{F}_k[\tilde{u}_k] \le \liminf_{j\to\infty} \mathscr{F}_{k(j)}[u_{k(j)}^{(j)}] = \liminf_{j\to\infty} \mathscr{F}_\infty[u_j].$$

Thus we have shown the lower semicontinuity of $\mathscr{F}_\infty$. □

As a consequence of the preceding proposition, the notion of Γ-convergence is not generated by a topology, because any constant sequence of a non-lower semicontinuous function has a different Γ-limit (we have already observed this fact in Example 13.1). It turns out, however, that in spaces of lower semicontinuous functions, the Γ-convergence is in fact generated by a topology, cf. Chapter 10 of [82] for details.

The most important property of Γ-convergence is that it entails the convergence of minima (or infima) and of the corresponding minimizers (or approximate minimizers) if the following notion of uniform coercivity is satisfied: A family $\{\mathscr{F}_k\}_k$ of functionals $\mathscr{F}_k \colon X \to \mathbb{R} \cup \{+\infty\}$ is called **equicoercive** if there exists a compact set $K \subset X$ with the property that

$$\inf_X \mathscr{F}_k = \inf_K \mathscr{F}_k \qquad \text{for all } k \in \mathbb{N}.$$

Clearly, without the equicoercivity no convergence of minima or minimizers can be expected, as the sequence $\mathscr{F}_k := -\delta_k$ in the space $X := \mathbb{R}$ shows.

Theorem 13.3. *Let $\mathscr{F}_k \colon X \to \mathbb{R} \cup \{+\infty\}$, $k \in \mathbb{N}$, be equicoercive functionals and assume that $\mathscr{F}_\infty = \Gamma\text{-}\lim_k \mathscr{F}_k$ exists. Then, $\mathscr{F}_\infty$ has a minimizer and*

$$\min_X \mathscr{F}_\infty = \lim_{k\to\infty} \inf_X \mathscr{F}_k.$$

In addition, all accumulation points of any precompact sequence $(u_k) \subset X$ with the property that $\liminf_{k\to\infty} \mathscr{F}_k[u_k] = \liminf_{k\to\infty} \inf_X \mathscr{F}_k$ are minimizers of $\mathscr{F}_\infty$.

Proof. Denote by K the compact set from the equicoercivity of the $\mathscr{F}_k$. For all $k \in \mathbb{N}$ choose $u_k \in K$ such that $|\mathscr{F}_k[u_k] - \inf_X \mathscr{F}_k| \le 1/k$. Then,

$$\liminf_{k\to\infty} \mathscr{F}_k[u_k] = \liminf_{k\to\infty} \inf_X \mathscr{F}_k.$$

Because all u_k lie in the compact set K, we can select a subsequence $(u_{k(j)})_j$ with $u_{k(j)} \to u_* \in X$ as $j \to \infty$ and

$$\lim_{j\to\infty} \mathscr{F}_{k(j)}[u_{k(j)}] = \liminf_{k\to\infty} \inf_X \mathscr{F}_k.$$

Then define the sequence $(\tilde{u}_k)_k$ as follows:

$$\tilde{u}_k := \begin{cases} u_{k(j)} & \text{if } k = k(j) \text{ for some } j \in \mathbb{N}, \\ u_* & \text{otherwise.} \end{cases}$$

We have $\tilde{u}_k \to u_*$ and, by the lim inf-inequality,

$$\inf_X \mathscr{F}_\infty \le \liminf_{k\to\infty} \mathscr{F}_k[\tilde{u}_k] \le \lim_{j\to\infty} \mathscr{F}_{k(j)}[u_{k(j)}] = \liminf_{k\to\infty} \inf_X \mathscr{F}_k. \tag{13.1}$$

Fix $\varepsilon > 0$. Let $u \in X$ be such that $\mathscr{F}_\infty[u] \le \inf_X \mathscr{F}_\infty + \varepsilon$ and take a recovery sequence $u_k \to u$ for u. We estimate

$$\limsup_{k\to\infty} \inf_X \mathscr{F}_k \le \limsup_{k\to\infty} \mathscr{F}_k[u_k] = \mathscr{F}_\infty[u] \le \inf_X \mathscr{F}_\infty + \varepsilon.$$

Let $\varepsilon \downarrow 0$ to get

$$\limsup_{k\to\infty} \inf_X \mathscr{F}_k \le \inf_X \mathscr{F}_\infty. \tag{13.2}$$

Combining (13.1) and (13.2), the first assertion of the theorem follows.

The second claim is clear if we take a sequence converging to a given accumulation point in the argument leading up to (13.1). □

For some arguments below it is useful to introduce the following two finer notions: The **(sequential) Γ-lower limit** $\Gamma\text{-}\liminf_k \mathscr{F}_k$ and the **(sequential) Γ-upper limit** $\Gamma\text{-}\limsup_k \mathscr{F}_k$ of the sequence of functionals $\mathscr{F}_k \colon X \to \mathbb{R} \cup \{+\infty\}$, $k \in \mathbb{N}$, are, respectively,

$$\Gamma\text{-}\liminf_k \mathscr{F}_k[u] := \inf \left\{ \liminf_{k\to\infty} \mathscr{F}_k[u_k] \;:\; u_k \to u \text{ in } X \right\},$$
$$\Gamma\text{-}\limsup_k \mathscr{F}_k[u] := \inf \left\{ \limsup_{k\to\infty} \mathscr{F}_k[u_k] \;:\; u_k \to u \text{ in } X \right\},$$

where $u \in X$.

The following lemma gives us an indirect way to establish that a sequence of functionals Γ-converges.

Lemma 13.4. $\Gamma\text{-}\liminf_k \mathscr{F}_k = \Gamma\text{-}\limsup_k \mathscr{F}_k = \mathscr{F}_\infty$ *if and only if* $\mathscr{F}_\infty = \Gamma\text{-}\lim_k \mathscr{F}_k$.

Proof. Assume first that $\Gamma\text{-}\liminf_k \mathscr{F}_k = \Gamma\text{-}\limsup_k \mathscr{F}_k = \mathscr{F}_\infty$. Then, for all $u_k \to u$ in X we have

$$\mathscr{F}_\infty[u] = \Gamma\text{-}\liminf_k \mathscr{F}_k[u] \le \liminf_{k\to\infty} \mathscr{F}_k[u_k],$$

that is, the lim inf-inequality holds. Moreover, for $u \in X$ and $n \in \mathbb{N}$ choose a sequence $(u_k^{(n)})_k$ with $u_k^{(n)} \to u$ as $k \to \infty$ and

$$\mathscr{F}_\infty[u] \ge \limsup_{k\to\infty} \mathscr{F}_k[u_k^{(n)}] - \frac{1}{n}.$$

It is always possible to find such a sequence since $\mathscr{F}_\infty[u] = \Gamma\text{-}\limsup_k \mathscr{F}_k[u]$. Now we inductively combine the sequences $(u_k^{(n)})_k$ into one sequence $\tilde{u}_k$ as follows: Let $(K(n))_n$ denote a growing sequence of indices such that

$$d(u_k^{(n)}, u) \le \frac{1}{n} \qquad \text{and} \qquad \mathscr{F}_\infty[u] \ge \mathscr{F}_k(u_k^{(n)}) - \frac{2}{n} \qquad \text{for all } k \ge K(n),$$

which exists by the definition of the upper limit. Then define

$$(\tilde{u}_k)_k := \left(u^{(1)}_{K(1)}, u^{(1)}_{K(1)+1}, \ldots, u^{(2)}_{K(2)}, u^{(2)}_{K(2)+1}, \ldots, u^{(l)}_{K(l)}, u^{(l)}_{K(l)+1}, \ldots\right)$$

and observe that $\tilde{u}_k \to u$ and $\mathscr{F}_\infty[u] \ge \limsup_{k\to\infty} \mathscr{F}_k[\tilde{u}_k]$. This proves the lim sup-inequality and hence the existence of a recovery sequence.

For the other direction assume $\mathscr{F}_\infty = \Gamma\text{-}\lim_k \mathscr{F}_k$. It is easy to see that for all $u \in X$ it holds that

$$\mathscr{F}_\infty[u] \le \Gamma\text{-}\liminf_k \mathscr{F}_k[u] \le \Gamma\text{-}\limsup_k \mathscr{F}_k[u] \le \mathscr{F}_\infty[u],$$

where the last inequality follows since recovery sequences are admissible in the definition of the Γ-upper limit. □

Recall from Chapter 7 that the **relaxation** of a functional $\mathscr{F} : X \to \mathbb{R} \cup \{+\infty\}$ is defined to be

$$\mathscr{F}_* := \sup\left\{ \mathscr{G} \;:\; \mathscr{G} \le \mathscr{F} \text{ and } \mathscr{G} \text{ is lower semicontinuous } \right\}.$$

This can be expressed via Γ-convergence as follows:

Proposition 13.5. *$\Gamma\text{-}\lim \mathscr{F} = \mathscr{F}_*$, where $\Gamma\text{-}\lim \mathscr{F}$ denotes the Γ-limit of the constant sequence $(\mathscr{F})_k$.*

Proof. First, we show that $\Gamma\text{-}\lim \mathscr{F}$ always exists. Indeed, by the definition of the Γ-lower limit we can for all $u \in X$ and all $\varepsilon > 0$ find a sequence $u_k \to u$ such that

$$\lim_{k\to\infty} \mathscr{F}[u_k] \le \Gamma\text{-}\liminf_k \mathscr{F}[u] + \varepsilon.$$

Using this sequence in the definition of the Γ-upper limit and letting $\varepsilon \to 0$, we conclude that $\Gamma\text{-}\limsup_k \mathscr{F}[u] \le \Gamma\text{-}\liminf_k \mathscr{F}[u]$. Then, Lemma 13.4 shows that $\Gamma\text{-}\lim_k \mathscr{F}$ exists.

For any lower semicontinuous function $\mathscr{G} \le \mathscr{F}$, we have $\mathscr{G} = \Gamma\text{-}\lim \mathscr{G}$ since for all $u_k \to u$ it holds that

$$\begin{aligned} \mathscr{G}[u] &\le \inf \left\{ \liminf_{k\to\infty} \mathscr{G}[u_k] \;:\; u_k \to u \right\} \\ &= \Gamma\text{-}\liminf_k \mathscr{G}[u] \\ &\le \Gamma\text{-}\limsup_k \mathscr{G}[u] \\ &\le \mathscr{G}[u] \end{aligned}$$

since the constant sequence is admissible in the definition of the Γ-upper limit. Thus $\mathscr{G} = \Gamma\text{-}\lim \mathscr{G}$ by Lemma 13.4. Consequently,

$$\mathscr{G} = \Gamma\text{-}\lim \mathscr{G} \le \Gamma\text{-}\lim \mathscr{F}.$$

Taking the supremum over all such $\mathscr{G}$, we obtain $\mathscr{F}_* \le \Gamma\text{-}\lim \mathscr{F}$.

To prove the converse inequality, note that $\Gamma\text{-}\lim \mathscr{F}$ is lower semicontinuous by Proposition 13.2 and $\Gamma\text{-}\lim \mathscr{F} \le \mathscr{F}$ (see above). Then it follows immediately that $\Gamma\text{-}\lim \mathscr{F} \le \mathscr{F}_*$. □

13.2 Sharp-Interface Limits

As an important example of singularly-perturbed functionals we consider the following, which we saw first in Section 1.9 in relation to phase transitions. For $u \in \mathrm{L}^1(\Omega)$ we set

$$\mathscr{F}_\varepsilon[u] := \begin{cases} \displaystyle\int_\Omega \frac{1}{\varepsilon} f(u(x)) + \varepsilon |\nabla u(x)|^2 \,\mathrm{d}x & \text{if } u \in \mathrm{W}^{1,2}(\Omega) \text{ and } \int_\Omega u \,\mathrm{d}x = \gamma, \\ +\infty & \text{otherwise,} \end{cases}$$

where $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain, $f : \mathbb{R} \to [0, \infty)$ is a continuous double-well potential with zeros (only) at $\alpha, \beta \in \mathbb{R}$ $(\alpha < \beta)$, $\gamma \in (\alpha|\Omega|, \beta|\Omega|)$, and $\varepsilon > 0$ is a small parameter.

For the definition of a limit candidate of the $\mathscr{F}_\varepsilon$ as $\varepsilon \downarrow 0$, first recall from (9.13) the definition of the perimeter of a Borel set $E \subset \Omega$, namely,

$$\mathrm{Per}_\Omega(E) = |D\mathbb{1}_E|(\Omega) = \sup \left\{ \int_E \operatorname{div} \varphi \,\mathrm{d}x \;:\; \varphi \in \mathrm{C}^1_c(\Omega; \mathbb{R}^d),\ \|\varphi\|_\infty \le 1 \right\}.$$

If E has a smooth boundary, then $\mathrm{Per}_\Omega(E) = \mathscr{H}^{d-1}(E \cap \Omega)$. Then set for $u \in \mathrm{L}^1(\Omega)$,

$$\mathscr{F}_0[u] := \begin{cases} \sigma_0 \operatorname{Per}_\Omega\big(\{ x \in \Omega : u(x) = \alpha \}\big) & \text{if } u \in \mathrm{BV}(\Omega; \{\alpha, \beta\}) \\ & \quad \text{and } \int_\Omega u \,\mathrm{d}x = \gamma, \\ +\infty & \text{otherwise,} \end{cases}$$

where

$$\sigma_0 := 2 \int_\alpha^\beta \sqrt{f(s)} \,\mathrm{d}s.$$

Note that in Section 1.9 we had $\alpha = -1$, $\beta = 1$ and also additionally required that $u(x) \in [-1, 1]$ for almost every $x \in \Omega$. As will become obvious from the proofs, all of the following will also apply to this situation.

The main result of this section is the following:

Theorem 13.6 (**Modica–Mortola 1977 [189, 190]**). *The functionals $\mathscr{F}_\varepsilon$ Γ-converge to $\mathscr{F}_0$ as $\varepsilon \downarrow 0$ with respect to the strong L^1-topology.*

For the proof without loss of generality we assume that $\alpha = -1$ and $\beta = 1$, which can be accomplished along the lines of the transformation applied in Section 1.9. To establish the sought Γ-convergence, we need to show

(a) the lim inf-inequality

$$\mathscr{F}_0[u] \le \liminf_{\varepsilon \downarrow 0} \mathscr{F}_\varepsilon[u_\varepsilon] \tag{13.3}$$

for all sequences $u_\varepsilon \to u$ in $\mathrm{L}^1(\Omega)$;

(b) for all $u \in \mathrm{L}^1(\Omega)$ the existence of a recovery sequence $(u_\varepsilon)_{\varepsilon>0} \subset \mathrm{L}^1(\Omega)$ such that $u_\varepsilon \to u$ in L^1 as $\varepsilon \downarrow 0$ and

$$\mathscr{F}_0[u] = \lim_{\varepsilon \downarrow 0} \mathscr{F}_\varepsilon[u_\varepsilon].$$

The lim inf-inequality turns out to be fairly straightforward to prove:

Proof of Theorem 13.6: lim inf-inequality. Let $u_\varepsilon \to u$ in $\mathrm{L}^1(\Omega)$ and assume that $\liminf_{\varepsilon\downarrow 0} \mathscr{F}_\varepsilon[u_\varepsilon] < \infty$ (otherwise there is nothing to show), whereby we may require that $u_\varepsilon \in \mathrm{W}^{1,2}(\Omega)$ for all $\varepsilon > 0$. Then, there exists a sequence $\varepsilon_n \downarrow 0$ (as $n \to \infty$) such that $u_{\varepsilon_n} \to u$ almost everywhere and Fatou's lemma implies

$$\int_\Omega f(u(x)) \,\mathrm{d}x \le \liminf_{n\to\infty} \int_\Omega f(u_{\varepsilon_n}(x)) \,\mathrm{d}x \le \liminf_{n\to\infty} \big(\varepsilon_n \mathscr{F}_{\varepsilon_n}[u_{\varepsilon_n}]\big) = 0.$$

Thus, $u(x) \in \{-1, 1\}$ almost everywhere.

Next, we observe from Young's inequality that

$$\begin{aligned} \liminf_{n\to\infty} \mathscr{F}_{\varepsilon_n}[u_{\varepsilon_n}] &\ge \liminf_{n\to\infty} \int_\Omega 2\sqrt{f(u_{\varepsilon_n}(x))} \cdot |\nabla u_{\varepsilon_n}(x)| \,\mathrm{d}x \\ &\ge \liminf_{n\to\infty} \int_\Omega 2|\nabla(h \circ u_{\varepsilon_n})(x)| \,\mathrm{d}x, \end{aligned}$$

where we have set

$$h(t) := \int_0^t \sqrt{f(s)}\,\mathrm{d}s.$$

From the L^1-lower semicontinuity of the total variation norm (see Problem 13.2) we then get

$$\begin{aligned}\liminf_{n\to\infty} \mathscr{F}_{\varepsilon_n}[u_{\varepsilon_n}] &\geq 2|D(h\circ u)|(\Omega)\\ &= 2(h(1)-h(-1))\,\mathrm{Per}_\Omega\big(\{\, x\in\Omega \,:\, u(x)=-1\,\}\big)\\ &= \mathscr{F}_0[u]\end{aligned}$$

since $\sigma_0 = 2(h(1)-h(-1))$. In particular, $u \in \mathrm{BV}(\Omega;\{-1,1\})$. This establishes the lim inf-inequality (13.3). □

For the existence of a recovery sequence, we first need a few technical preparations.

Lemma 13.7. *Let the Borel set $E \subset \Omega$ be of finite perimeter, i.e.,* $\mathrm{Per}_\Omega(E) < \infty$*, and assume that*

$$E \text{ and } \Omega\setminus E \text{ both contain a non-empty open ball.} \tag{13.4}$$

Then, there exists a sequence of open and bounded sets $E_n \subset \mathbb{R}^d$ with smooth boundaries such that

$$|E_n\cap\Omega| = |E|, \tag{13.5}$$

the (measure-theoretic) transversality condition

$$\mathscr{H}^{d-1}(\partial E_n \cap \partial\Omega) = 0 \tag{13.6}$$

holds, and

$$|(E\,\Delta\,E_n)\cap\Omega| \to 0, \quad \mathrm{Per}_\Omega(E_n)\to\mathrm{Per}_\Omega(E) \qquad \text{as } n\to\infty. \tag{13.7}$$

Here, $A\,\Delta\,B := (A\setminus B)\cup(B\setminus A)$ denotes the **symmetric difference** between the sets A and B.

Proof. The idea is to mollify $\mathbb{1}_E$ and then to select for the E_n suitable superlevel sets. The technical details, however, are somewhat involved.

Step 1. Let $u \in (\mathrm{BV}\cap\mathrm{L}^\infty)(\mathbb{R}^d)$ be an extension of $\mathbb{1}_E \in (\mathrm{BV}\cap\mathrm{L}^\infty)(\Omega)$ with the property that $|Du|(\partial\Omega) = 0$. Since Ω was assumed to be a bounded Lipschitz domain, such an extension of $\mathbb{1}_E$ always exists; this result is recalled in Section 10.3. Let $(\eta_\delta)_{\delta>0} \subset \mathrm{C}_c^\infty(\mathbb{R}^d)$ be a family of mollifiers as in Appendix A.5; in particular, $\operatorname{supp}\eta_\delta \subset B(0,\delta)$. Set

$$u_\delta := \eta_\delta \star u,$$

for which it holds that $u_\delta \to u$ in L^1 and in measure, as well as $|Du_\delta|(\mathbb{R}^d) \to |Du|(\mathbb{R}^d)$ as $\delta \downarrow 0$. The latter fact can be seen by applying the lower semicontinuity of the total variation norm (see Problem 13.2) to Ω and to $\mathbb{R}^d \setminus \overline{\Omega}$, and also using that $|Du|(\partial\Omega) = 0$ as well as properties of the extension operator.

By assumption, there exists an $\varepsilon > 0$ and $x_1, x_2 \in \mathbb{R}^d$ such that

$$B(x_1, 2\varepsilon) \subset E, \qquad B(x_2, 2\varepsilon) \subset \Omega \setminus E.$$

Then,

$$u_\delta = u \quad \text{on } \overline{B(x_1, \varepsilon)} \cup \overline{B(x_2, \varepsilon)} \qquad \text{if } \delta < \varepsilon.$$

Now, since $u_\delta \to u$ in measure, for every $n \in \mathbb{N}$ we may choose a positive $\delta_n < \min\{1/n, \varepsilon\}$ such that

$$\left|\left\{ x \in \Omega \;\boldsymbol{:}\; \left|u_{\delta_n}(x) - u(x)\right| \geq \frac{1}{n} \right\}\right| \leq \frac{1}{n}. \tag{13.8}$$

Next, we let

$$p_n := \operatorname*{ess\,inf}_{t \in (1/n, 1-1/n)} \operatorname{Per}_\Omega\big(\{ x \in \mathbb{R}^d \;\boldsymbol{:}\; u_{\delta_n}(x) > t \}\big),$$

which is defined to be the largest almost-everywhere lower bound of the Lebesgue-measurable function $t \mapsto \operatorname{Per}_\Omega(\{ x \in \mathbb{R}^d \;\boldsymbol{:}\; u_{\delta_n}(x) > t \})$ in the interval $(1/n, 1 - 1/n)$ (the right-hand side is Lebesgue-measurable in t since it can be written as the supremum over countably many continuous functions). For each $n \in \mathbb{N}$ choose $t_n \in (1/n, 1 - 1/n)$ such that the following three conditions hold:

(a) $\operatorname{Per}_\Omega\big(\{ x \in \mathbb{R}^d \;\boldsymbol{:}\; u_{\delta_n}(x) > t_n \}\big) \leq p_n + \dfrac{1}{n}$;
(b) $\nabla u_{\delta_n}(x) \neq 0$ for all $x \in \mathbb{R}^d$ such that $u_{\delta_n}(x) = t_n$;
(c) $\mathscr{H}^{d-1}\big(\{ x \in \partial\Omega \;\boldsymbol{:}\; u_{\delta_n}(x) = t_n \}\big) = 0$.

By definition, (a) holds for a non-negligible set of t's and (b) is satisfied for a Lebesgue-full set of t's by the Sard Theorem A.17. The fact that $\mathscr{H}^{d-1}(\partial\Omega) < \infty$ implies (c) for all but countably many t's. Then, with t_n defined, we set

$$D_n := \left\{ x \in \mathbb{R}^d \;\boldsymbol{:}\; u_{\delta_n}(x) > t_n \right\}, \qquad \lambda_n := |D_n \cap \Omega| - |E|,$$

and, letting $r_n > 0$ such that $|B(x_1, r_n)| = |B(x_2, r_n)| = |\lambda_n|$, we define

$$E_n := \begin{cases} D_n \setminus \overline{B(x_1, r_n)} & \text{if } \lambda_n > 0, \\ D_n & \text{if } \lambda_n = 0, \\ D_n \cup B(x_2, r_n) & \text{if } \lambda_n < 0. \end{cases}$$

Clearly, the sets E_n so defined are open and bounded. From (b) we furthermore infer that ∂E_n is smooth. It remains to show (13.5), (13.6), and (13.7).

Step 2. If $x \in (D_n \cap \Omega) \setminus E$, then $u_{\delta_n}(x) > t_n > 1/n$ and $u(x) = 0$, whereas if $x \in E \setminus (D_n \cap \Omega)$, then $u_{\delta_n}(x) \le t_n < 1 - 1/n$ and $u(x) = 1$. Hence, also using (13.8),

$$|\lambda_n| \le |(D_n \cap \Omega) \,\Delta\, E| \le \left| \left\{ x \in \Omega \,:\, \left|u_{\delta_n}(x) - u(x)\right| \ge \frac{1}{n} \right\} \right| \le \frac{1}{n}. \tag{13.9}$$

Consequently,

$$r_n \to 0 \qquad \text{as } n \to \infty. \tag{13.10}$$

Then, for n large enough, $r_n < \varepsilon/2$, so that $\overline{B(x_1, r_n)} \subset B(x_1, \varepsilon)$ and $\overline{B(x_2, r_n)} \subset B(x_2, \varepsilon)$. Moreover, since $\delta_n < \varepsilon$, we have $\overline{B(x_1, \varepsilon)} \subset D_n \cap \Omega$ and $\overline{B(x_2, \varepsilon)} \subset \Omega \setminus D_n$. We conclude that

$$\begin{aligned} |E_n \cap \Omega| &= \begin{cases} |D_n \cap \Omega| - |\overline{B(x_1, r_n)}| & \text{if } \lambda_n > 0, \\ |D_n \cap \Omega| & \text{if } \lambda_n = 0, \\ |D_n \cap \Omega| + |B(x_2, r_n)| & \text{if } \lambda_n < 0 \end{cases} \\ &= |E|, \end{aligned}$$

which shows (13.5) after discarding some elements at the beginning of the sequence (E_n).

Step 3. In a similar fashion to the previous step we deduce that if $\lambda_n \neq 0$, then

$$\partial E_n \cap \partial \Omega = (\partial D_n \cup \partial B(x_i, r_n)) \cap \partial \Omega$$

for either $i = 1$ or $i = 2$, depending on whether $\lambda_n > 0$ or $\lambda_n < 0$. On the other hand, $\partial B(x_i, r_n) \cap \partial \Omega = \emptyset$, so that by (c) we see that

$$\mathscr{H}^{d-1}(\partial E_n \cap \partial \Omega) = \mathscr{H}^{d-1}(\partial D_n \cap \partial \Omega) = 0.$$

This shows the transversality condition (13.6).

Step 4. To prove (13.7) we first observe that $\overline{B(x_1, \varepsilon)} \subset D_n \cap \Omega$ and $\overline{B(x_2, \varepsilon)} \subset \Omega \setminus D_n$. Then, by (13.9),

$$|(E_n \cap \Omega) \,\Delta (D_n \cap \Omega)| = |\lambda_n| \to 0 \qquad \text{as } n \to \infty.$$

Thus, again by (13.9),

$$\lim_{n\to\infty} |(E \,\Delta\, E_n) \cap \Omega| = \lim_{n\to\infty} |(E_n \cap \Omega) \,\Delta\, E| = \lim_{n\to\infty} |(D_n \cap \Omega) \,\Delta\, E| = 0.$$

This is the first part of (13.7).

Because $\overline{B(x_1, \varepsilon)} \subset D_n \cap \Omega$ and $\overline{B(x_2, \varepsilon)} \subset \Omega \setminus D_n$, we have

$$\mathrm{Per}_{\Omega}(E_n) = \mathrm{Per}_{\Omega}(D_n) + \mathscr{H}^{d-1}(\partial B(x_i, r_n)) \tag{13.11}$$

for $i = 1$ or $i = 2$ (depending on the sign of λ_n). Moreover, we have that $\mathbb{1}_{E_n} \to \mathbb{1}_E$ in $\mathrm{L}^1(\Omega)$ by the first part of (13.7), which was shown above. The perimeter is lower semicontinuous under this convergence, see Problem 13.2. Consequently,

$$\begin{aligned}\mathrm{Per}_{\Omega}(E) &\le \liminf_{n\to\infty} \mathrm{Per}_{\Omega}(E_n) \\ &= \liminf_{n\to\infty} \left[\mathrm{Per}_{\Omega}(D_n) + \mathscr{H}^{d-1}(\partial B(x_i, r_n))\right] \\ &= \liminf_{n\to\infty} \mathrm{Per}_{\Omega}(D_n),\end{aligned} \tag{13.12}$$

where we also employed (13.10). On the other hand, by (a) above,

$$\mathrm{Per}_{\Omega}(D_n) \le p_n + \frac{1}{n} \le \mathrm{Per}_{\Omega}\big(\{\, x \in \mathbb{R}^d : u_{\delta_n}(x) > t \,\}\big) + \frac{1}{n} \tag{13.13}$$

for all $n \in \mathbb{N}$ and almost every $t \in (1/n, 1 - 1/n)$.

The **Fleming–Rishel coarea formula in** BV states that for every $u \in \mathrm{BV}(\Omega)$ the set $\{\, x \in \Omega : u(x) > t \,\}$ has finite perimeter for $\mathscr{L}^1$-almost every $t \in \mathbb{R}$ and

$$|Du|(\Omega) = \int_{-\infty}^{+\infty} \mathrm{Per}_{\Omega}(\{\, x \in \Omega : u(x) > t \,\})\, \mathrm{d}t.$$

A proof of this fundamental fact can be found in Theorem 3.40 of [15]. We can now integrate (13.13) from $1/n$ to $1 - 1/n$ and employ the said coarea formula to get

$$\left(1 - \frac{2}{n}\right) \mathrm{Per}_{\Omega}(D_n) \le \int_{\Omega} |\nabla u_{\delta_n}| \, \mathrm{d}x + \frac{1}{n}\left(1 - \frac{2}{n}\right). \tag{13.14}$$

On the other hand, from $|Du_{\delta_n}|(\mathbb{R}^d) \to |Du|(\mathbb{R}^d)$, which we justified at the beginning of the proof, we deduce that

$$\lim_{n\to\infty} \int_{\Omega} |\nabla u_{\delta_n}| \, \mathrm{d}x = |Du|(\Omega) = \mathrm{Per}_{\Omega}(E).$$

We then see from (13.14) that

$$\limsup_{n\to\infty} \mathrm{Per}_{\Omega}(D_n) \le \mathrm{Per}_{\Omega}(E).$$

Combining this with (13.12), we arrive at

$$\mathrm{Per}_{\Omega}(E) \le \liminf_{n\to\infty} \mathrm{Per}_{\Omega}(D_n) \le \limsup_{n\to\infty} \mathrm{Per}_{\Omega}(D_n) \le \mathrm{Per}_{\Omega}(E),$$

and so, also using (13.11),

$$\lim_{n\to\infty} \mathrm{Per}_\Omega(E_n) = \lim_{n\to\infty} \mathrm{Per}_\Omega(D_n) = \mathrm{Per}_\Omega(E).$$

This is the second part of (13.7). □

If we dispense with the assumption (13.4), we can still prove the following result.

Lemma 13.8. *Let the Borel set $E \subset \Omega$ be of finite perimeter, i.e.,* $\mathrm{Per}_\Omega(E) < \infty$. *Then, there exists a sequence of sets $D_n \subset \mathbb{R}^d$ with smooth boundaries such that*

$$|(E \,\Delta\, D_n) \cap \Omega| \to 0, \quad \mathrm{Per}_\Omega(D_n) \to \mathrm{Per}_\Omega(E) \qquad \text{as } n \to \infty.$$

Proof. The proof is contained in the parts of the proof for the previous lemma relating to the sets D_n, for which the assumption (13.4) is not used. □

Lemma 13.9. *Let $E \subset \mathbb{R}^d$ be open with a smooth, compact, non-empty boundary and such that the measure-theoretic transversality condition*

$$\mathscr{H}^{d-1}(\partial E \cap \partial\Omega) = 0 \tag{13.15}$$

holds. Define the function $\delta_E\colon \mathbb{R}^d \to \mathbb{R}$ via

$$\delta_E(x) := \begin{cases} -\operatorname{dist}(x, \partial E) & \text{if } x \in E, \\ \operatorname{dist}(x, \partial E) & \text{if } x \notin E. \end{cases} \tag{13.16}$$

Then, δ_E is Lipschitz continuous, $|\nabla \delta_E| = 1$ almost everywhere, and for all $t \in \mathbb{R}$,

$$\lim_{t\to 0} \mathscr{H}^{d-1}(S_t \cap \Omega) = \mathscr{H}^{d-1}(\partial E \cap \Omega), \tag{13.17}$$

where $S_t := \left\{x \in \mathbb{R}^d : \delta_E(x) = t\right\}$.

Proof. Step 1. We first show for the sets

$$F_r := \left\{ x \in E \;:\; \operatorname{dist}(x, \partial E) = r \right\}, \qquad r > 0,$$

that

$$\lim_{r\downarrow 0} \mathscr{H}^{d-1}(F_r) = \mathscr{H}^{d-1}(\partial E). \tag{13.18}$$

For this we recall the geometric fact, proved in detail, for instance, in Section 14.6 of [136], that for small $r > 0$ there exists a diffeomorphism φ from $V_r := \{x \in E : 0 < \operatorname{dist}(x, \partial E) < r\}$ to $\partial E \times (0, r) \subset \mathbb{R}^{d+1}$ with

$$\det \nabla\varphi(x) = \prod_{i=1}^{d-1} \big(1 - \kappa_i(\hat{\varphi}(x)) \operatorname{dist}(x, \partial E)\big) \geq \mu > 0, \tag{13.19}$$

where $\hat{\varphi}(x)$ is the component of $\varphi(x)$ on ∂E (more precisely, $\hat{\varphi} := \pi \circ \varphi \subset \partial E$, where $\pi(y, s) := y$) and $\kappa_1(\hat{x}), \ldots, \kappa_{d-1}(\hat{x})$ denote the principal curvatures of ∂E at $\hat{x} \in \partial E$. Also, $x \mapsto \mathrm{dist}(x, \partial E)$ is smooth on $\overline{V}_r$ and, with the unit outward normal vector n on ∂E,

$$\nabla[\mathrm{dist}(x, \partial E)] = -n(\hat{\varphi}(x)) \qquad \text{for all } x \in \overline{V}_r.$$

Finally, if m_r is the normal vector to F_r, oriented outwards with respect to V_r, then

$$\nabla[\mathrm{dist}(x, \partial E)] = m_r(x) \qquad \text{for all } x \in F_r.$$

By the Gauss–Green theorem, noticing that $\partial V_r = F_r \cup \partial E$ (disjointly),

$$\begin{aligned}\int_{V_r} \Delta[\mathrm{dist}(\bullet, \partial E)] \, \mathrm{d}x &= \int_{F_r} \nabla[\mathrm{dist}(\bullet, \partial E)] \cdot m_r \, \mathrm{d}\mathscr{H}^{d-1} \\ &\quad + \int_{\partial E} \nabla[\mathrm{dist}(\bullet, \partial E)] \cdot n \, \mathrm{d}\mathscr{H}^{d-1} \\ &= \mathscr{H}^{d-1}(F_r) - \mathscr{H}^{d-1}(\partial E).\end{aligned}$$

Using the lower bound in (13.19), we have $\det \nabla\varphi^{-1} \leq \mu^{-1}$ and thus

$$\begin{aligned}|V_r| &= \int_{\partial E} \int_0^r \det \nabla\varphi^{-1}(y, s) \, \mathrm{d}s \, \mathrm{d}\mathscr{H}^{d-1}(y) \\ &\leq \frac{r\mathscr{H}^{d-1}(\partial E)}{\mu} \\ &\to 0 \qquad \text{as } r \downarrow 0.\end{aligned} \tag{13.20}$$

This directly implies (13.18).

Step 2. Next, we show that if the transversality condition (13.15) holds, then

$$\lim_{r\downarrow 0} \mathscr{H}^{d-1}(F_r \cap \Omega) = \mathscr{H}^{d-1}(\partial E \cap \Omega). \tag{13.21}$$

From (13.20) we get that $\mathbb{1}_{E \setminus V_r} \to \mathbb{1}_E$ in L^1. From the lower semicontinuity of the perimeter (see Problem 13.2) we thus infer that

$$\begin{aligned}\mathscr{H}^{d-1}(\partial E \cap \Omega) &= \mathrm{Per}_\Omega(E) \\ &\leq \liminf_{r\downarrow 0} \mathrm{Per}_\Omega(E \setminus V_r) \\ &= \liminf_{r\downarrow 0} \mathscr{H}^{d-1}(F_r \cap \Omega),\end{aligned} \tag{13.22}$$

where the last equality follows since $\partial(E \setminus V_r) = F_r$, which is a smooth manifold.

Conversely, we have

$$\mathscr{H}^{d-1}(F_r \cap \Omega) \le \mathscr{H}^{d-1}(F_r) - \mathscr{H}^{d-1}(F_r \cap (\mathbb{R}^d \setminus \Omega)).$$

By a similar argument as before,

$$\mathscr{H}^{d-1}(\partial E \cap (\mathbb{R}^d \setminus \overline{\Omega})) \le \liminf_{r \downarrow 0} \mathscr{H}^{d-1}(F_r \cap (\mathbb{R}^d \setminus \overline{\Omega})).$$

Thus, using (13.18) and the transversality condition (13.15),

$$\begin{aligned} \limsup_{r\downarrow 0} \mathscr{H}^{d-1}(F_r \cap \Omega) &\le \limsup_{r\downarrow 0} \mathscr{H}^{d-1}(F_r) - \liminf_{r\downarrow 0} \mathscr{H}^{d-1}(F_r \cap (\mathbb{R}^d \setminus \overline{\Omega})) \\ &\le \mathscr{H}^{d-1}(\partial E) - \mathscr{H}^{d-1}(\partial E \cap (\mathbb{R}^d \setminus \overline{\Omega})) \\ &= \mathscr{H}^{d-1}(\partial E \cap \Omega). \end{aligned}$$

Combining this with (13.22), we arrive at (13.21).

Step 3. Finally, we prove the statements about δ_E. Clearly, we have for all $x, y \in \mathbb{R}^d$ that $\delta_E(y) \le \delta_E(x) + |x - y|$, so δ_E is Lipschitz continuous and $|\nabla \delta_E| \le 1$ almost everywhere. On the other hand, for every $x \in \mathbb{R}^d$ there exists an $\bar{x} \in \partial E$ such that $|\delta_E(x)| = |x - \bar{x}|$. For every y on the connecting line between x and $\bar{x}$, it holds that $|\delta_E(y)| = |y - \bar{x}|$ and so $|\nabla \delta_E| = 1$ almost everywhere. Finally, (13.17) follows from Step 2 (i.e., from (13.21)) applied to E and $\mathbb{R}^d \setminus \overline{E}$. □

We can now complete the proof of the Modica–Mortola theorem.

Proof of Theorem 13.6: *Recovery sequence*. Since $\mathscr{F}_0[u] = +\infty$ for $u \in (\mathrm{L}^1 \setminus \mathrm{BV})(\Omega)$ or if u takes values other than ± 1, we only need to explicitly construct a recovery sequence for maps $u \in \mathrm{BV}(\Omega; \{-1, 1\})$, which we henceforth assume.

Step 1. Set

$$E := E_{-1} = \left\{ x \in \Omega \ : \ u(x) = -1 \right\} \subset \Omega,$$

so that

$$u = -\mathbb{1}_{E\cap\Omega} + \mathbb{1}_{\Omega\setminus E}, \qquad |E \cap \Omega| = |\Omega \setminus E| - \gamma,$$

where we recall that $\gamma \in (-|\Omega|, |\Omega|)$ is the parameter such that $\int_\Omega u \,\mathrm{d}x = \gamma$. Since u has bounded variation, we know that E is of finite perimeter in Ω. From Lemma 13.8 we infer that u may be approximated by $v = -\mathbb{1}_{F\cap\Omega} + \mathbb{1}_{\Omega\setminus F}$, where $F \subset \Omega$ is of finite perimeter and additionally has a smooth boundary. By the properties of the approximation, for any fixed $\eta > 0$ we may choose v to also satisfy

$$\left| \mathscr{F}_0[u] - \mathscr{F}_0[v] \right| < \eta.$$

Since $\gamma \in (-|\Omega|, |\Omega|)$, both $F \cap \Omega$ and $F \setminus \Omega$ contain a non-empty open ball (this also uses the smoothness of the boundary of ∂F). Thus, Lemma 13.7 becomes applicable and via another approximation step we may find $w = -\mathbb{1}_{G\cap\Omega} + \mathbb{1}_{\Omega\setminus G}$

such that $G \subset \Omega$ is of finite perimeter, has a smooth boundary, the measure-theoretic transversality condition

$$\mathscr{H}^{d-1}(\partial G \cap \partial \Omega) = 0$$

holds, and

$$\left|\mathscr{F}_0[v] - \mathscr{F}_0[w]\right| < \eta.$$

The preceding approximation arguments show that we only need to find a recovery sequence for $u \in \mathrm{BV}(\Omega; \{-1, 1\})$ with the property that $E = E_{-1} \subset \Omega$ is of finite perimeter, has a smooth boundary, and the measure-theoretic transversality condition

$$\mathscr{H}^{d-1}(\partial E \cap \partial \Omega) = 0$$

holds. Finally, we may further assume that E is in fact open since the boundary ∂E is a Lebesgue negligible set, on which we may modify the representative of our u freely.

Step 2. Define for $\varepsilon > 0$ and $s \in [-1, 1]$, $t \in \mathbb{R}$ the functions

$$\varphi_\varepsilon(s) := \int_{-1}^{s} \frac{\varepsilon}{\sqrt{\varepsilon + f(r)}} \,\mathrm{d}r, \qquad \psi_\varepsilon(t) := \begin{cases} -1 & \text{if } t \le 0, \\ \varphi_\varepsilon^{-1}(t) & \text{if } 0 < t < \varphi_\varepsilon(1), \\ 1 & \text{if } \varphi_\varepsilon(1) \le t. \end{cases}$$

This is possible since φ_ε as defined above is strictly increasing, hence invertible. Then, with $\delta_E \colon \mathbb{R}^d \to \mathbb{R}$ defined as in (13.16) for our set E as above, we set

$$u_\varepsilon(x) := \psi_\varepsilon(\delta_E(x) + \eta_\varepsilon), \qquad x \in \Omega,$$

where $\eta_\varepsilon \in [0, \varphi_\varepsilon(1)]$ is chosen such that

$$\int_\Omega u_\varepsilon(x) \,\mathrm{d}x = \int_\Omega u(x) \,\mathrm{d}x.$$

Indeed, since

$$\psi_\varepsilon(\delta_E(x)) \le u(x) \le \psi_\varepsilon(\delta_E(x) + \varphi_\varepsilon(1)), \qquad x \in \Omega,$$

and ψ_ε is continuous, such an η_ε always exists. Figure 13.1 illustrates this *transition layer* construction. We will show in the following that u_ε is a recovery sequence for u.

Step 3. We first prove that $u_\varepsilon \to u$ in L^1. For this we recall **Federer's coarea formula**, which generalizes Fubini's theorem: Let $w \in \mathrm{L}^1(\Omega)$ and $h \colon \Omega \to \mathbb{R}$ be Lipschitz. Then,

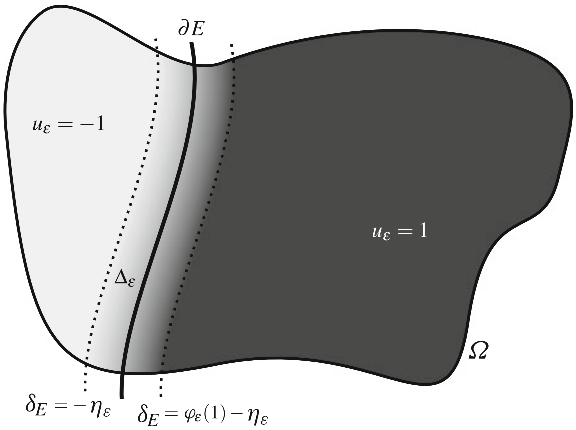


Fig. 13.1 The transition layer construction

$$\int_\Omega w(x)|\nabla h(x)|\,\mathrm{d}x = \int_{-\infty}^{+\infty} \int_{h^{-1}(t)} w(x)\,\mathrm{d}\mathscr{H}^{d-1}(x)\,\mathrm{d}t.$$

A proof of this result (in a more general form) can be found in Section 2.12 of [15].

In our situation, we first note that if $\delta(x) \leq -\eta_\varepsilon$ or $\delta_E(x) \geq \varphi_\varepsilon(1) - \eta_\varepsilon$ then $u_\varepsilon(x) = u(x)$. Define the *transition region* $\Delta_\varepsilon \subset \Omega$ via

$$\Delta_\varepsilon := \big\{ x \in \Omega \ : \ -\eta_\varepsilon \leq \delta_E(x) \leq \varphi_\varepsilon(1) - \eta_\varepsilon \big\}.$$

Using that $|\nabla \delta_E| = 1$ almost everywhere by Lemma 13.9, and Federer's coarea formula, we get

$$\begin{aligned}\int_\Omega |u_\varepsilon(x) - u(x)|\,\mathrm{d}x &= \int_{\Delta_\varepsilon} \big|\psi_\varepsilon(\delta_E(x) + \eta_\varepsilon) - u(x)\big|\,\mathrm{d}x \\ &\leq \int_{\Delta_\varepsilon} \big(|\psi_\varepsilon(\delta_E(x) + \eta_\varepsilon)| + 1\big) \cdot |\nabla \delta_E(x)|\,\mathrm{d}x \\ &= \int_{-\eta_\varepsilon}^{\varphi_\varepsilon(1)-\eta_\varepsilon} \big(|\psi_\varepsilon(t + \eta_\varepsilon)| + 1\big) \cdot \mathscr{H}^{d-1}(S_t \cap \Omega)\,\mathrm{d}t,\end{aligned}$$

where as in Lemma 13.9 we have set $S_t := \{ x \in \mathbb{R}^d \ : \ \delta_E(x) = t \}$. Then, with

$$g(s) := \sup_{t \in [-s,s]} \mathscr{H}^{d-1}(S_t \cap \Omega),$$

we can further estimate (also note $0 \leq \eta_\varepsilon \leq \varphi_\varepsilon(1)$ and $\varphi_\varepsilon(1) \leq 2\sqrt{\varepsilon}$)

$$\int_\Omega |u_\varepsilon(x) - u(x)|\,\mathrm{d}x \leq 2\varphi_\varepsilon(1) g(\varphi_\varepsilon(1)) \leq 4\sqrt{\varepsilon} g(2\sqrt{\varepsilon}).$$

On the other hand, by Lemma 13.9, $g(s) \to \mathscr{H}^{d-1}(\partial E \cap \Omega) < \infty$ as $s \to 0$. Thus, we conclude that $u_\varepsilon \to u$ as $\varepsilon \downarrow 0$.

Step 4. It remains to show that $\mathscr{F}_\varepsilon[u_\varepsilon] \to \mathscr{F}_0[u]$ as $\varepsilon \downarrow 0$. In fact, in light of the already established lim inf-inequality, it suffices to prove

$$\limsup_{\varepsilon \downarrow 0} \mathscr{F}_\varepsilon[u_\varepsilon] \le \mathscr{F}_0[u]. \tag{13.23}$$

Again using Federer's coarea formula, we get

$$\begin{aligned}
\mathscr{F}_\varepsilon[u_\varepsilon] &= \int_{\Delta_\varepsilon} \frac{1}{\varepsilon} f(\psi_\varepsilon(\delta_E(x) + \eta_\varepsilon)) + \varepsilon |\psi_\varepsilon'(\delta_E(x) + \eta_\varepsilon)|^2 \,\mathrm{d}x \\
&= \int_{-\eta_\varepsilon}^{\varphi_\varepsilon(1) - \eta_\varepsilon} \left[\frac{1}{\varepsilon} f(\psi_\varepsilon(t + \eta_\varepsilon)) + \varepsilon |\psi_\varepsilon'(t + \eta_\varepsilon)|^2 \right] \cdot \mathscr{H}^{d-1}(S_t \cap \Omega) \,\mathrm{d}t \\
&\le g(\varphi_\varepsilon(1)) \int_0^{\varphi_\varepsilon(1)} \frac{1}{\varepsilon} f(\psi_\varepsilon(t)) + \varepsilon |\psi_\varepsilon'(t)|^2 \,\mathrm{d}t.
\end{aligned}$$

We also compute

$$\psi_\varepsilon'(t) = \frac{1}{\varphi_\varepsilon'(\varphi_\varepsilon^{-1}(t))} = \frac{\sqrt{\varepsilon + f(\psi_\varepsilon(t))}}{\varepsilon} \qquad \text{for } 0 < t < \varphi_\varepsilon(1).$$

Then, continuing the above estimate,

$$\begin{aligned}
\mathscr{F}_\varepsilon[u_\varepsilon] &\le g(\varphi_\varepsilon(1)) \int_0^{\varphi_\varepsilon(1)} \frac{1}{\varepsilon} f(\psi_\varepsilon(t)) + \frac{\varepsilon + f(\psi_\varepsilon(t))}{\varepsilon} \,\mathrm{d}t \\
&\le 2g(\varphi_\varepsilon(1)) \int_0^{\varphi_\varepsilon(1)} \frac{\varepsilon + f(\psi_\varepsilon(t))}{\varepsilon} \,\mathrm{d}t \\
&= 2g(\varphi_\varepsilon(1)) \int_0^{\varphi_\varepsilon(1)} \sqrt{\varepsilon + f(\psi_\varepsilon(t))} \cdot \psi_\varepsilon'(t) \,\mathrm{d}t \\
&= 2g(\varphi_\varepsilon(1)) \int_{-1}^{1} \sqrt{\varepsilon + f(s)} \,\mathrm{d}s.
\end{aligned}$$

Taking the upper limit of the last expression as $\varepsilon \downarrow 0$ using Lemma 13.9, we arrive at

$$\limsup_{\varepsilon \downarrow 0} \mathscr{F}_\varepsilon[u_\varepsilon] \le 2\mathscr{H}^{d-1}(\partial E \cap \Omega) \int_{-1}^{1} \sqrt{f(s)} \,\mathrm{d}s = \sigma_0 \operatorname{Per}_\Omega(E) = \mathscr{F}_0[u].$$

Here we also employed the fact that $\mathscr{H}^{d-1}(\partial E \cap \Omega) = \operatorname{Per}_\Omega(E)$ since $\partial E \cap \Omega$ is smooth. Thus, (13.23) holds. □

Example 13.10. In our phase transition example from Section 1.9, we can now apply the Modica–Mortola Theorem 13.6 to see that the approximate functionals $\mathscr{F}_\varepsilon$ Γ-converge to $\mathscr{F}_0$ as $\varepsilon \downarrow 0$ with respect to L^1-convergence. We note that the constraint that $u(x) \in [\alpha, \beta]$ causes no problems since the recovery sequence constructed in the proof above satisfies this additional requirement anyway. In this sense, the $\mathscr{F}_\varepsilon$ approximate $\mathscr{F}_0$. We have thus shown that the regularized energy functionals converge to their sharp-interface limit as the strength of the regularization tends to zero.

13.3 Higher-Order Sharp-Interface Limits

In the theory of microstructure, which we presented in Chapters 8 and 9, we discussed how the differential inclusion

$$\begin{cases} u \in \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m), \\ \nabla u \in K \quad \text{in } \Omega \end{cases}$$

could be seen as an approximation to the minimization of the integral functional

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

if the set K is chosen as $K := \{ A \in \mathbb{R}^{m \times d} : f(A) = \min f \}$. In some situations of interest K consisted of at least two disjoint parts, usually either discrete points or $\mathrm{SO}(d)$-invariant wells. We saw in Chapter 9 that this could give rise to very complicated fine microstructure; in this context also note the Dolzmann–Müller Theorem 9.13.

If we want to incorporate higher-order energy contributions into our model, we should instead consider functionals of the (already normalized) form

$$\mathscr{F}_\varepsilon[u] := \begin{cases} \displaystyle\int_\Omega \frac{1}{\varepsilon} f(\nabla u(x)) + \varepsilon |\nabla^2 u(x)|^2 \,\mathrm{d}x & \text{if } u \in \mathrm{W}^{2,2}(\Omega; \mathbb{R}^m), \\ +\infty & \text{otherwise,} \end{cases}$$

where $f\colon \mathbb{R}^{m \times d} \to \mathbb{R}$ has multiple minimizing points or wells, and $\varepsilon > 0$ is a small parameter modeling the relative strength of the interfacial contribution. An important question is then whether these functionals Γ-converge to a limit as $\varepsilon \downarrow 0$. This question turns out to be quite difficult, essentially because the required curl-freeness of a recovery sequence and the vector-valued nature of the candidate maps necessitate more complicated constructions. We only quote the following two results:

Theorem 13.11 (Conti–Fonseca–Leoni 2002 [72]). *Let $\Omega \subset \mathbb{R}^d$ be a bounded and simply-connected Lipschitz domain. Assume that $f\colon \mathbb{R}^{m \times d} \to [0, \infty)$ is continuously differentiable, that is, $f(A) = 0$ if and only if $A \in \{A_1, A_2\}$, where $A_1, A_2 \in \mathbb{R}^{m \times d}$ with* $\mathrm{rank}(A_1 - A_2) = 1$, *and that $f(A) \to \infty$ as $|A| \to \infty$.*

Furthermore, suppose that the following technical condition holds:

$$f(A) \geq f(0|A_d),$$

where in the matrix $(0|A_d) \in \mathbb{R}^{m\times d}$ *the first* $(d-1)$ *columns of* $A = (A', A_d)$ *are replaced with zeros. Then, there is a* $\sigma > 0$ *such that the* $\mathscr{F}_\varepsilon$ Γ*-converge as* $\varepsilon \downarrow 0$ *with respect to the strong* $\mathrm{W}^{1,1}$*-topology to the functional*

$$\mathscr{F}_0[u] := \begin{cases} \sigma\,\mathrm{Per}_\Omega\big(\{x \in \Omega : \nabla u(x) = A_1\}\big) & \text{if } u \in \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m) \text{ and} \\ & \nabla u \in \mathrm{BV}(\Omega;\{A_1, A_2\}), \\ +\infty & \text{otherwise.} \end{cases}$$

Theorem 13.12 **(Conti–Schweizer 2006 [73]).** *Let* $\Omega \subset \mathbb{R}^2$ *be a bounded Lipschitz domain that is strictly star-shaped, i.e., there exists an* $x_0 \in \Omega$ *such that for all* $y \in \partial\Omega$ *the open segment from* x_0 *to* y *is contained in* Ω*. Assume that* $f : \mathbb{R}^{2\times 2} \to [0,\infty)$ *satisfies the following conditions:*

(i) $f(QA) = f(A)$ *for all* $A \in \mathbb{R}^{2\times 2}$, $Q \in \mathrm{SO}(2)$*;*
(ii) $f(A) = 0$ *if and only if* $A \in K := \mathrm{SO}(2)U_1 \cup \mathrm{SO}(2)U_2$ *for some* $U_1, U_2 \in \mathbb{R}^{2\times 2}$ *with* $\det U_1, \det U_2 > 0$ *and such that there exists a matrix* $Q \in \mathrm{SO}(2)$ *with* $\mathrm{rank}(U_1 - QU_2) = 1$*;*
(iii) f *satisfies the quadratic growth condition*

$$\mu\,\mathrm{dist}^2(A,K) \leq f(A) \leq M\,\mathrm{dist}^2(A,K), \qquad A \in \mathbb{R}^{2\times 2},$$

for some $\mu, M > 0$.

Then, there is a $\sigma > 0$ *such that the* $\mathscr{F}_\varepsilon$ Γ*-converge as* $\varepsilon \downarrow 0$ *with respect to the strong* L^1*-topology to the functional*

$$\mathscr{F}_0[u] := \begin{cases} \displaystyle\int_{J_{\nabla u}} g(n)\,\mathrm{d}\mathscr{H}^1 & \text{if } u \in \mathrm{W}^{1,1}(\Omega;\mathbb{R}^2) \text{ and } \nabla u \in \mathrm{BV}(\Omega;K), \\ +\infty & \text{otherwise.} \end{cases}$$

Here, $J_{\nabla u} \subset \Omega$ *is the jump set of* ∇u *with unit normal* $n\colon J_{\nabla u} \to \mathbb{S}^1$*, and*

$$g(n) := \inf\,\{\liminf_{n\to\infty} \mathscr{F}_{\varepsilon_n}[u_n; Q_n] : \varepsilon_n \to 0, u_n \to u_0^n \text{ in } \mathrm{L}^1\},$$

where Q_n *is the unit-volume square centered at the origin with two faces orthogonal to* $n \in \mathbb{S}^1$ *and*

$$u_0^n := \begin{cases} U_1 & \text{if } x \cdot n > 0, \\ QU_2 & \text{if } x \cdot n < 0, \end{cases}$$

with $Q \in \mathrm{SO}(2)$ *such that* $U_1 - QU_2 = a \otimes n$ *for some* $a \in \mathbb{R}^2$.

Note that by the Dolzmann–Müller Theorem 9.13 any u with $\mathscr{F}_0[u] < \infty$ is locally a simple laminate, so for all normals n to $J_{\nabla u}$ there exists a vector $a \in \mathbb{R}^2$ with $U_1 - QU_2 = a \otimes n$.

An analogous result for the 3D two-well problem is not known at present.

13.4 Periodic Homogenization

We now turn to the investigation of highly-oscillatory integrands. We define for all $\varepsilon > 0$ the functional

$$\mathscr{F}_\varepsilon[u] := \int_\Omega f\Big(\frac{x}{\varepsilon}, \nabla u(x)\Big)\,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m),$$

where $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain and $f : \mathbb{R}^d \times \mathbb{R}^{m\times d} \to [0, \infty)$ is a Carathéodory integrand that satisfies the standard p-growth and coercivity assumption

$$\mu|A|^p \le f(x, A) \le M(1 + |A|^p), \qquad (x, A) \in \mathbb{R}^d \times \mathbb{R}^{m\times d}, \tag{13.24}$$

for some $p \in (1, \infty)$ and $\mu, M > 0$, as well as the *periodicity condition*

$$x \mapsto f(x, A) \text{ is 1-periodic for all } A \in \mathbb{R}^{m\times d}.$$

Furthermore, we assume the following *local Lipschitz condition*:

$$|f(x, A) - f(x, B)| \le C(1 + |A|^{p-1} + |B|^{p-1})|A - B| \tag{13.25}$$

for all $x \in \Omega$, $A, B \in \mathbb{R}^{m\times d}$ and some constant $C > 0$. We remark that the last condition is in fact not needed if one follows a more involved proof, see [50, 52]. However, we showed in Lemma 5.6 that (13.25) holds if $f(x, \bullet)$ is quasiconvex for all x, which is necessary for weak lower semicontinuity by Proposition 5.18. So, unless we want to consider non-quasiconvex integrands, (13.25) is no restriction.

The main result of this section is the following *homogenization theorem*:

Theorem 13.13 (Braides 1985 & Müller 1987 [50, 199]). *The functionals $\mathscr{F}_\varepsilon$ Γ-converge as $\varepsilon \downarrow 0$ with respect to the weak $\mathrm{W}^{1,p}$-topology to the functional*

$$\mathscr{F}_0[u] := \int_\Omega f_{\mathrm{hom}}(\nabla u(x))\,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m),$$

*where $f_{\mathrm{hom}} : \mathbb{R}^{m\times d} \to [0, \infty)$ is the integrand given by the **asymptotic homogenization formula***

$$f_{\mathrm{hom}}(A) := \inf_{k\in\mathbb{N}} \inf_{\psi\in \mathrm{W}^{1,p}_{\mathrm{per}}((0,k)^d;\mathbb{R}^m)} \fint_{(0,k)^d} f(x, A + \nabla\psi(x))\,\mathrm{d}x, \qquad A \in \mathbb{R}^{m\times d}.$$

Moreover, in the definition of $\mathscr{F}_0$ the integrand f_{hom} may equivalently be replaced by

$$f_{\mathrm{hom},0}(A) := \inf_{k\in\mathbb{N}} \inf_{\psi\in \mathrm{W}^{1,p}_{0}((0,k)^d;\mathbb{R}^m)} \fint_{(0,k)^d} f(x, A + \nabla\psi(x))\,\mathrm{d}x, \qquad A \in \mathbb{R}^{m\times d}.$$

Here, $\mathrm{W}^{1,p}_{\mathrm{per}}((0,k)^d;\mathbb{R}^m)$ is the space of k-periodic functions in $\mathrm{W}^{1,p}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^m)$, namely the $\mathrm{W}^{1,p}_{\mathrm{loc}}$-closure of smooth k-periodic functions. Note that the formula for $f_{\mathrm{hom},0}$ does *not* simplify to the formula (7.1) for the quasiconvex envelope because the first argument of f is not held fixed. Indeed, it is intuitively evident that the optimal ψ needs to take into account how f varies in its first argument.

We will establish this theorem via several lemmas. For technical reasons we first consider $\mathscr{F}_0$ to be defined with $f_{\mathrm{hom},0}$ in place of f_{hom} and then show that $f_{\mathrm{hom},0} = f_{\mathrm{hom}}$. In the following we will also simply write "Fx" for the map $x \mapsto Fx$.

Lemma 13.14. *Let $F \in \mathbb{R}^{m\times d}$. Then, there exists a sequence $(u_\varepsilon) \subset \mathrm{W}^{1,p}_{Fx}(\Omega;\mathbb{R}^m)$ with $u_\varepsilon \rightharpoonup Fx$ in $\mathrm{W}^{1,p}$ as $\varepsilon \downarrow 0$ and*

$$\mathscr{F}_0[Fx] = \lim_{\varepsilon\downarrow 0} \mathscr{F}_\varepsilon[u_\varepsilon].$$

Proof. Fix $\delta > 0$ and choose $\psi_\delta \in \mathrm{W}^{1,p}_0((0,k)^d;\mathbb{R}^m)$ for some $k \in \mathbb{N}$ (depending on δ) such that

$$f_{\mathrm{hom},0}(F) \le \fint_{(0,k)^d} f(y, F + \nabla\psi_\delta(y))\,\mathrm{d}y \le f_{\mathrm{hom},0}(F) + \delta. \tag{13.26}$$

For $\eta > 0$ we denote by $\Omega^\eta \subset \Omega$ the set of all cubes from the regular lattice of open cubes $(0,\eta)^d + \eta\mathbb{Z}^d$ that are contained in Ω. Consider ψ_δ to be extended to all of $\mathbb{R}^d$ by periodicity and define

$$u_{\delta,\varepsilon}(x) := \begin{cases} Fx + \varepsilon\psi_\delta\Big(\dfrac{x}{\varepsilon}\Big) & \text{if } x \in \Omega^{\varepsilon k},\\ Fx & \text{if } x \in \Omega\setminus\Omega^{\varepsilon k}. \end{cases}$$

We have that $u_{\delta,\varepsilon} \in \mathrm{W}^{1,p}_{Fx}(\Omega;\mathbb{R}^m)$ and $u_{\delta,\varepsilon} \to Fx$ in L^p as $\varepsilon \downarrow 0$.

For every cube $Q \in (0,\varepsilon k)^d + \varepsilon k\mathbb{Z}^d$ that is contained in $\Omega^{\varepsilon k}$ we have

$$\begin{aligned}\fint_Q f\Big(\frac{x}{\varepsilon}, \nabla u_{\delta,\varepsilon}(x)\Big)\,\mathrm{d}x &= \fint_Q f\Big(\frac{x}{\varepsilon}, F + \nabla\psi_\delta\Big(\frac{x}{\varepsilon}\Big)\Big)\,\mathrm{d}x\\ &= \fint_{(0,k)^d} f(y, F + \nabla\psi_\delta(y))\,\mathrm{d}y.\end{aligned}$$

Combining this with (13.26) and summing over all cubes Q contained in $\Omega^{\varepsilon k}$, we get

$$|\Omega^{\varepsilon k}|f_{\mathrm{hom},0}(F) \le \int_{\Omega^{\varepsilon k}} f\Big(\frac{x}{\varepsilon}, \nabla u_{\delta,\varepsilon}(x)\Big)\,\mathrm{d}x \le |\Omega^{\varepsilon k}|\big(f_{\mathrm{hom},0}(F) + \delta\big).$$

We now let $\varepsilon \downarrow 0$ and use the growth bound on f together with $|\Omega \setminus \Omega^{\varepsilon k}| \to 0$ to conclude that

$$\mathscr{F}_0[Fx] \le \liminf_{\varepsilon\downarrow 0} \mathscr{F}_\varepsilon[u_{\delta,\varepsilon}] \le \limsup_{\varepsilon\downarrow 0} \mathscr{F}_\varepsilon[u_{\delta,\varepsilon}] \le \mathscr{F}_0[Fx] + \delta|\Omega|.$$

Thus, we have

$$\lim_{\delta\downarrow 0}\lim_{\varepsilon\downarrow 0}\Big[\big|\mathscr{F}_\varepsilon[u_{\delta,\varepsilon}] - \mathscr{F}_0[Fx]\big| + \|u_{\delta,\varepsilon} - Fx\|_{\mathrm{L}^p}\Big] = 0.$$

Using a diagonal sequence, see Problem 13.6, we can construct a function $\delta\colon (0,\infty) \to (0,\infty)$ such that for $u_\varepsilon := u_{\delta(\varepsilon),\varepsilon}$ it holds that

$$\mathscr{F}_\varepsilon[u_\varepsilon] \to \mathscr{F}_0[Fx] \qquad \text{and} \qquad u_\varepsilon \to Fx \quad \text{in } \mathrm{L}^p.$$

Since by the lower bound in (13.24), we have that (∇u_ε) is uniformly bounded in $\mathrm{L}^p(\Omega;\mathbb{R}^{m\times d})$, we may assume that also $u_\varepsilon \rightharpoonup Fx$ in $\mathrm{W}^{1,p}$ (the weak limit is already determined by $u_\varepsilon \to Fx$ in L^p). □

Lemma 13.15. *Let $(u_\varepsilon) \subset \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ with $u_\varepsilon \rightharpoonup Fx$ in $\mathrm{W}^{1,p}$ as $\varepsilon \downarrow 0$, where $F \in \mathbb{R}^{m\times d}$. Then,*

$$\mathscr{F}_0[Fx] \le \liminf_{\varepsilon\downarrow 0} \mathscr{F}_\varepsilon[u_\varepsilon].$$

Proof. Step 1. Assume first that $\Omega = Q$ is an open cube with all edges parallel to the coordinate axes and side length $s > 0$. Assume furthermore that the u_ε all have the same linear boundary values as the limit, i.e., $(u_\varepsilon) \subset \mathrm{W}^{1,p}_{Fx}(Q;\mathbb{R}^m)$ with $u_\varepsilon \rightharpoonup Fx$.

For fixed $\varepsilon > 0$ let $k \in \mathbb{N}$ be the smallest natural number such that $\varepsilon k \ge s + \varepsilon$ and denote by $Q^\varepsilon \supset Q$ an open cube with side length $k\varepsilon$ and such that all vertices of Q^ε lie in $\varepsilon\mathbb{Z}^d$, i.e., $Q^\varepsilon = \varepsilon z_0 + (0,\varepsilon k)^d$ for some $z_0 \in \mathbb{Z}^d$. Then extend u_ε continuously to Q^ε by setting

$$u_\varepsilon(x) := Fx \qquad \text{for } x \in Q^\varepsilon \setminus Q.$$

Hence,

$$\int_Q f\Big(\frac{x}{\varepsilon}, \nabla u_\varepsilon(x)\Big)\,\mathrm{d}x = \int_{Q^\varepsilon} f\Big(\frac{x}{\varepsilon}, \nabla u_\varepsilon(x)\Big)\,\mathrm{d}x - \int_{Q^\varepsilon\setminus Q} f\Big(\frac{x}{\varepsilon}, F\Big)\,\mathrm{d}x.$$

Since the second term vanishes as $\varepsilon \downarrow 0$ by the growth bounds on f and $|Q^\varepsilon \setminus Q| \le (s+2\varepsilon)^d - s^d \to 0$, we get with the change of variables $x = \varepsilon(z_0 + y)$ and with

$$\psi_\varepsilon(x) := u_\varepsilon(x) - Fx \in \mathrm{W}^{1,p}_0(Q^\varepsilon;\mathbb{R}^m),$$
$$\widetilde{\psi}_\varepsilon(y) := \frac{\psi_\varepsilon(\varepsilon(z_0+y))}{\varepsilon} \in \mathrm{W}^{1,p}_0((0,k)^d;\mathbb{R}^m)$$

that

$$\begin{aligned}\liminf_{\varepsilon\downarrow 0} \int_Q f\Big(\frac{x}{\varepsilon}, \nabla u_\varepsilon(x)\Big)\,\mathrm{d}x &\geq \liminf_{\varepsilon\downarrow 0} \int_{Q^\varepsilon} f\Big(\frac{x}{\varepsilon}, \nabla u_\varepsilon(x)\Big)\,\mathrm{d}x \\ &= \liminf_{\varepsilon\downarrow 0}\, (\varepsilon k)^d \fint_{(0,k)^d} f\big(y, F + \nabla\psi_\varepsilon(\varepsilon(z_0+y))\big)\,\mathrm{d}y \\ &= \liminf_{\varepsilon\downarrow 0}\, (\varepsilon k)^d \fint_{(0,k)^d} f(y, F + \nabla\widetilde{\psi}_\varepsilon(y))\,\mathrm{d}y \\ &\geq \liminf_{\varepsilon\downarrow 0}\, (\varepsilon k)^d f_{\mathrm{hom},0}(F) \\ &\geq |Q|\, f_{\mathrm{hom},0}(F).\end{aligned}$$

Here, we also used the definition of $f_{\mathrm{hom},0}$ and $\varepsilon k \geq s$. This proves the claim if $\Omega = Q$ and $(u_\varepsilon) \subset \mathrm{W}^{1,p}_{Fx}(Q;\mathbb{R}^m)$.

Step 2. We now assume that Ω is an arbitrary bounded Lipschitz domain, but we still require that the u_ε have linear boundary values, that is, $(u_\varepsilon) \subset \mathrm{W}^{1,p}_{Fx}(\Omega;\mathbb{R}^m)$ with $u_\varepsilon \rightharpoonup Fx$ in $\mathrm{W}^{1,p}$. Let $Q \Supset \Omega$ be a cube with all edges parallel to the coordinate axes. By the previous lemma applied in the domain $Q \setminus \overline{\Omega}$ there exists a sequence $(v_\varepsilon) \subset \mathrm{W}^{1,p}_{Fx}(Q \setminus \overline{\Omega};\mathbb{R}^m)$ with $v_\varepsilon \rightharpoonup Fx$ in $\mathrm{W}^{1,p}$ and

$$\int_{Q\setminus\overline{\Omega}} f\Big(\frac{x}{\varepsilon}, \nabla v_\varepsilon(x)\Big)\,\mathrm{d}x \to |Q \setminus \overline{\Omega}|\, f_{\mathrm{hom},0}(F) \qquad \text{as } \varepsilon \downarrow 0. \tag{13.27}$$

Define for $\varepsilon > 0$,

$$w_\varepsilon(x) := \begin{cases} u_\varepsilon(x) & \text{if } x \in \Omega, \\ v_\varepsilon(x) & \text{if } x \in Q \setminus \overline{\Omega}. \end{cases}$$

These maps lie in $\mathrm{W}^{1,p}(Q;\mathbb{R}^m)$ since the boundary values agree over the gluing boundary. By Step 1 we have that

$$|Q|\, f_{\mathrm{hom},0}(F) \leq \liminf_{\varepsilon\downarrow 0} \int_Q f\Big(\frac{x}{\varepsilon}, \nabla w_\varepsilon(x)\Big)\,\mathrm{d}x.$$

Hence, combining this with (13.27), we arrive at

$$\mathscr{F}_0[Fx] \leq \liminf_{\varepsilon\downarrow 0} \mathscr{F}_\varepsilon[u_\varepsilon].$$

Step 3. Finally, we remove the restriction on the boundary values via a cut-off procedure. So, suppose that $(u_\varepsilon) \subset \mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$ with $u_\varepsilon \rightharpoonup Fx$ in $\mathrm{W}^{1,p}$. Then let $n \in \mathbb{N}$ and let $\Omega_0 \Subset \Omega$ be Lipschitz subdomain. With $R := \mathrm{dist}(\Omega_0, \partial\Omega)$ set

$$\Omega_i := \left\{ x \in \Omega \;:\; \mathrm{dist}(x, \Omega_0) < \frac{iR}{n} \right\} \qquad \text{for } i = 1, \ldots, n.$$

Next, choose cut-off functions $\rho_i \in \mathrm{C}_0^\infty(\Omega; [0, 1])$ such that

$$\mathbb{1}_{\Omega_{i-1}} \le \rho_i \le \mathbb{1}_{\Omega_i}, \qquad |\nabla \rho_i| \le \frac{2n}{R}$$

and set

$$u_\varepsilon^{(i)}(x) := Fx + \rho_i(x)(u_\varepsilon(x) - Fx), \qquad x \in \Omega,$$

for which

$$\nabla u_\varepsilon^{(i)}(x) = F + \rho_i(x)(\nabla u_\varepsilon(x) - F) + (u_\varepsilon(x) - Fx) \otimes \nabla \rho_i(x).$$

From the growth bound on f we get, with an i, n-independent constant $C > 0$,

$$\begin{aligned}
\mathscr{F}_\varepsilon[u_\varepsilon^{(i)}] &= \int_{\Omega_{i-1}} f\Big(\frac{x}{\varepsilon}, \nabla u_\varepsilon(x)\Big) \,\mathrm{d}x \\
&\quad + \int_{\Omega_i \setminus \Omega_{i-1}} f\Big(\frac{x}{\varepsilon}, \nabla u_\varepsilon^{(i)}(x)\Big) \,\mathrm{d}x + \int_{\Omega \setminus \Omega_i} f\Big(\frac{x}{\varepsilon}, F\Big) \,\mathrm{d}x \\
&\le \int_\Omega f\Big(\frac{x}{\varepsilon}, \nabla u_\varepsilon(x)\Big) \,\mathrm{d}x \\
&\quad + C \int_{\Omega_i \setminus \Omega_{i-1}} 1 + |F|^p + |\nabla u_\varepsilon(x) - F|^p + \Big(\frac{2n}{R}\Big)^p |u_\varepsilon(x) - Fx|^p \,\mathrm{d}x \\
&\quad + C(1 + |F|^p)|\Omega \setminus \Omega_i| \\
&\le \mathscr{F}_\varepsilon[u_\varepsilon] + C \int_{\Omega_i \setminus \Omega_{i-1}} |\nabla u_\varepsilon(x) - F|^p \,\mathrm{d}x + C\Big(\frac{2n}{R}\Big)^p \|u_\varepsilon - Fx\|_{\mathrm{L}^p}^p \\
&\quad + C|\Omega \setminus \Omega_0|.
\end{aligned}$$

On the other hand, $u_\varepsilon^{(i)} \rightharpoonup Fx$ in $\mathrm{W}^{1,p}$ and so, by Step 2,

$$\begin{aligned}
\mathscr{F}_0[Fx] &\le \liminf_{\varepsilon \downarrow 0} \mathscr{F}_\varepsilon[u_\varepsilon^{(i)}] \\
&\le \liminf_{\varepsilon \downarrow 0} \left[\mathscr{F}_\varepsilon[u_\varepsilon] + C \int_{\Omega_i \setminus \Omega_{i-1}} |\nabla u_\varepsilon - F|^p \,\mathrm{d}x \right] + C|\Omega \setminus \Omega_0|.
\end{aligned}$$

Now sum this over $i = 1, \ldots, n$, use the superadditivity of the lower limit (that is, $\liminf_{j\to\infty} a_j + \liminf_{j\to\infty} b_j \le \liminf_{j\to\infty}(a_j + b_j)$), and divide by n to see that

$$\mathscr{F}_0[Fx] \le \liminf_{\varepsilon \downarrow 0} \mathscr{F}_\varepsilon[u_\varepsilon] + \frac{C}{n} \cdot \limsup_{\varepsilon \downarrow 0} \int_\Omega |\nabla u_\varepsilon - F|^p \,\mathrm{d}x + C|\Omega \setminus \Omega_0|.$$

Then we may let $n \to \infty$ and $\Omega_0 \uparrow \Omega$, that is, $|\Omega \setminus \Omega_0| \to 0$, to conclude. □

Lemma 13.16. *The integrand* $f_{\mathrm{hom},0}\colon \mathbb{R}^{m\times d} \to [0,\infty)$ *from Theorem 13.13 satisfies the growth condition* (13.24) *and the local Lipschitz condition* (13.25).

Proof. Step 1. Since $\psi = 0$ is admissible in the definition of $f_{\mathrm{hom},0}$, we immediately have $f_{\mathrm{hom},0}(x, A) \le M(1+|A|^p)$. For the lower bound at $A \in \mathbb{R}^{m\times d}$, fix $\delta > 0$ and choose $\psi_\delta \in \mathrm{W}^{1,p}_0((0,k)^d;\mathbb{R}^m)$ for some $k \in \mathbb{N}$ (depending on δ) such that

$$f_{\mathrm{hom},0}(A) + \delta \ge \fint_{(0,k)^d} f(y, A + \nabla\psi_\delta(y))\,\mathrm{d}y.$$

Then we can estimate, using (13.24) for f and Jensen's inequality,

$$f_{\mathrm{hom},0}(A) + \delta \ge \mu \fint_{(0,k)^d} |A + \nabla\psi_\delta(y)|^p\,\mathrm{d}y \ge \mu|A|^p,$$

which for $\delta \downarrow 0$ establishes the lower bound in (13.24) for $f_{\mathrm{hom},0}$.

Step 2. To show the local Lipschitz condition (13.25) for $f_{\mathrm{hom},0}$, we fix $A, B \in \mathbb{R}^{m\times d}$ and use Lemma 13.14 on the domain $(0,1)^d$ to get a (recovery) sequence $(u_\varepsilon) \subset \mathrm{W}^{1,p}((0,1)^d;\mathbb{R}^m)$ with $u_\varepsilon \rightharpoonup Ax$ in $\mathrm{W}^{1,p}$ and

$$f_{\mathrm{hom},0}(A) = \int_{(0,1)^d} f_{\mathrm{hom},0}(A)\,\mathrm{d}x = \lim_{\varepsilon\downarrow 0} \int_{(0,1)^d} f\Big(\frac{x}{\varepsilon}, \nabla u_\varepsilon(x)\Big)\,\mathrm{d}x.$$

From (13.24) for $f_{\mathrm{hom},0}$ (proved in Step 1) we infer that

$$\limsup_{\varepsilon\downarrow 0} \|\nabla u_\varepsilon\|^p_{\mathrm{L}^p} \le \frac{1}{\mu} f_{\mathrm{hom},0}(A) \le \frac{M}{\mu}(1+|A|^p). \tag{13.28}$$

Define

$$v_\varepsilon(x) := (B-A)x + u_\varepsilon(x), \qquad x \in (0,1)^d.$$

It is straightforward to see that

$$\limsup_{\varepsilon\downarrow 0} \|\nabla v_\varepsilon\|^p_{\mathrm{L}^p} \le C\big(\|\nabla u_\varepsilon\|^p_{\mathrm{L}^p} + |A|^p + |B|^p\big) \le C(1+|A|^p+|B|^p). \tag{13.29}$$

Furthermore, since $v_\varepsilon \rightharpoonup Bx$ in $\mathrm{W}^{1,p}$, Lemma 13.15 implies

$$f_{\mathrm{hom},0}(B) = \int_{(0,1)^d} f_{\mathrm{hom},0}(B)\,\mathrm{d}x \le \liminf_{\varepsilon\downarrow 0} \int_{(0,1)^d} f\Big(\frac{x}{\varepsilon}, \nabla v_\varepsilon(x)\Big)\,\mathrm{d}x.$$

Thus, also using the local Lipschitz continuity (13.25) for f,

$$
\begin{aligned}
&f_{\mathrm{hom},0}(B) - f_{\mathrm{hom},0}(A) \\
&\quad \le \liminf_{\varepsilon\downarrow 0} \int_{(0,1)^d} f\Bigl(\frac{x}{\varepsilon}, \nabla v_\varepsilon(x)\Bigr) - f\Bigl(\frac{x}{\varepsilon}, \nabla u_\varepsilon(x)\Bigr) \,\mathrm{d}x \\
&\quad \le C \cdot \liminf_{\varepsilon\downarrow 0} \int_{(0,1)^d} (1 + |\nabla u_\varepsilon(x)|^{p-1} + |\nabla v_\varepsilon(x)|^{p-1}) \cdot |\nabla u_\varepsilon(x) - \nabla v_\varepsilon(x)| \,\mathrm{d}x \\
&\quad \le C \cdot \limsup_{\varepsilon\downarrow 0} \bigl(1 + \|\nabla u_\varepsilon\|_{\mathrm{L}^p}^p + \|\nabla v_\varepsilon\|_{\mathrm{L}^p}^p\bigr)^{(p-1)/p} \cdot |A - B| \\
&\quad \le C(1 + |A|^{p-1} + |B|^{p-1})|A - B|,
\end{aligned}
$$

where we also used (13.28), (13.29) and Hölder's inequality. Interchanging the roles of A and B, we have thus shown (13.25) for $f_{\mathrm{hom},0}$. □

Lemma 13.17. *It holds that* $f_{\mathrm{hom}} = f_{\mathrm{hom},0}$.

Proof. Clearly, $f_{\mathrm{hom}} \le f_{\mathrm{hom},0}$ since $\mathrm{W}_0^{1,p}((0,k)^d;\mathbb{R}^m)$ is contained in the space $\mathrm{W}_{\mathrm{per}}^{1,p}((0,k)^d;\mathbb{R}^m)$.

To see the other inequality, let $A \in \mathbb{R}^{m\times d}$, $\psi \in \mathrm{W}_{\mathrm{per}}^{1,p}((0,k)^d;\mathbb{R}^m)$ and define

$$
u_\varepsilon(x) := Ax + \varepsilon\psi\Bigl(\frac{x}{\varepsilon}\Bigr), \qquad x \in (0,k)^d.
$$

Since $u_\varepsilon \rightharpoonup Ax$ in $\mathrm{W}^{1,p}$, we can apply Lemma 13.15 on the domain $(0,1)^d$ to the left-hand side to deduce that

$$
\begin{aligned}
f_{\mathrm{hom},0}(A) &\le \liminf_{\varepsilon\downarrow 0} \int_{(0,1)^d} f\Bigl(\frac{x}{\varepsilon}, \nabla u_\varepsilon(x)\Bigr) \,\mathrm{d}x \\
&= \liminf_{\varepsilon\downarrow 0} \fint_{(0,\varepsilon^{-1})^d} f(y, A + \nabla\psi(y)) \,\mathrm{d}y \\
&= \fint_{(0,k)^d} f(y, A + \nabla\psi(y)) \,\mathrm{d}y.
\end{aligned}
$$

Here, for the second to last equality we utilized the oscillatory nature of the function $f(x/\varepsilon, \nabla u_\varepsilon(x))$, see Problem 13.5. Taking the infimum over all $\psi \in \mathrm{W}_{\mathrm{per}}^{1,p}((0,k)^d;\mathbb{R}^m)$, we arrive at

$$
f_{\mathrm{hom},0}(A) \le f_{\mathrm{hom}}(A),
$$

which is the claim. □

Proof of Theorem 13.13. *Step 1: Liminf-inequality.* Let $u_\varepsilon \rightharpoonup u$ in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$. Given $\delta > 0$, it can be seen that there exists a finite partition of Ω into disjoint open sets $\Omega_i \subset \Omega$, $i = 1, \ldots, N$, and a negligible set $Z \subset \Omega$, that is,

$$
\Omega = Z \cup \bigcup_{i=1}^N \Omega_i, \qquad |Z| = 0,
$$

such that

$$\sum_{i=1}^{N} \int_{\Omega_i} |\nabla u(x) - A_i|^p \,\mathrm{d}x < \delta^p, \tag{13.30}$$

where

$$A_i := [\nabla u]_{\Omega_i} = \fint_{\Omega_i} \nabla u \,\mathrm{d}x.$$

Then set

$$v_\varepsilon(x) := A_i x + u_\varepsilon(x) - u(x) \qquad \text{if } x \in \Omega_i \ (i = 1, \ldots, N).$$

It is obvious that $v_\varepsilon \rightharpoonup A_i x$ in $\mathrm{W}^{1,p}(\Omega_i; \mathbb{R}^m)$.

Below we will show the estimates

$$\left| \mathscr{F}_\varepsilon[u_\varepsilon] - \sum_{i=1}^{N} \int_{\Omega_i} f\Bigl(\frac{x}{\varepsilon}, \nabla v_\varepsilon(x)\Bigr) \,\mathrm{d}x \right| \le C\delta, \tag{13.31}$$

$$\left| \mathscr{F}_0[u] - \sum_{i=1}^{N} \int_{\Omega_i} f_{\mathrm{hom}}(A_i) \,\mathrm{d}x \right| \le C\delta \tag{13.32}$$

for some constant $C > 0$ that does not depend on ε or δ. From Lemma 13.15 (extended to affine maps, which is trivial) in conjunction with Lemma 13.17, we get

$$\sum_{i=1}^{N} \int_{\Omega_i} f_{\mathrm{hom}}(A_i) \,\mathrm{d}x \le \liminf_{\varepsilon \downarrow 0} \sum_{i=1}^{N} \int_{\Omega_i} f\Bigl(\frac{x}{\varepsilon}, \nabla v_\varepsilon(x)\Bigr) \,\mathrm{d}x.$$

Thus, combining this with (13.31) and (13.32),

$$\mathscr{F}_0[u] \le \liminf_{\varepsilon \downarrow 0} \mathscr{F}_\varepsilon[u_\varepsilon] + C\delta.$$

As $\delta > 0$ was arbitrary, we arrive at the lim inf-inequality.

It remains to show (13.31) and (13.32). For the first assertion we estimate, using the local Lipschitz continuity (13.25) of f and Hölder's inequality for sums, that

$$\begin{aligned}
&\left| \mathscr{F}_\varepsilon[u_\varepsilon] - \sum_{i=1}^{N} \int_{\Omega_i} f\Bigl(\frac{x}{\varepsilon}, \nabla v_\varepsilon(x)\Bigr) \,\mathrm{d}x \right| \\
&\quad \le \sum_{i=1}^{N} \int_{\Omega_i} \left| f\Bigl(\frac{x}{\varepsilon}, \nabla u_\varepsilon(x)\Bigr) - f\Bigl(\frac{x}{\varepsilon}, \nabla v_\varepsilon(x)\Bigr) \right| \,\mathrm{d}x \\
&\quad \le C \sum_{i=1}^{N} \int_{\Omega_i} \bigl(1 + |\nabla u_\varepsilon(x)|^{p-1} + |\nabla v_\varepsilon(x)|^{p-1}\bigr) |\nabla u_\varepsilon(x) - \nabla v_\varepsilon(x)| \,\mathrm{d}x
\end{aligned}$$

$$\leq C \sum_{i=1}^{N} \left[\big\| 1 + |\nabla u_\varepsilon| + |\nabla v_\varepsilon| \big\|_{\mathrm{L}^p(\Omega_i)}^{p-1} \cdot \left(\int_{\Omega_i} |\nabla u(x) - A_i|^p \,\mathrm{d}x \right)^{1/p} \right]$$

$$\leq C \big\| 1 + |\nabla u_\varepsilon| + |\nabla v_\varepsilon| \big\|_{\mathrm{L}^p(\Omega)}^{p-1} \cdot \left(\sum_{i=1}^{N} \int_{\Omega_i} |\nabla u(x) - A_i|^p \,\mathrm{d}x \right)^{1/p}$$

$$\leq C\delta,$$

where for the last estimate we also employed (13.30). The second assertion is shown in a similar fashion using Lemma 13.16.

Step 2: Recovery sequence. Let $u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$. Then, for every $\delta > 0$ there exists a $u_\delta \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ that is countably piecewise affine and such that $\|u - u_\delta\|_{\mathrm{W}^{1,p}} < \delta$ (see Theorem A.29). By the same estimate (based on Lemma 13.16) as above, we get

$$\begin{aligned} \big| \mathscr{F}_0[u_\delta] - \mathscr{F}_0[u] \big| &\leq C \big(1 + \|\nabla u_\delta\|_{\mathrm{L}^p}^p + \|\nabla u\|_{\mathrm{L}^p}^p \big)^{(p-1)/p} \|\nabla u_\delta - \nabla u\|_{\mathrm{L}^p} \\ &\to 0 \qquad \text{as } \delta \downarrow 0. \end{aligned}$$

Suppose that $\Omega = Z \cup \bigcup_i \Omega_i$ is a decomposition of Ω, up to the negligible set Z, into disjoint open patches Ω_i ($i \in \mathbb{N}$), on every one of which u_δ is affine, say $\nabla u_\delta(x) = A_i \in \mathbb{R}^{m \times d}$ for $x \in \Omega_i$. We remark that the Ω_i of course depend on δ, but we suppress this in our notation, which should not cause any confusion. Now apply Lemma 13.14 (extended to affine maps) in conjunction with Lemma 13.17 in every Ω_i separately to get sequences $(u^{(i)}_{\delta,\varepsilon}) \subset \mathrm{W}^{1,p}(\Omega_i; \mathbb{R}^m)$ with $u^{(i)}_{\delta,\varepsilon}|_{\partial\Omega_i} = u_\delta|_{\partial\Omega_i}$ and $u^{(i)}_{\delta,\varepsilon} \rightharpoonup u_\delta$ in $\mathrm{W}^{1,p}(\Omega_i; \mathbb{R}^m)$ as $\varepsilon \downarrow 0$ (δ held fixed) as well as

$$\lim_{\varepsilon \downarrow 0} \int_{\Omega_i} f\Big(\frac{x}{\varepsilon}, \nabla u^{(i)}_{\delta,\varepsilon}(x) \Big) \,\mathrm{d}x = |\Omega_i| f_{\mathrm{hom}}(A_i).$$

Then set

$$u_{\delta,\varepsilon}(x) := u^{(i)}_{\delta,\varepsilon}(x) \qquad \text{if } x \in \Omega_i \ (i \in \mathbb{N}).$$

It holds that $u_{\delta,\varepsilon} \rightharpoonup u_\delta$ in $\mathrm{W}^{1,p}$ as $\varepsilon \downarrow 0$ and

$$\lim_{\varepsilon \downarrow 0} \int_{\Omega} f\Big(\frac{x}{\varepsilon}, \nabla u_{\delta,\varepsilon}(x) \Big) \,\mathrm{d}x = \int_{\Omega} f_{\mathrm{hom}}(\nabla u_\delta(x)) \,\mathrm{d}x. \tag{13.33}$$

For

$$v_{\delta,\varepsilon} := u_{\delta,\varepsilon} + u - u_\delta$$

we have $v_{\delta,\varepsilon} \rightharpoonup u$ in $\mathrm{W}^{1,p}$ as $\varepsilon \downarrow 0$ and δ held fixed.

From the Lipschitz assumption (13.25) we may derive the equicontinuity property

$$\begin{aligned} \big| \mathscr{F}_\varepsilon[u_{\delta,\varepsilon}] - \mathscr{F}_\varepsilon[v_{\delta,\varepsilon}] \big| &\leq C \big(1 + \|\nabla u_{\delta,\varepsilon}\|_{\mathrm{L}^p}^p + \|\nabla u_{\delta,\varepsilon}\|_{\mathrm{L}^p}^p \big)^{(p-1)/p} \|\nabla u_{\delta,\varepsilon} - \nabla v_{\delta,\varepsilon}\|_{\mathrm{L}^p} \\ &\leq C \|\nabla u - \nabla u_\delta\|_{\mathrm{L}^p}. \end{aligned}$$

Thus, also using (13.33),

$$\limsup_{\varepsilon\downarrow 0} \mathscr{F}_\varepsilon[v_{\delta,\varepsilon}] \le \limsup_{\varepsilon\downarrow 0} \mathscr{F}_\varepsilon[u_{\delta,\varepsilon}] + C\delta \le \mathscr{F}_0[u_\delta] + C\delta.$$

In a similar fashion (with Lemma 13.16 replacing (13.25)) one also gets

$$\mathscr{F}_0[u_\delta] - C\delta \le \liminf_{\varepsilon\downarrow 0} \mathscr{F}_\varepsilon[v_{\delta,\varepsilon}].$$

Consequently,

$$\lim_{\delta\downarrow 0}\lim_{\varepsilon\downarrow 0} \mathscr{F}_\varepsilon[v_{\delta,\varepsilon}] = \mathscr{F}_0[u].$$

Thus we can finish the proof by the diagonal argument from Problem 13.6, just like in the proof of Lemma 13.14. □

13.5 Convex Homogenization

In the convex case, the homogenization formula from Theorem 13.13 simplifies and we have the following theorem.

Theorem 13.18 (**Marcellini 1978 [179]**). *In the situation of Theorem 13.13, if additionally $A \mapsto f(x, A)$ is assumed to be convex for almost every $x \in \Omega$, then the asymptotic homogenization formula simplifies to the* ***cell problem formula***

$$f_{\mathrm{hom}}(A) = \inf_{\varphi\in \mathrm{W}^{1,p}_{\mathrm{per}}((0,1)^d;\mathbb{R}^m)} \int_{(0,1)^d} f(x, A + \nabla\varphi(x))\,\mathrm{d}x, \qquad A \in \mathbb{R}^{m\times d}.$$

Remark 13.19. This result for the convex case also holds for more general upper growth assumptions on f than the one in (13.24), but then the proof becomes more involved, see [199]. Moreover, for the scalar case $m = 1$, the convexity condition is not necessary, see Problem 13.9.

Proof. Let

$$\hat{f}_{\mathrm{hom}}(A) := \inf_{\varphi\in \mathrm{W}^{1,p}_{\mathrm{per}}((0,1)^d;\mathbb{R}^m)} \int_{(0,1)^d} f(x, A + \nabla\varphi(x))\,\mathrm{d}x, \qquad A \in \mathbb{R}^{m\times d}.$$

We will show in the following that

$$\inf_{\psi\in \mathrm{W}^{1,p}_{\mathrm{per}}((0,k)^d;\mathbb{R}^m)} \fint_{(0,k)^d} f(x, A + \nabla\psi(x))\,\mathrm{d}x \ge \hat{f}_{\mathrm{hom}}(A) \tag{13.34}$$

for every $A \in \mathbb{R}^{m\times d}$ and every $k \in \mathbb{N}$. From this the conclusion follows immediately since $f_{\mathrm{hom}} \le \hat{f}_{\mathrm{hom}}$ is trivially true.

By Problem 13.8 we may assume that f is smooth in its second argument. Since the values of both $f_{\hom}$ and $\hat{f}_{\hom}$ are defined via minimization problems with convex, coercive integrands, by Theorem 2.7 there exist minimizers $\psi_* \in \mathrm{W}^{1,p}_{\mathrm{per}}((0,k)^d;\mathbb{R}^m)$ and $\varphi_* \in \mathrm{W}^{1,p}_{\mathrm{per}}((0,1)^d;\mathbb{R}^m)$, respectively. Moreover, by a slight modification of Theorem 3.1 for periodic spaces of candidate functions, we have the Euler–Lagrange equation

$$\int_{(0,1)^d} \mathrm{D}_A f(x, A + \nabla\varphi_*(x)) : \nabla w(x)\,\mathrm{d}x = 0 \qquad \text{for all } w \in \mathrm{W}^{1,p}_{\mathrm{per}}((0,1)^d;\mathbb{R}^m),$$

see Problem 13.7. From the convexity of f we then get

$$\begin{aligned}\int_{(0,k)^d} & f(x, A + \nabla\psi_*(x)) - f(x, A + \nabla\varphi_*(x))\,\mathrm{d}x \\ &\geq \int_{(0,k)^d} \mathrm{D}_A f(x, A + \nabla\varphi_*(x)) : [\nabla\psi_*(x) - \nabla\varphi_*(x)]\,\mathrm{d}x.\end{aligned}$$

Set $v := \psi_* - \varphi_*$ and write the right-hand side as follows, using the 1-periodicity of φ_* and of f with respect to the first argument,

$$\begin{aligned}\int_{(0,k)^d} & \mathrm{D}_A f(x, A + \nabla\varphi_*(x)) : \nabla v(x)\,\mathrm{d}x \\ &= \sum_{z\in\{0,\dots,k-1\}^d} \int_{(0,1)^d} \mathrm{D}_A f(x+z, A + \nabla\varphi_*(x+z)) : \nabla v(x+z)\,\mathrm{d}x \\ &= \int_{(0,1)^d} \mathrm{D}_A f(x, A + \nabla\varphi_*(x)) : \nabla w(x)\,\mathrm{d}x,\end{aligned}$$

where

$$w(x) := \sum_{z\in\{0,\dots,k-1\}^d} v(x+z), \qquad x \in (0,1)^d.$$

Then, w is 1-periodic, as can be checked easily, and thus, by the Euler–Lagrange equation,

$$\begin{aligned}\int_{(0,k)^d} & f(x, A + \nabla\psi_*(x)) - f(x, A + \nabla\varphi_*(x))\,\mathrm{d}x \\ &\geq \int_{(0,k)^d} \mathrm{D}_A f(x, A + \nabla\varphi_*(x)) : \nabla w(x)\,\mathrm{d}x \\ &= 0.\end{aligned}$$

This shows (13.34) and completes the proof. □

We remark that in the general non-convex case the reduction to the cell problem on $(0, 1)^d$ is not possible. A counterexample can be found in [199] or in Section 14.4 of [52].

13.6 Quadratic Homogenization

Finally, we give a more concrete formula for f_{hom} in the case when f is quadratic.

Theorem 13.20. *In the situation of Theorem 13.13 (for $p = 2$), assume that f has the form*

$$f(x, A) = A : \mathbf{S}(x)\, A, \qquad (x, A) \in \mathbb{R}^d \times \mathbb{R}^{m\times d},$$

with a symmetric, uniformly positive definite, and 1-periodic fourth-order tensor field $\mathbf{S}(x) = \mathbf{S}^{ik}_{jl}(x)$ that is measurable in $x \in \mathbb{R}^d$. Then, there exists a symmetric and positive definite fourth-order tensor $\mathbf{S}_{\mathrm{hom}} = [\mathbf{S}_{\mathrm{hom}}]^{ik}_{jl}$ such that

$$f_{\mathrm{hom}}(A) = A : \mathbf{S}_{\mathrm{hom}}\, A, \qquad A \in \mathbb{R}^{m\times d}.$$

Moreover, the entries $[\mathbf{S}_{\mathrm{hom}}]^{ik}_{jl}$ ($i, k \in \{1, \dots, m\}$, $j, l \in \{1, \dots, d\}$) of the tensor $\mathbf{S}_{\mathrm{hom}}$ are given as

$$[\mathbf{S}_{\mathrm{hom}}]^{ik}_{jl} = \int_{(0,1)^d} (\mathrm{e}_i \otimes \mathrm{e}_j + \nabla\varphi_{i,j}(x)) : \mathbf{S}(x)\, (\mathrm{e}_k \otimes \mathrm{e}_l + \nabla\varphi_{k,l}(x))\, \mathrm{d}x, \qquad (13.35)$$

*where $\varphi_{i,j} \in \mathrm{W}^{1,2}_{\mathrm{per}}((0, 1)^d; \mathbb{R}^m)$ is the (unique up to constants) weak solution of the **cell problem** PDE*

$$\begin{cases} -\operatorname{div}\left[\mathbf{S}(x)(\mathrm{e}_i \otimes \mathrm{e}_j + \nabla\varphi_{i,j}(x))\right] = 0, & x \in (0, 1)^d, \\ \varphi_{i,j} \text{ has periodic boundary values.} \end{cases}$$

We also explicitly formulate this theorem for the scalar case:

Corollary 13.21. *In the situation of Theorem 13.13 (for $p = 2$), assume that $m = 1$ and that f has the form*

$$f(x, \xi) = \xi^T S(x)\, \xi, \qquad (x, \xi) \in \mathbb{R}^d \times \mathbb{R}^d,$$

with a symmetric, uniformly positive definite, and 1-periodic measurable matrix function $S\colon \mathbb{R}^d \to \mathbb{R}^{d\times d}$. Then, there exists a symmetric and positive definite matrix $S_{\mathrm{hom}} \in \mathbb{R}^{d\times d}$ such that

$$f_{\mathrm{hom}}(\xi) = \xi^T S_{\mathrm{hom}}\, \xi, \qquad \xi \in \mathbb{R}^d.$$

Moreover, the entries $[S_{\mathrm{hom}}]_l^k$ *($k, l \in \{1, \ldots, d\}$) of the matrix* S_{hom} *are given as*

$$[S_{\mathrm{hom}}]_l^k = \int_{(0,1)^d} (\mathrm{e}_k + \nabla\varphi_k(x))^T S(x)\, (\mathrm{e}_l + \nabla\varphi_l(x))\, \mathrm{d}x, \tag{13.36}$$

where $\varphi_k \in \mathrm{W}^{1,2}_{\mathrm{per}}((0,1)^d)$ *is the (unique up to constants) weak solution of the cell problem PDE*

$$\begin{cases} -\operatorname{div}\left[S(x)(\mathrm{e}_k + \nabla\varphi_k(x))\right] = 0, & x \in (0,1)^d, \\ \varphi_k \text{ has periodic boundary values.} \end{cases}$$

We will use the following elementary characterization of quadratic forms, whose proof is the task of Problem 13.10.

Lemma 13.22. *Let* X *be a Banach space and let* $F\colon X \to [0,\infty)$*. Then,* F *is a quadratic form, that is,* $F(x) = B(x,x)$ *with* $B\colon X \times X \to \mathbb{R}$ *bilinear, if and only if it satisfies the following conditions:*

(i) $F(0) = 0$;
(ii) $F(tx) = t^2 F(x)$ *for all* $x \in X$, $t > 0$;
(iii) $F(x+y) + F(x-y) \le 2F(x) + 2F(y)$ *for every* $x, y \in X$.

From the preceding lemma we may infer an abstract result about Γ-convergence:

Proposition 13.23. *Let* X *be a Banach space and let* $\mathscr{F}_k\colon X \to [0,\infty)$, $k \in \mathbb{N}$, *be positive definite quadratic forms that* Γ*-converge to* $\mathscr{F}_\infty\colon X \to \mathbb{R}$*. Then,* $\mathscr{F}_\infty$ *is also a positive definite quadratic form.*

Proof. It is easy to see that $\mathscr{F}_\infty \ge 0$. We will show the conditions (i)–(iii) of Lemma 13.22 for $\mathscr{F}_\infty$.

For (i), we notice by the lim inf-inequality that

$$\mathscr{F}_\infty[0] \le \liminf_{j\to\infty} \mathscr{F}_k[0] = 0.$$

Thus $\mathscr{F}_\infty[0] = 0$.

For (ii), let $u \in X$ and $t > 0$. Take a recovery sequence $(u_k) \subset X$ of u. Clearly, $tu_k \to tu$ in X. Thus, using the lim inf-inequality, the positive 2-homogeneity of the $\mathscr{F}_k$'s, and the recovery property of (u_k),

$$\mathscr{F}_\infty[tu] \le \liminf_{k\to\infty} \mathscr{F}_k[tu_k] = t^2 \cdot \lim_{k\to\infty} \mathscr{F}_k[u_k] = t^2 \mathscr{F}_\infty[u].$$

Hence,

$$\mathscr{F}_\infty[u] \le \frac{1}{t^2}\mathscr{F}_\infty[tu] \le \mathscr{F}_\infty[u],$$

from which we conclude that $\mathscr{F}_\infty[tu] = t^2 \mathscr{F}_\infty[u]$.

Finally, for (iii), we take $u, v \in X$ and let $(u_k), (v_k) \subset X$ be recovery sequences for u and v, respectively, Then it suffices to observe by the lim inf-inequality and the recovery property of $(u_k), (v_k)$ together with property (iii) for $\mathscr{F}_k$ that

$$\begin{aligned}
&\mathscr{F}_\infty[u+v] + \mathscr{F}_\infty[u-v] - 2\mathscr{F}_\infty[u] - 2\mathscr{F}_\infty[v] \\
&\quad \le \liminf_{j\to\infty} \big(\mathscr{F}_k[u_k+v_k] + \mathscr{F}_k[u_k-v_k] - 2\mathscr{F}_k[u_k] - 2\mathscr{F}_k[v_k]\big) \\
&\quad \le 0.
\end{aligned}$$

This shows (iii) for $\mathscr{F}_\infty$. Hence, by Lemma 13.22, $\mathscr{F}_\infty$ is a quadratic form. □

Proof of Theorem 13.20. Via Proposition 13.23 we immediately have

$$f_{\mathrm{hom}}(A) = A : \mathbf{S}_{\mathrm{hom}}\, A = \sum_{i,k=1}^{m} \sum_{j,l=1}^{d} [\mathbf{S}_{\mathrm{hom}}]_{jl}^{ik} A_j^i A_l^k$$

for a positive definite fourth-order tensor $\mathbf{S}_{\mathrm{hom}} = [\mathbf{S}_{\mathrm{hom}}]_{jl}^{ik}$ as in the statement of the theorem. It only remains to show the formula (13.35).

From Marcellini's Theorem 13.18 we know that for $A \in \mathbb{R}^{m\times d}$ it holds that

$$f_{\mathrm{hom}}(A) = \inf_{\varphi\in \mathrm{W}^{1,p}_{\mathrm{per}}((0,1)^d;\mathbb{R}^m)} \int_{(0,1)^d} (A+\nabla\varphi(x)) : \mathbf{S}(x)\,(A+\nabla\varphi(x))\,\mathrm{d}x.$$

The minimizer $\varphi_A \in \mathrm{W}^{1,p}_{\mathrm{per}}((0,1)^d;\mathbb{R}^m)$ to this minimization problem exists, is unique up to constants (by a suitable adaptation of Proposition 2.10), and satisfies the Euler–Lagrange equation (see Problem 13.7)

$$-\operatorname{div}\big[\mathbf{S}(x)(A+\nabla\varphi_A(x))\big] = 0, \qquad x \in (0,1)^d,$$

weakly (with the row-wise divergence). Moreover, φ_A is linear in the matrix $A \in \mathbb{R}^{m\times d}$, as can be seen directly from the Euler–Lagrange equation. Thus, abbreviating

$$\varphi_{i,j} := \varphi_{\mathrm{e}_i\otimes \mathrm{e}_j},$$

we may write

$$\varphi_A = \sum_{i=1}^{m} \sum_{j=1}^{d} A_j^i \varphi_{i,j}.$$

Hence,

$$f_{\text{hom}}(A) = \sum_{i,k=1}^{m} \sum_{j,l=1}^{d} [\mathbf{S}_{\text{hom}}]_{jl}^{ik} A_j^i A_l^k$$
$$= \int_{(0,1)^d} (A + \nabla\varphi_A(x)) : \mathbf{S}(x)\,(A + \nabla\varphi_A(x))\,\mathrm{d}x.$$

Thus, $f_{\text{hom}}(A)$ is equal to

$$\sum_{i,k=1}^{m} \sum_{j,l=1}^{d} \left(\int_{(0,1)^d} (\mathrm{e}_i \otimes \mathrm{e}_j + \nabla\varphi_{i,j}(x)) : \mathbf{S}(x)\,(\mathrm{e}_k \otimes \mathrm{e}_l + \nabla\varphi_{k,l}(x))\,\mathrm{d}x \right) A_j^i A_l^k,$$

which implies (13.35). □

Example 13.24. We now consider the scalar case as in Corollary 13.21 and assume furthermore that

$$S(x) = h_1(x_1) \cdots h_d(x_d)\mathrm{Id}, \qquad x \in \mathbb{R}^d,$$

where all $h_1, \ldots, h_d \colon \mathbb{R} \to [0, \infty)$ are 1-periodic Borel functions such that there exist constants $\alpha, \beta > 0$ with $\alpha \le h_i \le \beta$. In this case, we use the formula (13.36) to infer that

$$[S_{\text{hom}}]_k^k = \int_{(0,1)^d} h_1(x_1) \cdots h_d(x_d)\,|\mathrm{e}_k + \nabla\varphi_k(x))|^2\,\mathrm{d}x, \tag{13.37}$$

where $\varphi_k \in \mathrm{W}^{1,2}_{\text{per}}((0,1)^d)$ is the (unique up to additive constants) weak solution of the cell problem

$$-\operatorname{div}\left[h_1(x_1) \cdots h_d(x_d)(\mathrm{e}_k + \nabla\varphi_k(x))\right] = 0, \qquad x \in (0,1)^d. \tag{13.38}$$

We can also deduce that, with a slight abuse of notation, $\varphi_k(x) = \varphi_k(x_k)$. Indeed, choosing the ansatz that φ_k only depends on x_k, the cell problem reads

$$[h_k(s)(1 + \varphi_k'(s))]' = 0, \qquad s \in (0,1), \tag{13.39}$$

which has a unique periodic solution (up to additive constants). Since the solution to (13.38) is unique (up to additive constants), it must be given by this $\varphi_k(x) = \varphi_k(x_k)$. From the special structure of φ_k, we then deduce that

$$[S_{\text{hom}}]_l^k = \int_{(0,1)^d} h_1(x_1) \cdots h_d(x_d)(\mathrm{e}_k \cdot \mathrm{e}_l)\,\mathrm{d}x = 0 \qquad \text{if } k \neq l.$$

Furthermore, as a consequence of (13.39),

$$h_k(s)(1 + \varphi_k'(s)) \equiv \text{const},$$

and since $\int_0^1 1 + \varphi_k' \, \mathrm{d}s = 1$, we may conclude that

$$1 + \varphi_k'(t) = \frac{1}{h_k(t)} \left(\int_0^1 \frac{1}{h_k(s)} \, \mathrm{d}s \right)^{-1}.$$

Plugging this into (13.37), we calculate

$$\begin{aligned}[S_{\mathrm{hom}}]_k^k &= \int_{(0,1)^d} h_1(x_1) \cdots h_d(x_d) \frac{1}{h_k(x_k)^2} \left(\int_0^1 \frac{1}{h_k(s)} \, \mathrm{d}s \right)^{-2} \mathrm{d}x \\ &= \overline{h_1} \cdots \overline{h_{k-1}} \cdot \underline{h_k} \cdot \overline{h_{k+1}} \cdots \overline{h_d},\end{aligned}$$

where we denoted the *arithmetic mean* of h_i by

$$\overline{h_i} := \int_0^1 h_i(s) \, \mathrm{d}s$$

and the *harmonic mean* of h_i by

$$\underline{h_i} := \left(\int_0^1 \frac{1}{h_i(s)} \, \mathrm{d}s \right)^{-1}.$$

Example 13.25. In Section 1.10 we were led to identify the variational limit as $\varepsilon \downarrow 0$ of the functionals

$$\begin{aligned}\mathscr{F}_\varepsilon[u] &:= \int_\Omega \frac{1}{2} \mathscr{E}u(x) : \mathbf{C}\Big(\frac{x}{\varepsilon}\Big) \, \mathscr{E}u(x) - b(x) \cdot u(x) \, \mathrm{d}x \\ &= \int_\Omega f\Big(\frac{x}{\varepsilon}, \nabla u(x)\Big) - b(x) \cdot u(x) \, \mathrm{d}x,\end{aligned}$$

where

$$\mathbf{C}(x) = \mathbf{C}_1 + (\mathbf{C}_2 - \mathbf{C}_1) h(x_1), \qquad x \in \mathbb{R}^3,$$

and

$$h(t) := \begin{cases} 0 & \text{if } t - \lfloor t \rfloor \le \theta, \\ 1 & \text{if } t - \lfloor t \rfloor > \theta, \end{cases}$$

for some $\theta \in (0, 1)$, and $b \in \mathrm{L}^2(\Omega; \mathbb{R}^3)$ (say). From Theorem 13.20 in conjunction with the straightforward fact that the addition of strongly continuous ε-independent functionals commutes with Γ-convergence, we infer that indeed the $\mathscr{F}_\varepsilon$ Γ-converge with respect to the weak topology in $\mathrm{W}^{1,2}(\Omega; \mathbb{R}^3)$ to the functional

$$\mathscr{F}_0[u] = \int_\Omega \frac{1}{2} \mathscr{E}u(x) : \mathbf{C}_{\mathrm{hom}} \, \mathscr{E}u(x) - b(x) \cdot u(x) \, \mathrm{d}x$$

for some symmetric and positive definite fourth-order tensor $\mathbf{C}_{\text{hom}} = [\mathbf{C}_{\text{hom}}]^{ik}_{jl}$.

We finally consider the special case when

$$\mathbf{C}_1 = \alpha\,\mathbf{I}, \qquad \mathrm{C}_2 = \beta\,\mathbf{I}$$

where $\alpha, \beta > 0$ and $\mathbf{I}$ denotes the tensor such that $A : \mathbf{I}B = A : B$ (that is, $\mathbf{I}^{ik}_{jl} = \delta_{ik}\delta_{jl}$). In this case, a simple computation shows that

$$f(x, A) = \big(\alpha + (\beta - \alpha)h(x_1)\big)|A|^2, \qquad (x, A) \in \mathbb{R}^3 \times \mathbb{R}^{3\times 3}.$$

Then, we can adapt the previous example to see that

$$\begin{aligned}
[\mathbf{C}_{\text{hom}}]^{ii}_{11} &= \left(\frac{\theta}{\alpha} + \frac{1-\theta}{\beta}\right)^{-1}, \\
[\mathbf{C}_{\text{hom}}]^{ii}_{jj} &= \theta\alpha + (1-\theta)\beta \qquad \text{if } j \neq 1,
\end{aligned}$$

and $[\mathbf{C}_{\text{hom}}]^{ik}_{jl} = 0$ for all other values of i, j, k, l. Equivalently,

$$[\mathbf{C}_{\text{hom}}]^{ik}_{jl} = \delta_{ik}\delta_{jl}\left[\left(\frac{\theta}{\alpha} + \frac{1-\theta}{\beta}\right)^{-1}\delta_{j1} + \big(\theta\alpha + (1-\theta)\beta\big)(1-\delta_{j1})\right].$$

In particular, $\mathbf{C}_{\text{hom}}$ can no longer be written as $\gamma\mathbf{I}$ for some $\gamma \in \mathbb{R}$. This is also not surprising since the anisotropic lamination structure (in the first coordinate direction) is reflected in $\mathbf{C}_{\text{hom}}$.

Notes and Historical Remarks

The theory of Γ-convergence was founded by Ennio De Giorgi in the 1970s and nowadays is widely used in the calculus of variations. All results from Section 13.1 are essentially due to De Giorgi, see [89]. The book [51] provides a first introduction to Γ-convergence with many applications to homogenization theory, phase transition, and free discontinuity problems. The encyclopedic work [82] treats Γ-convergence in general topological spaces and also considers many applications to integral functionals.

For the Modica–Mortola Theorem 13.6 we closely follow the original works [189, 190], with some minor modifications. Our proof of Theorem 13.13 is the one in [199], which was obtained independently of Braides' proof in [50].

Another powerful technique in the theory of Γ-convergence is the *compactness method*: It can often be shown by a compactness argument that the Γ-limit of a (sub)sequence of integral functionals exists as some abstract functional. The task is then to identify this Γ-limit. For this, one can use a technique, which seems to be due to De Giorgi–Letta, Fusco, and Braides, where one first shows that the

Γ-limit parametrized on the domain is a measure, which then in a further step is seen to be given by an integral. The book [52] uses this technique to present many homogenization results in a unified framework. In fact, Chapter 14 of [52] proves Theorem 13.13 and several extensions using this technique.

In a non-variational context, that is, on the level of PDEs, several techniques have been developed for homogenization, which can also cope with completely non-periodic situations. We only mention G-convergence (see [82]), H-convergence (see [8]) and two-scale convergence (see [216] and also [8]). Here also the Div-curl Lemma 8.32 finds its natural home.

Problems

13.1. Let X be a complete metric space. Show that the Γ-limit of the functionals $\mathscr{F}_k \colon X \to \mathbb{R}$, $k \in \mathbb{N}$, if it exists, is uniquely determined.

13.2. Show that for a sequence $(u_j) \subset \mathrm{BV}(\Omega)$ and $u \in \mathrm{L}^1(\Omega)$ with $u_j \to u$ in L^1 it holds that

$$|Du|(\Omega) \le \liminf_{j\to\infty} |Du|(\Omega).$$

Also show that the perimeter $\mathrm{Per}_\Omega(E)$ is lower semicontinuous with respect to the convergence of sets $E_j \to E$ defined as $\mathbb{1}_{E_j} \to \mathbb{1}_E$ in $\mathrm{L}^1(\Omega)$. *Hint:* Write the total variation norm and the perimeter as a supremum over L^1-continuous functionals.

13.3. Let $\mathscr{F}_k \colon X \to \mathbb{R} \cup \{+\infty\}$, $k \in \mathbb{N}$, be equicoercive functionals on a complete metric space. Prove that $\Gamma\text{-}\liminf_k \mathscr{F}_k$ admits a minimizer and that

$$\min_X \, \Gamma\text{-}\liminf_k \mathscr{F}_k = \liminf_{k\to\infty} \, \inf_X \mathscr{F}_k.$$

Also show that

$$\limsup_{k\to\infty} \, \inf_X \mathscr{F}_k \le \inf_X \, \Gamma\text{-}\limsup_k \mathscr{F}_k.$$

Find a sequence of equicoercive functionals such that the Γ-upper limit does not fulfill the reverse inequality (this is contrary to the situation for the Γ-lower limit).

13.4. Find a non-separable and complete metric space (X, d) and a sequence of functionals $\mathscr{F}_k \colon X \to \mathbb{R} \cup \{+\infty\}$ such that no subsequence of the $\mathscr{F}_k$'s Γ-converges. *Hint:* Consider $X := \{-1, 1\}^{\mathbb{N}}$ and observe that Γ-convergence in X is equivalent to pointwise convergence.

13.5. Prove that if $g \in \mathrm{L}^p_{\mathrm{loc}}(\mathbb{R}^d)$ is k-periodic for some $k \in \mathbb{N}$, then the maps $h_\varepsilon(x) := g(x/\varepsilon)$ converge weakly in $\mathrm{L}^1_{\mathrm{loc}}$ to the constant map

$$h_0(x) := \fint_{(0,k)^d} g(y) \,\mathrm{d}y, \qquad x \in \mathbb{R}^d.$$

Hint: Inspect the proof of Lemma 4.15.

13.6. Let $f: (0,\infty)\times(0,\infty)\to\mathbb{R}$. Show that there exists a function $\delta: (0,\infty)\to(0,\infty)$ such that $\delta(\varepsilon)\to 0$ as $\varepsilon\downarrow 0$ and

$$\limsup_{\varepsilon\downarrow 0} f(\varepsilon,\delta(\varepsilon)) \le \limsup_{\delta\downarrow 0}\limsup_{\varepsilon\downarrow 0} f(\varepsilon,\delta).$$

13.7. Let f be as in Marcellini's Homogenization Theorem 13.18. Show that the problem

$$\begin{cases} \text{Minimize} & \displaystyle\int_{(0,1)^d} f(x, F+\nabla\varphi(x))\,\mathrm{d}x \\ \text{over all} & \varphi\in \mathrm{W}^{1,p}_{\mathrm{per}}((0,1)^d;\mathbb{R}^m) \end{cases}$$

has a solution $\varphi_*\in \mathrm{W}^{1,p}_{\mathrm{per}}((0,1)^d;\mathbb{R}^m)$. Moreover, prove that φ_* is a weak solution of the Euler–Lagrange equation

$$-\operatorname{div}\big[\mathrm{D}_A f(x, F+\nabla\varphi_*(x))\big] = 0, \qquad x\in(0,1)^d,$$

that is,

$$\int_\Omega \mathrm{D}_A f(x, F+\nabla\varphi_*(x)) : \nabla\psi(x)\,\mathrm{d}x = 0 \qquad \text{for all } \psi\in \mathrm{W}^{1,p}_{\mathrm{per}}((0,1)^d;\mathbb{R}^m).$$

13.8. Let $\Omega\subset\mathbb{R}^d$ be a bounded Lipschitz domain and let $f:\mathbb{R}^d\times\mathbb{R}^{m\times d}\to[0,\infty)$ be a Carathéodory integrand. Define the **partial regularization** $f_\delta:\mathbb{R}^d\times\mathbb{R}^{m\times d}\to[0,\infty)$ of f as

$$f_\delta(x,A) := \int \eta_\delta(A-B)f(x,B)\,\mathrm{d}B, \qquad (x,A)\in\mathbb{R}^d\times\mathbb{R}^{m\times d},$$

where $(\eta_\delta)_{\delta>0}$ is a radially symmetric and positive family of mollifiers on $\mathbb{R}^{m\times d}$. Show that $f_\delta\to f$ pointwise. Show also that if f satisfies any of the following conditions, then so does f_δ:

(i) $\mu|A|^p \le f(x,A)\le M(1+|A|^p)$ for all $(x,A)\in\mathbb{R}^d\times\mathbb{R}^{m\times d}$ and some $p\in(1,\infty)$, $\mu, M>0$;
(ii) $x\mapsto f(x,A)$ is 1-periodic for all $A\in\mathbb{R}^{m\times d}$;
(iii) $|f(x,A)-f(x,B)|\le C(1+|A|^{p-1}+|B|^{p-1})|A-B|$ for all $x\in\Omega$, $A,B\in\mathbb{R}^{m\times d}$ and some $C>0$.

13.9. In the situation of Marcellini's Homogenization Theorem 13.18, show that if $m=1$ then

$$f_{\mathrm{hom}}(A) = \inf_{\varphi\in \mathrm{W}^{1,p}_{\mathrm{per}}((0,k)^d)} \int_{(0,k)^d} f^{**}(x, A+\nabla\varphi(x))\,\mathrm{d}x,$$

where f^{**} denotes the convex envelope of f with respect to the second argument. Conclude that for $m = 1$ Marcellini's Homogenization Theorem 13.18 also holds without assuming the convexity of f in the second argument. *Hint:* Extend a relaxation theorem to periodic integrands.

13.10. Prove Lemma 13.22.

Appendix A
Prerequisites

This appendix recalls some notation and results that are needed throughout the book.

A.1 Linear Algebra

We first review some facts from linear algebra and matrix analysis, see [151] for an advanced course.

For a $(m \times d)$-matrix $A \in \mathbb{R}^{m\times d}$ we denote by A^j_k the element in the j'th row and k'th column ($j = 1, \ldots, m$; $k = 1, \ldots, d$). In this book the matrix space $\mathbb{R}^{m\times d}$ always comes equipped with the **Frobenius matrix (inner) product**

$$A : B := \operatorname{tr}(A^T B) = \operatorname{tr}(AB^T) = \sum_{j,k} A^j_k B^j_k, \qquad A, B \in \mathbb{R}^{m\times d},$$

which is just the Euclidean product if we identify such matrices with vectors in $\mathbb{R}^{md}$. This inner product induces the **Frobenius matrix norm**

$$|A| := \sqrt{\sum_{j,k} (A^j_k)^2}, \qquad A \in \mathbb{R}^{m\times d}.$$

While of course all norms on the finite-dimensional space $\mathbb{R}^{m\times d}$ are equivalent, some finer arguments require us to specify a matrix norm; if nothing else is stated, we always use the Frobenius norm.

The Frobenius norm can also be expressed as

$$|A| = \sqrt{\sum_i \sigma_i(A)^2}, \qquad A \in \mathbb{R}^{m\times d},$$

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8

where $\sigma_i(A) \geq 0$ is the i'th singular value of A, $i = 1, \ldots, \min\{d, m\}$. For this, recall that every matrix $A \in \mathbb{R}^{m \times d}$ has a **(real) singular value decomposition**

$$A = P \Sigma Q^T$$

for orthogonal matrices $P \in \mathbb{R}^{m \times m}$, $Q \in \mathbb{R}^{d \times d}$ ($P^{-1} = P^T$, $Q^{-1} = Q^T$), and a diagonal matrix

$$\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_{\min\{d,m\}}) = \left(\begin{array}{ccc} \sigma_1 & & \\ & \ddots & \\ & & \sigma_{\min\{d,m\}} \\ \hline 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{array}\right) \in \mathbb{R}^{m \times d}$$

with only positive diagonal entries

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0 = \sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_{\min\{d,m\}},$$

called the **singular values** of A, where r is the rank of A.

From the above expression using the singular values, it follows immediately that the Frobenius norm is orthogonally invariant, that is, for all $A \in \mathbb{R}^{m \times d}$ and all orthogonal $P \in \mathbb{R}^{m \times m}$, $Q \in \mathbb{R}^{d \times d}$ it holds that

$$|PA| = |A| = |AQ|.$$

A special matrix in $\mathbb{R}^{m \times d}$ is the **tensor product** of the vectors $a \in \mathbb{R}^m$, $b \in \mathbb{R}^d$, which is defined as

$$a \otimes b := ab^T \in \mathbb{R}^{m \times d}.$$

Occasionally, we will also use $a \otimes b$ for a a column vector and b a row-vector to denote the matrix product ab. While technically incorrect, this notation emphasizes that the result is a *matrix*. We recall the following elementary fact: Let $A \in \mathbb{R}^{m \times d}$. Then, rank $A \leq 1$ if and only if there exist vectors $a \in \mathbb{R}^m$, $b \in \mathbb{R}^d$ such that $A = a \otimes b$. The tensor product also interacts well with the Frobenius norm:

$$|a \otimes b| = |a| \cdot |b|, \qquad a \in \mathbb{R}^m,\ b \in \mathbb{R}^d,$$

where in $\mathbb{R}^m$ and $\mathbb{R}^d$ we use the usual Euclidean norm.

A fundamental inequality involving the determinant is the **Hadamard inequality**: Let $A \in \mathbb{R}^{d \times d}$ with columns $A_j \in \mathbb{R}^d$ ($j = 1, \ldots, d$). Then,

$$|\det A| \leq \prod_{j=1}^{d} |A_j| \leq |A|^d.$$

An analogous formula holds with the rows of A.

For $A \in \mathbb{R}^{d\times d}$ the **cofactor matrix** $\operatorname{cof} A \in \mathbb{R}^{d\times d}$ of A is the matrix whose (j,k)'th entry is $(-1)^{j+k} M^{\neg j}_{\neg k}(A)$ with $M^{\neg j}_{\neg k}(A)$ being the (j,k)**-minor** of A, i.e., the determinant of the matrix that originates from A by deleting the j'th row and the k'th column. For

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

we have

$$\operatorname{cof} A = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}.$$

One important formula (and another way to define the cofactor matrix) is

$$\frac{\mathrm{d}}{\mathrm{d}A} \det A = \operatorname{cof} A.$$

Furthermore, **Cramer's rule** entails that

$$A(\operatorname{cof} A)^T = (\operatorname{cof} A)^T A = (\det A)\mathrm{Id}, \qquad A \in \mathbb{R}^{d\times d},$$

where Id denotes the identity matrix. In particular, if A is invertible,

$$A^{-1} = \frac{(\operatorname{cof} A)^T}{\det A}. \tag{A.1}$$

From this we deduce that $\operatorname{cof} A$ is invertible if A is. Sometimes, the matrix $(\operatorname{cof} A)^T$ is called the **adjugate matrix** to A in the literature (we will not use this terminology here, however).

One particular consequence of Cramer's rule is **Jacobi's formula**, which says that for any continuously differentiable function $A(t)\colon \mathbb{R} \to \mathbb{R}^{d\times d}$ it holds that

$$\frac{\mathrm{d}}{\mathrm{d}t} \det A(t) = \operatorname{cof} A(t) : \frac{\mathrm{d}A(t)}{\mathrm{d}t} = \operatorname{tr}\Big[(\operatorname{cof} A(t))^T \frac{\mathrm{d}A(t)}{\mathrm{d}t}\Big].$$

In particular, if $A(t) = A_0 + tB$ ($A_0, B \in \mathbb{R}^{d\times d}$), we have

$$\frac{\mathrm{d}}{\mathrm{d}t} \det[A_0 + tB] = \operatorname{tr}\Big[(\operatorname{cof} A_0)^T B + t(\operatorname{cof} B)^T B\Big] = (\operatorname{cof} A_0) : B + td \det B.$$

As a consequence, we derive that if $\frac{\mathrm{d}}{\mathrm{d}t} \det[A_0 + tB]$ is constant, then necessarily $\det B = 0$.

The **special orthogonal group** $\mathrm{SO}(d)$ is defined as

$$\mathrm{SO}(d) := \left\{ Q \in \mathbb{R}^{d\times d} \;:\; Q \text{ invertible}, \; Q^{-1} = Q^T, \det Q = 1 \right\}.$$

It has the following useful property, which can be verified via (A.1):

$$\operatorname{cof} Q = Q \qquad \text{for all } Q \in \mathrm{SO}(d).$$

Any $Q \in \mathrm{SO}(2) \subset \mathbb{R}^{2\times 2}$ (a rotation) has the form

$$Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

for some $\theta \in [0, 2\pi)$. For the convex hull $\mathrm{SO}(2)^{**}$ of $\mathrm{SO}(2)$ one may compute

$$\mathrm{SO}(2)^{**} = \left\{ A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \;:\; a^2 + b^2 \le 1 \right\}.$$

Moreover, a Taylor expansion and the fact that the Lie algebra of the Lie group SO(2) is the vector space of all skew-symmetric matrices yields

$$\operatorname{dist}(\mathrm{Id} + A, \mathrm{SO}(2)) \le \frac{1}{2}|A + A^T| + C|A|^2 \tag{A.2}$$

for some $C > 0$.

We also recall a special case of the theorem on the **Jordan normal form** for real (2×2)-matrices: Let $A \in \mathbb{R}^{2\times 2}$. Then, there exists an invertible matrix $S \in \mathbb{R}^{2\times 2}$ such that

$$S^{-1}AS = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \qquad \text{or} \qquad S^{-1}AS = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

with $a, c \in \mathbb{R}$ and $b \in \{0, 1\}$

Finally, any square matrix $A \in \mathbb{R}^{d\times d}$ has a real **polar decomposition**

$$A = QS,$$

where Q is orthogonal ($Q^{-1} = Q^T$) and S is symmetric and positive definite.

In a few instances we also deal with fourth-order **tensors** $\mathbf{T} = \mathbf{T}^{ik}_{jl}$ ($i, k = 1, \ldots, m$; $j, l = 1, \ldots, d$). They define bilinear forms on $\mathbb{R}^{m\times d}$ via

$$A : \mathbf{T}B := \sum_{i,k} \sum_{j,l} \mathbf{T}^{ik}_{jl} A^i_j B^k_l.$$

We call $\mathbf{T}$ symmetric and positive definite if the corresponding bilinear form has these properties.

A.2 Functional Analysis

We assume that the reader has a solid foundation in the basic notions of functional analysis such as Banach spaces and their duals, weak/weak* convergence (and topology), reflexivity, and weak/weak* compactness. Note that in this book we mostly only need convergence of sequences and rarely more advanced topological concepts. One very thorough reference for most of this material is [74].

Let X be a Banach space. The application of $x^* \in X^*$ to $x \in X$ is often expressed via the **duality pairing** $\langle x, x^* \rangle := x^*(x)$. We write $x_j \rightharpoonup x$ in X for **weak convergence**, that is, $\langle x_j, x^* \rangle \to \langle x, x^* \rangle$ for all $x^* \in X^*$, and $x_j^* \overset{*}{\rightharpoonup} x^*$ in X^* for **weak* convergence**, that is, $\langle x, x_j^* \rangle \to \langle x, x^* \rangle$ for all $x \in X$. The weak* topology is metrizable on norm-bounded sets in the dual to a separable Banach space. Likewise, in reflexive and separable Banach spaces the weak topology is metrizable on norm-bounded sets. In this context we note that in Banach spaces topological weak compactness is equivalent to sequential weak compactness by the Eberlein–Šmulian theorem. Also recall that the norm in a Banach space is lower semicontinuous with respect to weak convergence: If $x_j \rightharpoonup x$ in X, then $\|x\| \leq \liminf_{j\to\infty} \|x_j\|$.

Theorem A.1 (**Hahn–Banach separation theorem**). *Let X be a Banach space and let $K, F \subset X$ be disjoint, non-empty, and convex subsets of X such that K is compact and F is closed. Then, K and F can be separated by a hyperplane, that is, there exists an $x^* \in X^*$ such that*

$$\sup_{x\in K} \langle x, x^* \rangle < \inf_{x\in F} \langle x, x^* \rangle.$$

Theorem A.2 (**Weak compactness**). *Let X be a separable, reflexive Banach space. Then, norm-bounded sets in X are sequentially weakly precompact.*

Theorem A.3 (**Banach–Alaoglu**). *Let X be a separable Banach space. Then, norm-bounded sets in the dual space X^* are weakly* sequentially precompact.*

Weak convergence can be "improved" to strong convergence in the following way (see Section I.1.2 in [106] for a proof):

Lemma A.4 (**Mazur**). *Let $x_j \rightharpoonup x$ in a Banach space X. Then, there exists a sequence $(y_j) \subset X$ of convex combinations,*

$$y_j = \sum_{n=j}^{N(j)} \theta_n^{(j)} x_n, \qquad \theta_n^{(j)} \in [0, 1], \qquad \sum_{n=j}^{N(j)} \theta_n^{(j)} = 1$$

such that $y_j \to x$ in X.

A.3 Measure Theory

We assume that the reader is familiar with the notion of Lebesgue- and Borel-measurability, negligible sets, and L^p-spaces; a good introduction is [236]. For the d-dimensional Lebesgue measure we write $\mathscr{L}^d$ or $\mathscr{L}^d_x$ if we want to stress the integration variable. Often, however, the Lebesgue measure of a Borel- or Lebesgue-measurable set $A \subset \mathbb{R}^d$ is simply denoted by $|A|$. We also write ω_d for the volume of the d-dimensional unit ball.

The **indicator function** of a subset $A \in \mathbb{R}^d$ is

$$\mathbb{1}_A(x) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{otherwise,} \end{cases} \qquad x \in \mathbb{R}^d.$$

In the following we recall some basic results that are needed throughout the book.

Lemma A.5 *Let $B \subset \mathbb{R}^d$ be a Borel set. A function $f : B \to \mathbb{R}^N$ is Lebesgue-measurable if and only if there exists a sequence of simple functions*

$$f_j := \sum_{k=1}^{K(j)} v_k^{(j)} \mathbb{1}_{E_k^{(j)}}$$

such that

$$f_j \to f \quad \text{pointwise} \qquad \text{as } j \to \infty,$$

where $K(j) \in \mathbb{N}$, the $E_k^{(j)} \subset B$ are Lebesgue-measurable sets with $\bigcup_{k=1}^{K(j)} E_k^{(j)} = B$, and $v_k^{(j)} \in \mathbb{R}^N$ for all j, k.

Lemma A.6 **(Fatou).** *Let $f_j : \mathbb{R}^d \to [0, +\infty]$, $j \in \mathbb{N}$, be Lebesgue-measurable functions. Then,*

$$\int \liminf_{j\to\infty} f_j(x) \, dx \le \liminf_{j\to\infty} \int f_j(x) \, dx.$$

Lemma A.7 **(Monotone convergence).** *Let $f_j : \mathbb{R}^d \to [0, +\infty]$, $j \in \mathbb{N}$, be Lebesgue-measurable functions with $f_j(x) \uparrow f(x)$ for almost every $x \in \Omega$. Then, $f : \mathbb{R}^d \to [0, +\infty]$ is measurable and*

$$\int f(x) \, dx = \lim_{j\to\infty} \int f_j(x) \, dx.$$

Lemma A.8 *If $f_j \to f$ (strongly) in $\mathrm{L}^p(\mathbb{R}^d)$, $p \in [1, \infty]$, that is,*

$$\|f_j - f\|_{\mathrm{L}^p} \to 0 \qquad \text{as } j \to \infty,$$

then there exists a subsequence (not explicitly labeled; we do not choose different indices for subsequences if the original sequence is discarded at the same time) such that $f_j \to f$ pointwise almost everywhere.

Theorem A.9 (**Lebesgue dominated convergence theorem**). *Let $f_j : \mathbb{R}^d \to \mathbb{R}^N$, $j \in \mathbb{N}$, be Lebesgue-measurable functions such that there exists an* L^p*-integrable majorant $g \in \mathrm{L}^p(\mathbb{R}^d)$ for some $p \in [1, \infty)$, that is,*

$$|f_j| \le g \qquad \textit{for all } j \in \mathbb{N}.$$

If $f_j \to f$ pointwise almost everywhere for some $f : \mathbb{R}^d \to \mathbb{R}$, then also $f_j \to f$ in L^p; in particular, $f \in \mathrm{L}^p(\mathbb{R}^d)$.

The following strengthening of Lebesgue's theorem is often useful in the calculus of variations:

Theorem A.10 (**Pratt**). *Let $f_j : \mathbb{R}^d \to \mathbb{R}^N$, $j \in \mathbb{N}$, be Lebesgue-measurable functions. If $f_j \to f$ pointwise almost everywhere (or in measure) and there exists a sequence $(g_j) \subset \mathrm{L}^1(\mathbb{R}^d)$ with $g_j \to g$ in L^1 such that $|f_j| \le g_j$, then $f_j \to f$ in L^1.*

The following convergence theorem is of fundamental significance:

Theorem A.11 (**Vitali**). *Let $\Omega \subset \mathbb{R}^d$ be bounded and let $(f_j) \subset \mathrm{L}^p(\Omega; \mathbb{R}^m)$, $p \in [1, \infty)$. Assume furthermore that the following two conditions hold:*

(i) No oscillations: *$f_j \to f$ in measure, that is, for all $\delta > 0$,*

$$\bigl|\bigl\{\, x \in \Omega \;:\; |f_j(x) - f(x)| > \delta \,\bigr\}\bigr| \to 0 \qquad \textit{as } j \to \infty.$$

(ii) No concentrations: *the family $\{f_j\}_j$ is L^p-equiintegrable.*

Then, $f_j \to f$ in L^p.

Here, $\{f_j\}_j \subset \mathrm{L}^p(\Omega; \mathbb{R}^m)$ is called L^p**-equiintegrable** if one of the following equivalent conditions is satisfied:

(i) $\displaystyle\lim_{R\uparrow\infty} \sup_{j\in\mathbb{N}} \int_{\{|f_j|>R\}} |f_j|^p \,\mathrm{d}x = 0;$

(ii) $\displaystyle\lim_{R\uparrow\infty} \limsup_{j\to\infty} \int_{\{|f_j|>R\}} |f_j|^p \,\mathrm{d}x = 0;$

(iii) for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all Borel sets $B \subset \Omega$ with $|B| < \delta$ we have

$$\sup_{j\in\mathbb{N}} \int_B |f_j|^p \,\mathrm{d}x < \varepsilon.$$

Theorem A.12 (**Dunford–Pettis**). *Let $\Omega \subset \mathbb{R}^d$ be bounded and open. A norm-bounded family $\{f_j\}_{j\in\mathbb{N}} \subset \mathrm{L}^1(\Omega)$ is equiintegrable if and only if it is weakly sequentially precompact in $\mathrm{L}^1(\Omega)$.*

We remark that the usual formulation of the Dunford–Pettis theorem only mentions *topological* precompactness. The statement above follows by also utilizing the Eberlein–Šmulian theorem (see Chapter V in [74]).

Theorem A.13 (Egorov). *Let $\Omega \subset \mathbb{R}^d$ be a bounded Borel set and let $f_j \colon \Omega \to \mathbb{R}^N$, $j \in \mathbb{N}$, be Lebesgue-measurable functions. If $f_j \to f$ pointwise almost everywhere, then for every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset \Omega$ with $|\Omega \setminus K_\varepsilon| \leq \varepsilon$ and such that $f_j \to f$ uniformly in K_ε.*

Theorem A.14 (Radon–Riesz). *Let $p \in (1, \infty)$ and let $(f_j) \subset \mathrm{L}^p(\Omega; \mathbb{R}^m)$ with $f_j \rightharpoonup f$ (weak convergence) as well as $\|f_j\|_{\mathrm{L}^p} \to \|f\|_{\mathrm{L}^p}$. Then, $f_j \to f$ in L^p.*

The following *covering theorem* is a handy tool for several constructions, see Theorem 2.19 in [15] for a proof:

Theorem A.15 (Vitali covering theorem). *Let $\Omega, D \subset \mathbb{R}^d$ be open and bounded. Then, there exist $a_k \in \Omega$, $r_k > 0$, where $k \in \mathbb{N}$, such that we may write Ω as the disjoint union*

$$\Omega = Z \;\cup\; \bigcup_{k=1}^{\infty} \overline{D(a_k, r_k)}, \qquad D(a_k, r_k) := a_k + r_k D,$$

with $Z \subset \Omega$ a Lebesgue-negligible set ($|Z| = 0$). Moreover, if for almost every $x \in \Omega$ we are given a real number $r(x) > 0$, then we may additionally require of the cover that $r_k < r(a_k)$ for all $k \in \mathbb{N}$.

Theorem A.16 (Lusin). *Let $\Omega \subset \mathbb{R}^d$ be a bounded Borel set and let $f \colon \Omega \to \mathbb{R}^N$ be Lebesgue-measurable. Then, for every $\varepsilon > 0$ there exists a compact set $K \subset \Omega$ such that $|\Omega \setminus K| \leq \varepsilon$ and $f|_K$ is continuous.*

Theorem A.17 (Sard). *Let $f \colon \mathbb{R}^d \to \mathbb{R}$ be d times continuously differentiable. Then,*

$$\mathscr{L}^1(f(S)) = 0, \qquad \textit{where} \qquad S := \bigl\{\, x \in \mathbb{R}^d \;\colon\; \nabla f(x) = 0 \,\bigr\}.$$

We also need other measures than Lebesgue measure on subsets of $\mathbb{R}^N$ (this will usually be either $\mathbb{R}^d$ or a matrix space $\mathbb{R}^{m \times d}$, which is identified with $\mathbb{R}^{md}$). All of these abstract measures will be **Borel measures**, that is, they are defined on the **Borel σ-algebra** $\mathfrak{B}(\mathbb{R}^N)$ of $\mathbb{R}^N$, which is the smallest σ-algebra that contains all the open sets. All positive Borel measures defined on $\mathbb{R}^N$ that do not take the value $+\infty$ are collected in the set $\mathscr{M}^+(\mathbb{R}^N)$ of **(finite) positive Radon measures**; its subclass of **probability measures** is $\mathscr{M}^1(\mathbb{R}^N)$. We remark that all σ-finite measures on $\mathbb{R}^N$ are in fact **inner regular**, meaning that for every Borel set $B \subset \mathbb{R}^N$ it holds that

$$\mu(B) = \sup\bigl\{\, \mu(K) \;\colon\; K \subset B \text{ compact} \,\bigr\}.$$

A **local Radon measure** μ is a set function $\mu \colon \mathfrak{B}(\mathbb{R}^N) \to [0, +\infty]$ such that μ restricted to (subsets of) any compact set is a finite Radon measure. In this case we write $\mu \in \mathscr{M}^+_{\mathrm{loc}}(\mathbb{R}^N)$. We also use $\mathscr{M}^+(U)$, $\mathscr{M}^+_{\mathrm{loc}}(U)$, $\mathscr{M}^1(U)$ for the subset of measures that only charge $U \subset \mathbb{R}^N$, that is, the measure of the complement of U is zero. A good reference for (advanced) measure theory is [15].

For $h\colon \mathbb{R}^N \to \mathbb{R}$ and $\mu \in \mathscr{M}^+(\mathbb{R}^N)$ we define the **duality pairing**

$$\langle h, \mu \rangle := \int h(A) \,\mathrm{d}\mu(A)$$

whenever this integral makes sense. The following notation is convenient for the **barycenter** of a (finite, positive) Borel measure $\mu \in \mathscr{M}^+(\mathbb{R}^N)$:

$$[\mu] := \langle \mathrm{id}, \mu \rangle = \int A \,\mathrm{d}\mu(A).$$

We also define the **support** of $\mu \in \mathscr{M}^+(\mathbb{R}^N)$ by

$$\operatorname{supp} \mu := \left\{ x \in \mathbb{R}^N \ \mathbf{:}\ \mu(B(x,r)) > 0 \text{ for all } r > 0 \right\},$$

where $B(x_0, r) \subset \mathbb{R}^N$ is the ball with center x_0 and radius $r > 0$. The **restriction** of a Borel measure $\mu \in \mathscr{M}^+(\mathbb{R}^N)$ to a Borel set $A \subset \mathbb{R}^N$ is

$$(\mu \llcorner A)(B) := \mu(A \cap B) \qquad \text{for any Borel set } B \subset \mathbb{R}^N.$$

Probability measures and convex functions interact well:

Lemma A.18 (**Jensen inequality**). *For all probability measures $\mu \in \mathscr{M}^1(\mathbb{R}^N)$ and all convex $h\colon \mathbb{R}^N \to \mathbb{R}$ it holds that*

$$h([\mu]) \le \int h(A) \, d\mu(A).$$

We say that a sequence $(\mu_j) \subset \mathscr{M}^+(\mathbb{R}^N)$ **converges weakly*** in $\mathscr{M}^+(\mathbb{R}^N)$ to $\mu \in \mathscr{M}^+(\mathbb{R}^N)$, in symbols "$\mu_j \overset{*}{\rightharpoonup} \mu$", if $\langle \psi, \mu_j \rangle \to \langle \psi, \mu \rangle$ for all $\psi \in \mathrm{C}_0(\mathbb{R}^N)$. We speak of **local weak* convergence** if $\langle \psi, \mu_j \rangle \to \langle \psi, \mu \rangle$ for all $\psi \in \mathrm{C}_c(\mathbb{R}^N)$. A sequence $(\mu_j) \subset \mathscr{M}^+(\mathbb{R}^N)$ with $\sup_j \mu_j(\mathbb{R}^N) < \infty$ has a weakly* converging subsequence by Theorem A.2.

We also recall a useful convergence lemma:

Lemma A.19 *Let $\mu_j \overset{*}{\rightharpoonup} \mu$ in $\mathscr{M}^+_{\mathrm{loc}}(\mathbb{R}^N)$. Then for every lower semicontinuous function $g\colon \mathbb{R}^N \to [0, \infty]$ it holds that*

$$\int g \, d\mu \le \liminf_{j\to\infty} \int g \, d\mu_j,$$

and for every upper semicontinuous function $h\colon \mathbb{R}^N \to [0, \infty)$ with compact support it holds that

$$\int h \, d\mu \ge \limsup_{j\to\infty} \int h \, d\mu_j.$$

In particular, for $U \subset \mathbb{R}^N$ open and $K \subset \mathbb{R}^N$ compact,

$$\mu(U) \le \liminf_{j\to\infty} \mu_j(U) \qquad \textit{and} \qquad \mu(K) \ge \limsup_{j\to\infty} \mu_j(K).$$

Similar definitions and statements apply to $\mathscr{M}^+(U)$, $\mathscr{M}^+_{\mathrm{loc}}(U)$, $\mathscr{M}^1(U)$.

Finally, we recall a very useful "continuity" property of measurable functions:

Theorem A.20 *Let $f \in \mathrm{L}^1(\mathbb{R}^N, \mu)$, that is, f is μ-integrable, where $\mu \in \mathscr{M}^+(\mathbb{R}^N)$. Then, μ-almost every $x_0 \in \mathbb{R}^N$ is a* ***Lebesgue point*** *of f with respect to μ, that is,*

$$\lim_{r\downarrow 0} ⨍_{B(x_0,r)} |f(x) - f(x_0)| \, d\mu(x) = 0,$$

where $⨍_{B(x_0,r)} := |B(x_0,r)|^{-1} \int_{B(x_0,r)}$.

We denote by $\mathscr{H}^s$ the s-dimensional Hausdorff measure, $0 \le s < \infty$. For the definition of this measure and the associated notion of $\mathscr{H}^s$**-rectifiable sets** (which is not important for most of this book), we refer to [15].

A.4 Vector Measures

In this section we exhibit a few aspects of the theory of **vector (Radon) measures** (often just called "measures" in this book), which are σ-additive set functions $\mu\colon \mathfrak{B}(\mathbb{R}^d) \to \mathbb{R}^N$ (in particular, $\mu(\emptyset) = 0$). All such μ are collected in the space $\mathscr{M}(\mathbb{R}^d;\mathbb{R}^N)$; likewise define $\mathscr{M}(\Omega;\mathbb{R}^N)$ and $\mathscr{M}(\overline{\Omega};\mathbb{R}^N)$ for an open set $\Omega \subset \mathbb{R}^d$. We will also use **local vector measures**, defined analogously to the above, which are collected in the set $\mathscr{M}_{\mathrm{loc}}(\Omega;\mathbb{R}^N)$.

If for the target dimension we have $N = 1$, then we simply write $\mathscr{M}(\Omega)$ instead of $\mathscr{M}(\Omega;\mathbb{R})$; the elements of this space are called **signed (Radon) measures**, but in this case we also usually just speak of "measures".

The **total variation measure** of $\mu \in \mathscr{M}(\mathbb{R}^d;\mathbb{R}^N)$ is the positive measure $|\mu| \in \mathscr{M}^+(\mathbb{R}^d)$ defined as

$$|\mu|(B) := \sup\left\{ \sum_{k=1}^{\infty} |\mu(B_k)| \;:\; B = \bigcup_{k=1}^{\infty} B_k \text{ as a disjoint union of Borel sets} \right\}.$$

It can be shown that for all open sets $U \subset \mathbb{R}^d$ it holds that

$$|\mu|(U) = \sup\left\{ \int \psi \cdot \mathrm{d}\mu \;:\; \psi \in \mathrm{C}_c(U;\mathbb{R}^N),\ \|\psi\|_\infty \le 1 \right\}, \tag{A.3}$$

where the dot "$\cdot$" indicates that μ's values are to be scalar-multiplied with the values of ψ (e.g. $\int (v_0 \mathbb{1}_B)\cdot d\mu = v_0 \cdot \mu(B)$ for $v_0 \in \mathbb{R}^N$ and B a Borel set). See Proposition 1.47 in [15] for a proof. We also set $\operatorname{supp} \mu := \operatorname{supp} |\mu|$.

There is an alternative, dual, view on vector measures, expressed in the following important theorem:

Theorem A.21 **(Riesz representation theorem).** *The space of vector Radon measures $\mathscr{M}(\mathbb{R}^d;\mathbb{R}^N)$ is isometrically isomorphic to the dual space $\mathrm{C}_0(\mathbb{R}^d;\mathbb{R}^N)^*$ via the duality pairing*

$$\langle \varphi, \mu \rangle = \int \varphi \cdot d\mu, \qquad \varphi \in \mathrm{C}_0(\mathbb{R}^d;\mathbb{R}^N),\ \mu \in \mathscr{M}(\mathbb{R}^d;\mathbb{R}^N).$$

As an easy, often convenient, consequence, we can *define* measures through their *action* on $\mathrm{C}_0(\mathbb{R}^d;\mathbb{R}^N)$. Note, however, that we then need to check the boundedness $|\langle \varphi, \mu\rangle| \leq \|\varphi\|_\infty$ for all $\varphi \in \mathrm{C}_0(\mathbb{R}^d;\mathbb{R}^N)$.

If an element $\mu \in \mathrm{C}_0(\mathbb{R}^d)^*$ is additionally positive, that is, $\langle \varphi, \mu\rangle \geq 0$ for $\varphi \geq 0$, and normalized, that is, $\langle \mathbb{1}, \mu\rangle = 1$ (here, $\mathbb{1} = 1$ on the whole space), then the μ from the Riesz representation theorem is a *probability measure*, $\mu \in \mathscr{M}^1(\mathbb{R}^N)$.

The weak* convergence of vector measures is defined exactly as for positive measures, namely by considering vector measures as elements of $\mathrm{C}_0(\mathbb{R}^d;\mathbb{R}^N)^*$. Sometimes, for a norm-bounded sequence $(v_j) \subset \mathrm{L}^1(\Omega;\mathbb{R}^N)$, we will say that "$v_j \overset{*}{\rightharpoonup} \mu$ in $\mathscr{M}(\Omega;\mathbb{R}^N)$" when really we mean $v_j \mathscr{L}^d \llcorner \Omega \overset{*}{\rightharpoonup} \mu$ in $\mathscr{M}(\Omega;\mathbb{R}^N)$. A sequence $(\mu_j) \subset \mathscr{M}(\Omega;\mathbb{R}^N)$ with $\sup_j |\mu_j|(\Omega) < \infty$ has a weakly* converging subsequence by Theorem A.2.

The following lemma is proved in Proposition 1.62 (b) of [15].

Lemma A.22 *Let $\mu_j \overset{*}{\rightharpoonup} \mu$ in $\mathscr{M}(\mathbb{R}^d;\mathbb{R}^N)$ and assume that $|\mu_j| \overset{*}{\rightharpoonup} \Lambda \in \mathscr{M}^+(\mathbb{R}^d)$. If $K \subset \mathbb{R}^d$ is compact and $\Lambda(\partial K) = 0$, then $\mu_j(K) \to \mu(K)$. Moreover, if $h\colon \mathbb{R}^d \to \mathbb{R}$ is a bounded Borel function with compact support and a Λ-negligible set of discontinuity points, then*

$$\int h \, d\mu_j \to \int h \, d\mu.$$

Of fundamental importance is the following theorem:

Theorem A.23 **(Besicovitch differentiation theorem).** *Given $\mu \in \mathscr{M}(\mathbb{R}^d;\mathbb{R}^N)$ and $\nu \in \mathscr{M}^+(\mathbb{R}^d)$, for ν-almost every $x_0 \in \mathbb{R}^d$ in the support of ν, the limit*

$$\frac{d\mu}{d\nu}(x_0) := \lim_{r \downarrow 0} \frac{\mu(B(x_0,r))}{\nu(B(x_0,r))}$$

*exists in $\mathbb{R}^N$ and is called the **Radon–Nikodým derivative** of μ with respect to ν, Moreover, the **Lebesgue–Radon–Nikodým decomposition** of μ is given as*

$$\mu = \frac{d\mu}{d\nu}\nu + \mu^s.$$

Here, $\mu^s = \mu \llcorner E$ is ***singular*** *with respect to ν (that is, μ^s is concentrated on a ν-negligible set), where*

$$E := (\mathbb{R}^d \setminus \operatorname{supp} \nu) \cup \left\{ x \in \operatorname{supp} \nu \ : \ \lim_{r \downarrow 0} \frac{|\mu|(B(x,r))}{\nu(B(x,r))} = \infty \right\}.$$

Finally, for a Borel measure $\mu \in \mathcal{M}(\Omega; \mathbb{R}^N)$ and a surjective Borel map $\varphi \colon \Omega \to \Omega' \subset \mathbb{R}^n$, we define the **push-forward measure** of μ under φ via

$$\varphi_\# \mu := \mu \circ \varphi^{-1} \in \mathcal{M}(\Omega'; \mathbb{R}^N).$$

We have the following transformation formula for any $g \colon \Omega' \to \mathbb{R}$:

$$\int_{\Omega'} g \ \mathrm{d}(\varphi_\# \mu) = \int_{\Omega} g \circ \varphi \ \mathrm{d}\mu,$$

provided these integrals are defined.

A.5 Sobolev and Other Function Spaces

We give a brief overview of Sobolev spaces, see [176] or [111] for more detailed accounts and proofs.

In all of the following we assume that $\Omega \subset \mathbb{R}^d$ is a Lipschitz domain, that is, Ω is open, bounded, connected, and has a boundary that is the union of finitely many Lipschitz manifolds. We also let $p \in [1, \infty]$, unless otherwise indicated. As usual, we denote by $\mathrm{C}(\Omega) = \mathrm{C}^0(\Omega)$, $\mathrm{C}^k(\Omega)$, $k = 1, 2, \ldots$, the spaces of continuous and k times continuously differentiable functions. The spaces $\mathrm{C}^k(\overline{\Omega})$ contain the $\mathrm{C}^k(\Omega)$-functions such that all l'th-order derivatives for $l \leq k$ can be continuously extended to $\overline{\Omega}$. As norms in these spaces we have

$$\|u\|_{\mathrm{C}^k} := \sum_{|\alpha| \leq k} \|\partial^\alpha u\|_\infty, \qquad u \in \mathrm{C}^k(\Omega), \qquad k = 0, 1, 2, \ldots,$$

where $\|\bullet\|_\infty$ is the supremum norm. Here, the sum is over all **multi-indices** $\alpha \in (\mathbb{N} \cup \{0\})^d$ with $|\alpha| := \alpha_1 + \cdots + \alpha_d \leq k$, and

$$\partial^\alpha := \partial_1^{\alpha_1} \partial_2^{\alpha_2} \cdots \partial_d^{\alpha_d}$$

is the α-derivative operator.

Similarly, we define the linear space $\mathrm{C}^\infty(\Omega)$ of infinitely-often differentiable functions, but this cannot be equipped with a complete norm. A subscript "$_c$" indicates that all functions u in the respective function space (e.g. $\mathrm{C}_c^\infty(\Omega)$) must have their **support**

$$\operatorname{supp} u := \overline{\{x \in \Omega : u(x) \neq 0\}}$$

compactly contained in Ω (so, $\operatorname{supp} u \subset \Omega$ and $\operatorname{supp} u$ compact). For the compact containment of a bounded set A in an open set B we write $A \Subset B$, which means that $\overline{A} \subset B$. In this way, the previous condition could be written as $\operatorname{supp} u \Subset \Omega$. We denote by $\mathrm{C}_0^k(\Omega)$ the closure of $\mathrm{C}_c^\infty(\Omega)$ in $\mathrm{C}^k(\Omega)$. All the spaces $\mathrm{C}_0^k(\Omega)$ are separable.

For $k \in \mathbb{N}$ a positive integer and $p \in [1, \infty]$, the **Sobolev space** $\mathrm{W}^{k,p}(\Omega)$ is defined to contain all functions $u \in \mathrm{L}^p(\Omega)$ such that the **weak derivative** $\partial^\alpha u$ exists and lies in $\mathrm{L}^p(\Omega)$ for all multi-indices $\alpha \in (\mathbb{N} \cup \{0\})^d$ with $|\alpha| \leq k$. This means that for every such α, there is a (unique) function $v_\alpha \in \mathrm{L}^p(\Omega)$ satisfying

$$\int v_\alpha \cdot \psi \,\mathrm{d}x = (-1)^{|\alpha|} \int u \cdot \partial^\alpha \psi \,\mathrm{d}x \qquad \text{for all } \psi \in \mathrm{C}_c^\infty(\Omega),$$

and we write $\partial^\alpha u$ for this v_α. The uniqueness follows from the Fundamental Lemma 3.10 of the calculus of variations. Clearly, if $u \in \mathrm{C}^k(\Omega)$, then all k'th-order weak derivatives coincide with their classical counterparts. As norm in $\mathrm{W}^{k,p}(\Omega)$, $p \in [1, \infty)$, we use

$$\|u\|_{\mathrm{W}^{k,p}} := \Bigl(\sum_{|\alpha| \leq k} \|\partial^\alpha u\|_{\mathrm{L}^p}^p \Bigr)^{1/p}, \qquad u \in \mathrm{W}^{k,p}(\Omega).$$

For $p = \infty$, we set

$$\|u\|_{\mathrm{W}^{k,\infty}} := \max_{|\alpha| \leq k} \|\partial^\alpha u\|_{\mathrm{L}^\infty}, \qquad u \in \mathrm{W}^{k,\infty}(\Omega).$$

Under these norms, the sets $\mathrm{W}^{k,p}(\Omega)$ become Banach spaces.

For $u \in \mathrm{W}^{1,p}(\Omega)$ we further define the **weak gradient** and **weak divergence**,

$$\nabla u := (\partial_1 u, \partial_2 u, \ldots, \partial_d u), \qquad \operatorname{div} u := \partial_1 u + \partial_2 u + \cdots + \partial_d u.$$

Concerning the boundary values of Sobolev functions we have:

Theorem A.24 (Trace). *For $p \in [1, \infty]$ there exists a linear **trace operator***

$$\mathrm{tr}_\Omega \colon \mathrm{W}^{1,p}(\Omega) \to \mathrm{L}^p(\partial\Omega)$$

such that

$$\mathrm{tr}_\Omega(\varphi) = \varphi|_{\partial\Omega} \qquad \text{if } \varphi \in \mathrm{C}(\overline{\Omega}).$$

We write $\mathrm{tr}_\Omega(u)$ simply as $u|_{\partial\Omega}$. For $p \in (1, \infty)$ the operator tr_Ω is bounded and weakly continuous between $\mathrm{W}^{1,p}(\Omega)$ and $\mathrm{L}^p(\partial\Omega)$.

For $p \in (1, \infty)$ denote the image of $\mathrm{W}^{1,p}(\Omega)$ under tr_Ω by $\mathrm{W}^{1-1/p,p}(\partial\Omega)$, which is called the **trace space** of $\mathrm{W}^{1,p}(\Omega)$. The norms on $\mathrm{W}^{1-1/p,p}(\partial\Omega)$ involve fractional derivatives, see [176] for details. For $p = 1$, the trace space is $\mathrm{L}^1(\partial\Omega, \mathscr{H}^{d-1} \llcorner \partial\Omega)$, which we will denote by just $\mathrm{L}^1(\partial\Omega)$.

We write $\mathrm{W}^{1,p}_0(\Omega)$ for the linear subspace of $\mathrm{W}^{1,p}(\Omega)$ consisting of all $\mathrm{W}^{1,p}$-functions with zero boundary values (in the sense of trace). More generally, we use $\mathrm{W}^{1,p}_g(\Omega)$ with $g \in \mathrm{W}^{1-1/p,p}(\partial\Omega)$ (with the convention $\mathrm{W}^{0,1}(\partial\Omega) = \mathrm{L}^1(\partial\Omega)$ in the case $p = 1$), for the affine subspace of all $\mathrm{W}^{1,p}$-functions with boundary trace g.

The following are some properties of Sobolev spaces, stated for simplicity only for the first-order space $\mathrm{W}^{1,p}(\Omega)$.

Theorem A.25 **(Extension).** *Every $u \in \mathrm{W}^{1,p}(\Omega)$ can be extended to $\bar{u} \in \mathrm{W}^{1,p}(\mathbb{R}^d)$ with $\|\bar{u}\|_{\mathrm{W}^{1,p}(\mathbb{R}^d)} \le C\|u\|_{\mathrm{W}^{1,p}(\mathbb{R}^d)}$, where $C = C(\Omega, p) > 0$ is a constant.*

Theorem A.26 **(Poincaré inequalities).** *Let $u \in \mathrm{W}^{1,p}(\Omega)$.*

(i) If $u|_{\partial\Omega} = 0$, then

$$\|u\|_{\mathrm{L}^p} \le C\|\nabla u\|_{\mathrm{L}^p},$$

where $C = C(\Omega, p) > 0$ is a constant.

(ii) Setting $[u]_\Omega := \fint_\Omega u \, dx$, it furthermore holds that

$$\|u - [u]_\Omega\|_{\mathrm{L}^p} \le C\|\nabla u\|_{\mathrm{L}^p},$$

where $C = C(\Omega, p) > 0$ is a constant.

Theorem A.27 **(Sobolev embedding).** *Let $u \in \mathrm{W}^{1,p}(\Omega)$.*

(i) If $p < d$, then $u \in \mathrm{L}^{p^*}(\Omega)$, where

$$p^* := \frac{dp}{d-p},$$

and there is a constant $C = C(\Omega, p) > 0$ such that

$$\|u\|_{\mathrm{L}^{p^*}} \le C\|u\|_{\mathrm{W}^{1,p}}.$$

(ii) If $p = d$, then $u \in \mathrm{L}^q(\Omega)$ for all $1 \le q < \infty$ and

$$\|u\|_{\mathrm{L}^q} \le C\|u\|_{\mathrm{W}^{1,p}},$$

where $C = C(\Omega, p, q) > 0$ is a constant.

(iii) If $p > d$, then $u \in \mathrm{C}(\Omega)$ and

$$\|u\|_\infty \le C\|u\|_{\mathrm{W}^{1,p}},$$

where $C = C(\Omega, p) > 0$ is a constant.

The second part can in fact be made more precise by considering embeddings into Hölder spaces, see Section 5.6.3 in [111] for details.

Theorem A.28 **(Rellich–Kondrachov).** *Let* $(u_j) \subset \mathrm{W}^{1,p}(\Omega)$ *with* $u_j \rightharpoonup u$ *in* $\mathrm{W}^{1,p}$.

(i) *If* $p < d$, *then* $u_j \to u$ *in* $\mathrm{L}^q(\Omega)$ *for any* $q < p^* = dp/(d-p)$.
(ii) *If* $p = d$, *then* $u_j \to u$ *in* $\mathrm{L}^q(\Omega)$ *for any* $q < \infty$.
(iii) *If* $p > d$, *then* $u_j \to u$ *uniformly (i.e., in the supremum norm).*

Theorem A.29 **(Density).** *For every* $p \in [1, \infty)$, $u \in \mathrm{W}^{1,p}(\Omega)$, *and all* $\varepsilon > 0$ *there exists a map* $v \in (\mathrm{W}^{1,p} \cap \mathrm{C}^\infty)(\Omega)$ *with* $v|_{\partial\Omega} = u|_{\partial\Omega}$ *and* $\|u - v\|_{\mathrm{W}^{1,p}} < \varepsilon$. *Moreover, there also exists a countably piecewise affine* $w \in (\mathrm{W}^{1,p} \cap \mathrm{C})(\Omega)$ *with* $w|_{\partial\Omega} = u|_{\partial\Omega}$ *and* $\|u - w\|_{\mathrm{W}^{1,p}} < \varepsilon$.

Here, a map $w \colon D \to \mathbb{R}$ is called **countably piecewise affine** if there exists a disjoint partition of Ω into countably many open sets D_k ($k \in \mathbb{N}$), up to a negligible set, i.e., $\Omega = Z \cup \bigcup_k D_k$, where $|Z| = 0$, such that $w|_{D_k}$ is affine.

For $0 < \gamma \le 1$ a function $u \colon \Omega \to \mathbb{R}$ is γ**-Hölder-continuous function**, in symbols $u \in \mathrm{C}^{0,\gamma}(\Omega)$, if

$$\|u\|_{\mathrm{C}^{0,\gamma}} := \|u\|_\infty + \sup_{\substack{x,y\in\Omega \\ x\neq y}} \frac{|u(x) - u(y)|}{|x-y|^\gamma} < \infty.$$

Functions in $\mathrm{C}^{0,1}(\Omega)$ are called **Lipschitz continuous**. We construct the higher-order spaces $\mathrm{C}^{k,\gamma}(\Omega)$ analogously.

Theorem A.30 **(Rademacher).** *Let* $\Omega \subset \mathbb{R}^d$ *be an open, bounded, and convex set. Then, the space* $\mathrm{W}^{1,\infty}(\Omega)$ *consists precisely of all Lipschitz maps on* Ω, *the Lipschitz constant is equal to the* $\mathrm{W}^{1,\infty}$*-norm, and for* $u \in \mathrm{W}^{1,\infty}(\Omega)$ *the classical gradient* ∇u *exists almost everywhere in* Ω *and agrees with the weak gradient.*

Next, we define a **family of mollifiers** as follows: Let $\eta \in \mathrm{C}_c^\infty(\mathbb{R}^d)$ be radially symmetric and positive. Then, the family $(\eta_\delta)_{\delta>0}$ is defined as follows:

$$\eta_\delta(x) := \frac{1}{\delta^d}\eta\left(\frac{x}{\delta^d}\right), \qquad x \in \mathbb{R}^d.$$

For $u \in \mathrm{W}^{k,p}(\mathbb{R}^d)$, where $k \in \mathbb{N} \cup \{0\}$, and $p \in [1, \infty]$, we define the **mollification** $u_\delta \in \mathrm{W}^{k,p}(\mathbb{R}^d)$ of u as the **convolution** between η_δ and u, i.e.,

$$u_\delta(x) := (\eta_\delta \star u)(x) := \int \eta_\delta(x-y)u(y)\,\mathrm{d}y, \qquad x \in \mathbb{R}^d.$$

Lemma A.31 *For every* $p \in [1, \infty)$, *if* $u \in \mathrm{W}^{1,p}(\mathbb{R}^d)$, *then* $u_\delta \to u$ *in* $\mathrm{W}^{1,p}$ *as* $\delta \downarrow 0$.

Analogous results also hold for continuously differentiable functions.

Lemma A.32 **(Young's inequality for convolutions).** *Let* $u \in \mathrm{L}^p(\mathbb{R}^d)$, $v \in \mathrm{L}^q(\mathbb{R}^d)$ *and let* $p, q, r \in [1, \infty]$ *be such that*

$$1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}.$$

Then,

$$\|u \star v\|_{\mathrm{L}^r} \leq \|u\|_{\mathrm{L}^p} \cdot \|v\|_{\mathrm{L}^q}.$$

Finally, all the above notions and theorems continue to hold for vector-valued functions $u = (u^1, \ldots, u^m)^T : \Omega \to \mathbb{R}^m$ and in this case we set

$$\nabla u := \begin{pmatrix} \partial_1 u^1 & \partial_2 u^1 & \cdots & \partial_d u^1 \\ \partial_1 u^2 & \partial_2 u^2 & \cdots & \partial_d u^2 \\ \vdots & \vdots & & \vdots \\ \partial_1 u^m & \partial_2 u^m & \cdots & \partial_d u^m \end{pmatrix}.$$

We use the spaces $\mathrm{C}(\Omega; \mathbb{R}^m)$, $\mathrm{C}^k(\Omega; \mathbb{R}^m)$, $\mathrm{W}^{k,p}(\Omega; \mathbb{R}^m)$, $\mathrm{C}^{k,\gamma}(\Omega; \mathbb{R}^m)$ with analogous definitions as in the scalar-valued case; for matrices like ∇u we use the Frobenius matrix norm and similarly for higher-order tensors.

Occasionally, we employ *local* versions of the spaces defined above, namely $\mathrm{C}_{\mathrm{loc}}(\Omega)$, $\mathrm{C}^k_{\mathrm{loc}}(\Omega)$, $\mathrm{W}^{k,p}_{\mathrm{loc}}(\Omega)$, $\mathrm{C}^{k,\gamma}_{\mathrm{loc}}(\Omega)$, where the defining norm is only finite on every compact subset of Ω.

Finally, we quote the following two classical results about extensions of functions:

Theorem A.33 **(Tietze).** *Let* X *be a metric space, let* $F \subset X$ *be closed, and assume that* $f : F \to \mathbb{R}^m$ *is continuous. Then,* f *can be extended to a continuous* $\bar{f} : X \to \mathbb{R}^m$. *If* f *is bounded, then* $\bar{f}$ *can also be chosen as bounded.*

Theorem A.34 **(Kirszbraun).** *Let* $\Omega \subset \mathbb{R}^d$ *and let* $f : \Omega \to \mathbb{R}^m$ *be a Lipschitz continuous map. Then,* f *can be extended to* $\bar{f} : \mathbb{R}^d \to \mathbb{R}^m$ *with the same Lipschitz constant as* f.

A.6 Harmonic Analysis

In this book we only need a few basics of Fourier analysis and the Mihlin multiplier theorem. A thorough introduction can be found in [138, 139].

Define for $u \in \mathrm{L}^1(\mathbb{R}^d)$ (or vector-valued u) the **Fourier transform** $\hat{u} = \mathscr{F}u \in \mathrm{L}^\infty(\mathbb{R}^d)$ as follows:

$$\hat{u}(\xi) := \mathscr{F}u(\xi) := \int_{\mathbb{R}^d} u(x) \mathrm{e}^{-2\pi \mathrm{i} x \cdot \xi} \, \mathrm{d}x, \qquad \xi \in \mathbb{R}^d.$$

We also define the **inverse Fourier transform** $\check{v} = \mathscr{F}^{-1} v$ for $v \in \mathrm{L}^1(\mathbb{R}^d)$ to be

$$\check{u}(x) := \mathscr{F}^{-1}v(x) := \int_{\mathbb{R}^d} v(\xi)\mathrm{e}^{2\pi \mathrm{i} x\cdot\xi}\,\mathrm{d}x, \qquad x \in \mathbb{R}^d.$$

One can extend $\mathscr{F}$, $\mathscr{F}^{-1}$ to the space $\mathrm{L}^2(\mathbb{R}^d)$ via the **Plancherel identity**,

$$\|\hat{u}\|_{\mathrm{L}^2} = \|u\|_{\mathrm{L}^2}. \tag{A.4}$$

Moreover, we have the **Parseval relation**

$$\int u \cdot \overline{v}\,\mathrm{d}x = \int \hat{u} \cdot \overline{\hat{v}}\,\mathrm{d}\xi \tag{A.5}$$

for all $u, v \in \mathrm{L}^2(\mathbb{R}^d)$; the same relations hold for $\mathbb{C}^N$-valued functions.

The following is a classical result concerning the $(\mathrm{L}^p \to \mathrm{L}^p)$-boundedness of Fourier multiplier operators, see, for instance, [38, 138] (Theorem 6.1.6) for a proof.

Theorem A.35 (**Mihlin multiplier theorem**). *Let $m \in \mathrm{C}^{\lfloor d/2\rfloor+1}(\mathbb{R}^d \setminus \{0\}; \mathbb{C})$ satisfy*

$$|\partial^\alpha m(\xi)| \le K|\xi|^{-|\alpha|}, \qquad \xi \in \mathbb{R}^d \setminus \{0\},$$

for all multi-indices $\alpha \in \mathbb{N}_0^d$ with $|\alpha| := |\alpha_1| + \cdots + |\alpha_d| \le \lfloor d/2\rfloor + 1$ ($\lfloor t\rfloor$ denotes the largest integer less than or equal to $t \in \mathbb{R}$) and some $K > 0$. Then,

$$Tu := \mathscr{F}^{-1}[m(\xi)\hat{u}(\xi)],$$

which for $u \in \mathrm{L}^2(\mathbb{R}^d)$ is well-defined via the Plancherel identity (A.4), extends to a bounded operator $T\colon \mathrm{L}^p(\mathbb{R}^d) \to \mathrm{L}^p(\mathbb{R}^d)$ for all $p \in (1, \infty)$, which satisfies the estimate

$$\|T\|_{\mathrm{L}^p\to\mathrm{L}^p} \le C\max\{p, (p-1)^{-1}\}K,$$

where $C = C(d) > 0$ is a constant. Furthermore, for $p = 1$ the weak-type estimate

$$\left|\left\{ x \in \mathbb{R}^d \;:\; |(Tu)(x)| \ge t \right\}\right| \le \frac{CK}{t}\|u\|_{\mathrm{L}^1}$$

holds for all $t > 0$ and a constant $C = C(d) > 0$.

As a special case, the conclusions of the preceding theorem hold for any positively 0-homogeneous smooth multiplier $m\colon \mathbb{R}^d \setminus \{0\} \to \mathbb{C}$.

We will also use the (centered) **maximal function** $Mf\colon \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ of $f\colon \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, which is defined as

$$(Mf)(x_0) := \sup_{r>0} \fint_{B(x_0,r)} |f(x)|\,\mathrm{d}x, \qquad x_0 \in \mathbb{R}^d.$$

We quote the following results about the maximal function, whose proofs can be found in [138, 177, 247] (in particular, (iii) is essentially contained in Lemma 1.68 of [177]):

Theorem A.36 *The following statements are true:*

(i) If $p \in (1, \infty]$, *then*
$$\|Mf\|_{\mathrm{L}^p} \leq C\|f\|_{\mathrm{L}^p},$$
where $C = C(d, p) > 0$ *is a constant.*

(ii) If $p \in [1, \infty)$, *then the weak-type estimate*
$$\left|\left\{ x \in \mathbb{R}^d \ \boldsymbol{:}\ |Mf| \geq t \right\}\right| \leq \frac{C}{t^p} \int_{\{|f| \geq t/2\}} |f|^p \, dx \leq \frac{C}{t^p}\|u\|_{\mathrm{L}^p}^p$$
holds for all $t > 0$ *and a constant* $C = C(d) > 0$.

(iii) For every $K > 0$ *and* $f \in \mathrm{W}^{1,p}(\mathbb{R}^d; \mathbb{R}^m)$, $p \in (1, \infty]$, *the maximal function* Mf *is Lipschitz continuous on the set* $\{M(|f| + |\nabla f|) < K\}$ *and its Lipschitz constant is bounded by* CK, *where* $C = C(d, m, p) > 0$ *is a constant.*

References

1. Acerbi, E., Fusco, N.: Semicontinuity problems in the calculus of variations. Arch. Ration. Mech. Anal. **86**, 125–145 (1984)
2. Acerbi, E., Fusco, N.: A regularity theorem for minimizers of quasiconvex integrals. Arch. Ration. Mech. Anal. **99**, 261–281 (1987)
3. Acerbi, E., Fusco, N.: An approximation lemma for $W^{1,p}$ functions. Material Instabilities in Continuum Mechanics (Edinburgh. 1985–1986), Oxford Science Publications, pp. 1–5. Oxford University Press, Oxford (1988)
4. Alberti, G.: Rank one property for derivatives of functions with bounded variation. Proc. R. Soc. Edinb. Sect. A **123**, 239–274 (1993)
5. Alberti, G., Csörnyei, M., Preiss, D.: Structure of null sets in the plane and applications. In: Proceedings of the Fourth European Congress of Mathematics (Stockholm, 2004), pp. 3–22. European Mathematical Society (2005)
6. Alibert, J.J., Bouchitté, G.: Non-uniform integrability and generalized Young measures. J. Convex Anal. **4**, 129–147 (1997)
7. Alibert, J.J., Dacorogna, B.: An example of a quasiconvex function that is not polyconvex in two dimensions. Arch. Ration. Mech. Anal. **117**, 155–166 (1992)
8. Allaire, G.: Shape Optimization by the Homogenization Method. Applied Mathematical Sciences, vol. 146. Springer, Berlin (2002)
9. Allard, W.K.: On the first variation of a varifold. Ann. Math. **95**, 417–491 (1972)
10. Almgren Jr., F.J.: Existence and regularity almost everywhere of solutions to elliptic variational problems among surfaces of varying topological type and singularity structure. Ann. Math. **87**, 321–391 (1968)
11. Almgren, Jr., F.J.: Plateau's Problem: An Invitation to Varifold Geometry, Revised Edition. Student Mathematical Library, vol. 13. American Mathematical Society, New York (2001)
12. Ambrosio, L., Coscia, A., Dal Maso, G.: Fine properties of functions with bounded deformation. Arch. Ration. Mech. Anal. **139**, 201–238 (1997)
13. Ambrosio, L., Dal Maso, G.: On the relaxation in $\mathrm{BV}(\Omega; R^m)$ of quasi-convex integrals. J. Funct. Anal. **109**, 76–97 (1992)
14. Ambrosio, L., De Giorgi, E.: Un nuovo tipo di funzionale del calcolo delle variazioni. Atti Acc. Naz. dei Lincei, Rend. Cl. Sc. Fis. Mat. Natur. **82**, 199–210 (1988)
15. Ambrosio, L., Fusco, N., Pallara, D.: Functions of Bounded Variation and Free-Discontinuity Problems. Oxford Mathematical Monographs. Oxford University Press, Oxford (2000)
16. Arroyo-Rabasa, A., De Philippis, G., Rindler, F.: Lower semicontinuity and relaxation of linear-growth integral functionals under PDE constraints. Adv. Calc. Var. (2017). To appear, arXiv:1701.02230

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8

17. Astala, K., Faraco, D.: Quasiregular mappings and Young measures. Proc. R. Soc. Edinb. Sect. A **132**, 1045–1056 (2002)
18. Astala, K., Iwaniec, T., Prause, I., Saksman, E.: Burkholder integrals, Morrey's problem and quasiconformal mappings. J. Am. Math. Soc. **25**, 507–531 (2012)
19. Attouch, H., Buttazzo, G., Michaille, G.: Variational Analysis in Sobolev and BV Spaces - Applications to PDEs and Optimization. Society for Industrial and Applied Mathematics, MPS-SIAM Series on Optimization (2006)
20. Aubin, J.P., Cellina, A.: Differential Inclusions. Set-Valued Maps and Viability Theory. Grundlehren der mathematischen Wissenschaften, vol. 264. Springer, Berlin (1984)
21. Aumann, R.J., Hart, S.: Bi-convexity and bi-martingales. Israel J. Math. **54**, 159–180 (1986)
22. Auslender, A., Teboulle, M.: Asymptotic Cones and Functions in Optimization and Variational Inequalities. Springer Monographs in Mathematics. Springer, Berlin (2003)
23. Baía, M., Krömer, S., Kružík, M.: Generalized $W^{1,1}$-Young measures and relaxation of problems with linear growth. SIAM J. Math. Anal. **50**, 1076–1119 (2018)
24. Balder, E.J.: A general approach to lower semicontinuity and lower closure in optimal control theory. SIAM J. Control Optim. **22**, 570–598 (1984)
25. Ball, J.M.: Convexity conditions and existence theorems in nonlinear elasticity. Arch. Ration. Mech. Anal. **63**, 337–403 (1976/77)
26. Ball, J.M.: Global invertibility of Sobolev functions and the interpenetration of matter. Proc. R. Soc. Edinb. Sect. A **88**, 315–328 (1981)
27. Ball, J.M.: A version of the fundamental theorem for Young measures. PDEs and Continuum Models of Phase Transitions (Nice, 1988). Lecture Notes in Physics, vol. 344, pp. 207–215. Springer, Berlin (1989)
28. Ball, J.M.: Some open problems in elasticity. Geometry. Mechanics, and Dynamics, pp. 3–59. Springer, Berlin (2002)
29. Ball, J.M., Currie, J.C., Olver, P.J.: Null Lagrangians, weak continuity, and variational problems of arbitrary order. J. Funct. Anal. **41**, 135–174 (1981)
30. Ball, J.M., James, R.D.: Fine phase mixtures as minimizers of energy. Arch. Ration. Mech. Anal. **100**, 13–52 (1987)
31. Ball, J.M., James, R.D.: Proposed experimental tests of a theory of fine microstructure and the two-well problem. Phil. Trans. R. Soc. Lond. A **338**, 389–450 (1992)
32. Ball, J.M., James, R.D.: Incompatible sets of gradients and metastability. Arch. Ration. Mech. Anal. **218**, 1363–1416 (2015)
33. Ball, J.M., Kirchheim, B., Kristensen, J.: Regularity of quasiconvex envelopes. Calc. Var. Partial Differ. Equ. **11**, 333–359 (2000)
34. Ball, J.M., Mizel, V.J.: One-dimensional variational problems whose minimizers do not satisfy the Euler-Lagrange equation. Arch. Ration. Mech. Anal. **90**, 325–388 (1985)
35. Ball, J.M., Murat, F.: Remarks on Chacon's biting lemma. Proc. Am. Math. Soc. **107**, 655–663 (1989)
36. Beck, L.: Elliptic Regularity Theory. A First Course. Lecture Notes of the Unione Matematica Italiana, vol. 19. Springer, Berlin (2016)
37. Weak lower semicontinuity of integral functionals and applications: Benešová, B., Kruží k, M. SIAM Rev. **59**, 703–766 (2017)
38. Bergh, J., Löfström, J.: Interpolation Spaces. Grundlehren der mathematischen Wissenschaften, vol. 223. Springer, Berlin (1976)
39. Berliocchi, H., Lasry, J.M.: Intégrandes normales et mesures paramétrées en calcul des variations. Bull. Soc. Math. France **101**, 129–184 (1973)
40. Bernoulli, J.: Problema novum ad cujus solutionem Mathematici invitantur. In: Acta Eruditorum, p. 269. Lipsiæ, apud J. Grossium & J.F. Gletitschium (1696). https://hdl.handle.net/2027/mdp.39015067096282?urlappend=%3Bseq=285. Hathi Trust Digital Library. Accessed 15 May 2017
41. Bhattacharya, K.: Microstructure of Martensite. Oxford Series on Materials Modelling. Oxford University Press, Oxford (2003)

42. Bhattacharya, K., Dolzmann, G.: Relaxation of some multi-well problems. Proc. R. Soc. Edinb. Sect. A **131**, 279–320 (2001)
43. Bhattacharya, K., Firoozye, N.B., James, R.D., Kohn, R.V.: Restrictions on microstructure. Proc. R. Soc. Edinb. Sect. A **124**, 843–878 (1994)
44. Bildhauer, M.: Convex Variational Problems. Lecture Notes in Mathematics, vol. 1818. Springer, Berlin (2003)
45. Blåsjö, V.: The isoperimetric problem. Am. Math. Monthly **112**, 526–566 (2005)
46. Bojarski, B., Iwaniec, T.: Analytical foundations of the theory of quasiconformal mappings in $\mathbf{R}^n$. Ann. Acad. Sci. Fenn. Ser. A I Math. **8**, 257–324 (1983)
47. Bouchitté, G., Fonseca, I., Malý, J.: The effective bulk energy of the relaxed energy of multiple integrals below the growth exponent. Proc. R. Soc. Edinb. Sect. A **128**(3), 463–479 (1998)
48. Bouchitté, G., Fonseca, I., Mascarenhas, L.: A global method for relaxation. Arch. Ration. Mech. Anal. **145**, 51–98 (1998)
49. Bousquet, P., Mariconda, C., Treu, G.: On the Lavrentiev phenomenon for multiple integral scalar variational problems. J. Funct. Anal. **266**, 5921–5954 (2014)
50. Braides, A.: Homogenization of some almost periodic coercive functional. Rend. Accad. Naz. Sci. XL Mem. Mat. **9**, 313–321 (1985)
51. Braides, A.: Γ-Convergence for Beginners. Oxford Lecture Series in Mathematics and its Applications, vol. 22. Oxford University Press, Oxford (2002)
52. Braides, A., Defranceschi, A.: Homogenization of Multiple Integrals. Oxford Lecture Series in Mathematics and its Applications, vol. 12. Oxford University Press, Oxford (1998)
53. Brooks, J.K., Chacon, R.V.: Continuity and compactness of measures. Adv. Math. **37**, 16–26 (1980)
54. Cahn, J.W., Hilliard, J.E.: Free energy of a nonuniform system. I. Interfacial free energy. J. Chem. Phys. **28**, 258–267 (1958)
55. Campanato, S.: Hölder continuity of the solutions of some nonlinear elliptic systems. Adv. Math. **48**, 16–43 (1983)
56. Carstensen, C., Roubíček, T.: Numerical approximation of Young measures in non-convex variational problems. Numer. Math. **84**, 395–415 (2000)
57. Castaing, C., de Fitte, P.R., Valadier, M.: Young Measures on Topological Spaces: With Applications in Control Theory and Probability Theory. Mathematics and its Applications. Kluwer, Dordrecht (2004)
58. Cellina, A.: On the differential inclusion $x' \in [-1, +1]$. Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. **69**, 1–6 (1980)
59. Cellina, A., Perrotta, S.: On a problem of potential wells. J. Convex Anal. **2**, 103–115 (1995)
60. Chandler, R.E.: Hausdorff Compactifications. Lecture Notes in Pure and Applied Mathematics, vol. 23. Marcel Dekker (1976)
61. Chaudhuri, N., Müller, S.: Rigidity estimate for two incompatible wells. Calc. Var. Partial Differ. Equ. **19**, 379–390 (2004)
62. Chermisi, M., Conti, S.: Multiwell rigidity in nonlinear elasticity. SIAM J. Math. Anal. **42**, 1986–2012 (2010)
63. Chlebík, M., Kirchheim, B.: Rigidity for the four gradient problem. J. Reine Angew. Math. **551**, 1–9 (2002)
64. Ciarlet, P.G.: Mathematical Elasticity, vol. 1: Three Dimensional Elasticity. North-Holland (1988)
65. Ciarlet, P.G., Geymonat, G.: Sur les lois de comportement en élasticité non linéaire compressible. C. R. Acad. Sci. Paris Sér. II Méc. Phys. Chim. Sci. Univers. Sci. Terre **295**, 423–426 (1982)
66. Ciarlet, P.G., Gratie, L., Mardare, C.: Intrinsic methods in elasticity: a mathematical survey. Discret. Contin. Dyn. Syst. **23**, 133–164 (2009)
67. Ciarlet, P.G., Nečas, J.: Injectivity and self-contact in nonlinear elasticity. Arch. Ration. Mech. Anal. **97**, 171–188 (1987)
68. Clarke, F.: Functional Analysis, Calculus of Variations and Optimal Control. Graduate Texts in Mathematics, vol. 264. Springer, Berlin (2013)

69. Conti, S., Dolzmann, G., Kirchheim, B.: Existence of Lipschitz minimizers for the three-well problem in solid-solid phase transitions. Ann. Inst. H. Poincaré Anal. Non Linéaire **24**, 953–962 (2007)
70. Conti, S., Dolzmann, G., Kirchheim, B., Müller, S.: Sufficient conditions for the validity of the Cauchy-Born rule close to $\mathrm{SO}(n)$. J. Eur. Math. Soc. **8**, 515–530 (2006)
71. Conti, S., Faraco, D., Maggi, F.: A new approach to counterexamples to L^1 estimates: Korn's inequality, geometric rigidity, and regularity for gradients of separately convex functions. Arch. Ration. Mech. Anal. **175**(2), 287–300 (2005)
72. Conti, S., Fonseca, I., Leoni, G.: A Γ-convergence result for the two-gradient theory of phase transitions. Commun. Pure Appl. Math. **55**, 857–936 (2002)
73. Conti, S., Schweizer, B.: Rigidity and gamma convergence for solid-solid phase transitions with SO(2) invariance. Commun. Pure Appl. Math. **59**, 830–868 (2006)
74. Conway, J.B.: A Course in Functional Analysis. Graduate Texts in Mathematics, vol. 96, 2nd edn. Springer, Berlin (1990)
75. Dacorogna, B.: Weak Continuity and Weak Lower Semicontinuity for Nonlinear Functionals. Lecture Notes in Mathematics, vol. 922. Springer, Berlin (1982)
76. Dacorogna, B.: Direct Methods in the Calculus of Variations. Applied Mathematical Sciences, vol. 78, 2nd edn. Springer, Berlin (2008)
77. Dacorogna, B.: Introduction to the Calculus of Variations, 2nd edn. Imperial College Press, London (2009)
78. Dacorogna, B., Marcellini, P.: A counterexample in the vectorial calculus of variations. Material Instabilities in Continuum Mechanics (Edinburgh. 1985–1986), Oxford Science Publications, pp. 77–83. Oxford University Press, Oxford (1988)
79. Dacorogna, B., Marcellini, P.: General existence theorems for Hamilton-Jacobi equations in the scalar and vectorial cases. Acta Math. **178**, 1–37 (1997)
80. Dacorogna, B., Marcellini, P.: Implicit Partial Differential Equations. Progress in Nonlinear Differential Equations and their Applications, vol. 37. Birkhäuser (1999)
81. Dafermos, C.M.: Hyperbolic Conservation Laws in Continuum Physics. Grundlehren der mathematischen Wissenschaften, vol. 325, 3rd edn. Springer, Berlin (2010)
82. Dal Maso, G.: An Introduction to Γ-Convergence. Progress in Nonlinear Differential Equations and their Applications, vol. 8. Birkhäuser (1993)
83. Dal Maso, G., DeSimone, A., Mora, M.G., Morini, M.: Time-dependent systems of generalized Young measures. Netw. Heterog. Media **2**, 1–36 (2007)
84. Dal Maso, G., DeSimone, A., Mora, M.G., Morini, M.: Globally stable quasistatic evolution in plasticity with softening. Netw. Heterog. Media **3**, 567–614 (2008)
85. Dal Maso, G., DeSimone, A., Mora, M.G., Morini, M.: A vanishing viscosity approach to quasistatic evolution in plasticity with softening. Arch. Ration. Mech. Anal. **189**, 469–544 (2008)
86. Dal Maso, G., Modica, L.: Integral functionals determined by their minima. Rend. Sem. Mat. Univ. Padova **76**, 255–267 (1986)
87. De Giorgi, E.: Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. **3**, 25–43 (1957)
88. De Giorgi, E.: Un esempio di estremali discontinue per un problema variazionale di tipo ellittico. Boll. Un. Mat. Ital. **1**, 135–137 (1968)
89. De Giorgi, E., Franzoni, T.: Su un tipo di convergenza variazionale. Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. **58**, 842–850 (1975)
90. De Lellis, C.: A note on Alberti's rank-one theorem. Transport Tations and Multi-D Hyperbolic Conservation Laws (Bologna, 2005). Lecture Notes of the Unione Matematica Italiana, vol. 5, pp. 61–74. Springer, Berlin (2008)
91. De Lellis, C., Székelyhidi Jr., L.: The Euler equations as a differential inclusion. Ann. Math. **170**, 1417–1436 (2009)
92. De Philippis, G., Rindler, F.: On the structure of $\mathscr{A}$-free measures and applications. Ann. Math. **184**, 1017–1039 (2016)

93. De Philippis, G., Rindler, F.: Characterization of generalized Young measures generated by symmetric gradients. Arch. Ration. Mech. Anal. **224**, 1087–1125 (2017)
94. Delladio, S.: Lower semicontinuity and continuity of functions of measures with respect to the strict convergence. Proc. R. Soc. Edinb. Sect. A **119**, 265–278 (1991)
95. Demoulini, S.: Young measure solutions for nonlinear evolutionary systems of mixed type. Ann. Inst. H. Poincaré Anal. Non Linéaire **14**, 143–162 (1997)
96. Diestel, J., Uhl, Jr., J.J.: Vector Measures. Mathematical Surveys, vol. 15. American Mathematical Society (1977)
97. Dinculeanu, N.: Vector Measures. International Series of Monographs in Pure and Applied Mathematics, vol. 95. Pergamon Press, Oxford (1967)
98. DiPerna, R.J.: Compensated compactness and general systems of conservation laws. Trans. Am. Math. Soc. **292**, 383–420 (1985)
99. DiPerna, R.J., Majda, A.J.: Oscillations and concentrations in weak solutions of the incompressible fluid equations. Commun. Math. Phys. **108**, 667–689 (1987)
100. Dolzmann, G.: Variational Methods for Crystalline Microstructure - Analysis and Computation. Lecture Notes in Mathematics, vol. 1803. Springer, Berlin (2003)
101. Dolzmann, G., Kirchheim, B., Müller, S., Šverák, V.: The two-well problem in three dimensions. Calc. Var. Partial Differ. Equ. **10**, 21–40 (2000)
102. Dolzmann, G., Kristensen, J.: Higher integrability of minimizing Young measures. Calc. Var. Partial Differ. Equ. **22**, 283–301 (2005)
103. Dolzmann, G., Müller, S.: Microstructures with finite surface energy: the two-well problem. Arch. Ration. Mech. Anal. **132**, 101–141 (1995)
104. Duggin, M.J., Rachinger, W.A.: The nature of the martensite transformation in a coppernickel-aluminium alloy. Acta Metallurgica **12**, 529–535 (1964)
105. E, W., Ming, P.: Cauchy–Born rule and the stability of crystalline solids: static problems. Arch. Ration. Mech. Anal. **183**, 241–297 (2007)
106. Ekeland, I., Temam, R.: Convex Analysis and Variational Problems. North-Holland (1976)
107. Enami, K., Nenno, S.: Memory effect in Ni-36.8 at. pct. Al martensite. Metall. Trans. **2**, 1487 (1971)
108. Ericksen, J.L.: On the Cauchy-Born rule. Math. Mech. Solids **13**, 199–220 (2008)
109. Evans, L.C.: Quasiconvexity and partial regularity in the calculus of variations. Arch. Ration. Mech. Anal. **95**, 227–252 (1986)
110. Evans, L.C.: Weak Convergence Methods for Nonlinear Partial Differential Equations. CBMS Regional Conference Series in Mathematics, vol. 74. American Mathematical Society (1990)
111. Evans, L.C.: Partial Differential Equations. Graduate Studies in Mathematics, vol. 19, 2nd edn. American Mathematical Society (2010)
112. Evans, L.C., Gariepy, R.F.: Measure Theory and Fine Properties of Functions. Studies in Advanced Mathematics, 2nd edn. CRC Press, Boca Raton (2015)
113. Falconer, K.: Fractal Geometry, 3rd edn. Wiley, New Jersey (2014)
114. Faraco, D., Székelyhidi Jr., L.: Tartar's conjecture and localization of the quasiconvex hull in $\mathbb{R}^{2\times 2}$. Acta Math. **200**, 279–305 (2008)
115. Federer, H.: Geometric Measure Theory. Grundlehren der mathematischen Wissenschaften, vol. 153. Springer, Berlin (1969)
116. Federer, H., Fleming, W.H.: Normal and integral currents. Ann. Math. **72**, 458–520 (1960)
117. Ferriero, A.: The Lavrentiev phenomenon in the Calculus of Variations. Ph.D. thesis, Universitá degli Studi di Milano-Bicocca (2004)
118. Ferriero, A.: A direct proof of the Tonelli's partial regularity result. Discret. Contin. Dyn. Syst. **32**, 2089–2099 (2012)
119. Folland, G.B.: Real Analysis. Pure and Applied Mathematics, 2nd edn. Wiley, New Jersey (1999)
120. Fonseca, I.: The lower quasiconvex envelope of the stored energy function for an elastic crystal. J. Math. Pures Appl. **9**(67), 175–195 (1988)
121. Fonseca, I., Kružík, M.: Oscillations and concentrations generated by $\mathscr{A}$-free mappings and weak lower semicontinuity of integral functionals. ESAIM Control Optim. Calc. Var. **16**, 472–502 (2010)

122. Fonseca, I., Leoni, G.: Modern Methods in the Calculus of Variations: L^p Spaces. Springer, Berlin (2007)
123. Fonseca, I., Müller, S.: Quasi-convex integrands and lower semicontinuity in L^1. SIAM J. Math. Anal. **23**, 1081–1098 (1992)
124. Fonseca, I., Müller, S.: Relaxation of quasiconvex functionals in $\mathrm{BV}(\Omega, R^p)$ for integrands $f(x, u, \nabla u)$. Arch. Ration. Mech. Anal. **123**, 1–49 (1993)
125. Fonseca, I., Müller, S., Pedregal, P.: Analysis of concentration and oscillation effects generated by gradients. SIAM J. Math. Anal. **29**, 736–756 (1998)
126. Francfort, G.A.: An introduction to H-measures and their applications. In: Variational Problems in Materials Science (Trieste, 2004). Progress in Nonlinear Differential Equations and their Applications, vol. 68, pp. 85–110. Birkhäuser (2006)
127. Friesecke, G., James, R.D., Müller, S.: A theorem on geometric rigidity and the derivation of nonlinear plate theory from three-dimensional elasticity. Commun. Pure Appl. Math. **55**, 1461–1506 (2002)
128. Friesecke, G., Theil, F.: Validity and failure of the Cauchy-Born hypothesis in a two-dimensional mass-spring lattice. J. Nonlinear Sci. **12**, 445–478 (2002)
129. Gelfand, I.M., Fomin, S.V.: Calculus of Variations. Prentice-Hall, Revised English edition translated and edited by Richard A. Silverman (1963)
130. Gérard, P.: Microlocal defect measures. Commun. Partial Differ. Equ. **16**, 1761–1794 (1991)
131. Giaquinta, M., Hildebrandt, S.: Calculus of Variations. I – The Lagrangian Formalism. Grundlehren der mathematischen Wissenschaften, vol. 310. Springer, Berlin (1996)
132. Giaquinta, M., Hildebrandt, S.: Calculus of Variations. II – The Hamiltonian Formalism. Grundlehren der mathematischen Wissenschaften, vol. 311. Springer, Berlin (1996)
133. Giaquinta, M., Martinazzi, L.: An introduction to the regularity theory for elliptic systems, harmonic maps and minimal graphs. Lecture Notes. Scuola Normale Superiore di Pisa (New Series), vol. 11, 2nd edn. Edizioni della Normale, Pisa (2012)
134. Giaquinta, M., Modica, G., Souček, J.: Cartesian Currents in the Calculus of Variations. I. Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 37. Springer, Berlin (1998)
135. Giaquinta, M., Modica, G., Souček, J.: Cartesian Currents in the Calculus of Variations. II. Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 38. Springer, Berlin (1998)
136. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order. Grundlehren der mathematischen Wissenschaften, vol. 224. Springer, Berlin (1998)
137. Giusti, E.: Direct Methods in the Calculus of Variations. World Scientific, Singapore (2003)
138. Grafakos, L.: Classical Fourier Analysis. Graduate Texts in Mathematics, vol. 249, 3rd edn. Springer, Berlin (2014)
139. Grafakos, L.: Modern Fourier Analysis. Graduate Texts in Mathematics, vol. 250, 3rd edn. Springer, Berlin (2014)
140. Gratwick, R.: Singular sets and the Lavrentiev phenomenon. Proc. R. Soc. Edinb. Sect. A **145**, 513–533 (2015)
141. Gratwick, R.: Variations, approximation, and low regularity in one dimension. Calc. Var. Partial Differ. Equ. (2017). To appear
142. Gratwick, R., Preiss, D.: A one-dimensional variational problem with continuous Lagrangian and singular minimizer. Arch. Ration. Mech. Anal. **202**, 177–211 (2011)
143. Gratwick, R., Sychev, M.A., Tersenov, A.S.: Regularity and singularity phenomena for one-dimensional variational problems with singular ellipticity. Pure Appl. Funct. Anal. **1**, 397–416 (2016)
144. Gromov, M.: Convex integration of differential relations. I. Izv. Akad. Nauk SSSR Ser. Mat. **37**, 329–343 (1973)
145. Gromov, M.: Partial Differential Relations. Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 9. Springer, Berlin (1986)
146. Gurtin, M.E.: On phase transitions with bulk, interfacial, and boundary energy. Arch. Ration. Mech. Anal. **96**, 243–264 (1986)
147. Gurtin, M.E.: Some results and conjectures in the gradient theory of phase transitions. In: Metastability and Incompletely Posed Problems (Minneapolis, Minn., 1985). The IMA Volumes in Mathematics and its Applications, vol. 3, pp. 135–146. Springer, Berlin (1987)

148. Henao, D., Mora-Corral, C.: Invertibility and weak continuity of the determinant for the modelling of cavitation and fracture in nonlinear elasticity. Arch. Ration. Mech. Anal. **197**, 619–655 (2010)
149. Hilbert, D.: Mathematische Probleme – Vortrag, gehalten auf dem internationalen Mathematiker-Kongreß zu Paris 1900. Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse pp. 253–297 (1900)
150. Hildebrandt, S., Tromba, A.: Mathematics and Optimal Form. Scientific American Library (1985)
151. Horn, R.A., Johnson, C.R.: Matrix Analysis, 2nd edn. Cambridge University Press, Cambridge (2013)
152. Isett, P.: A Proof of Onsager's Conjecture. arXiv:1608.08301
153. James, R.D.: Displacive phase transformations in solids. J. Mech. Phys. Solids **34**, 359–394 (1986)
154. Jodeit Jr., M., Olver, P.J.: On the equation grad $f = M \operatorname{grad} g$. Proc. R. Soc. Edinb. Sect. A **116**, 341–358 (1990)
155. Kałamajska, A., Kružík, M.: Oscillations and concentrations in sequences of gradients. E-SAIM Control Optim. Calc. Var. **14**, 71–104 (2008)
156. Kinderlehrer, D.: Remarks about equilibrium configurations of crystals. Material Instabilities in Continuum Mechanics (Edinburgh. 1985–1986), Oxford Science Publications, pp. 217–241. Oxford University Press, Oxford (1988)
157. Kinderlehrer, D., Pedregal, P.: Characterizations of Young measures generated by gradients. Arch. Ration. Mech. Anal. **115**, 329–365 (1991)
158. Kinderlehrer, D., Pedregal, P.: Gradient Young measures generated by sequences in Sobolev spaces. J. Geom. Anal. **4**, 59–90 (1994)
159. Kirchheim, B.: Lipschitz minimizers of the 3-well problem having gradients of bounded variation. Preprint 12, Max-Planck-Institut für Mathematik in den Naturwissenschaften (1998)
160. Kirchheim, B.: Rigidity and Geometry of Microstructures. Lecture notes 16, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig (2003)
161. Kirchheim, B., Kristensen, J.: Automatic convexity of rank-1 convex functions. C. R. Math. Acad. Sci. Paris **349**, 407–409 (2011)
162. Kirchheim, B., Kristensen, J.: On rank-one convex functions that are homogeneous of degree one. Arch. Ration. Mech. Anal. **221**, 527–558 (2016)
163. Kristensen, J.: Lower semicontinuity of quasi-convex integrals in BV. Calc. Var. Partial Differ. Equ. **7**(3), 249–261 (1998)
164. Kristensen, J.: Lower semicontinuity in spaces of weakly differentiable functions. Math. Ann. **313**, 653–710 (1999)
165. Kristensen, J.: On the non-locality of quasiconvexity. Ann. Inst. H. Poincaré Anal. Non Linéaire **16**, 1–13 (1999)
166. Kristensen, J., Melcher, C.: Regularity in oscillatory nonlinear elliptic systems. Math. Z. **260**, 813–847 (2008)
167. Kristensen, J., Mingione, G.: The singular set of Lipschitzian minima of multiple integrals. Arch. Ration. Mech. Anal. **184**, 341–369 (2007)
168. Kristensen, J., Rindler, F.: Characterization of generalized gradient Young measures generated by sequences in $W^{1,1}$ and BV. Arch. Ration. Mech. Anal. **197**, 539–598 (2010); Erratum: Vol. 203, pp. 693–700 (2012)
169. Kristensen, J., Rindler, F.: Relaxation of signed integral functionals in BV. Calc. Var. Partial Differ. Equ. **37**, 29–62 (2010)
170. Kružík, M.: On the composition of quasiconvex functions and the transposition. J. Convex Anal. **6**, 207–213 (1999)
171. Kružík, M., Roubíček, T.: Explicit characterization of L^p-Young measures. J. Math. Anal. Appl. **198**, 830–843 (1996)
172. Kružík, M., Roubíček, T.: On the measures of DiPerna and Majda. Math. Bohem. **122**, 383–399 (1997)

173. Kuiper, N.H.: On C^1-isometric imbeddings. I, II. Nederl. Akad. Wetensch. Proc. Ser. A. 58 =. Indag. Math. **17**(545–556), 683–689 (1955)
174. Larsen, C.J.: Quasiconvexification in $W^{1,1}$ and optimal jump microstructure in BV relaxation. SIAM J. Math. Anal. **29**, 823–848 (1998)
175. Lavrentiev, M.: Sur quelques problèmes du calcul des variations. Ann. Mat. Pura. Appl. **4**, 7–28 (1926)
176. Leoni, G.: A First Course in Sobolev Spaces. Graduate Studies in Mathematics, vol. 105. American Mathematical Society (2009)
177. Malý, J., Ziemer, W.P.: Fine Regularity of Solutions of Elliptic Partial Differential Equations. Mathematical Surveys and Monographs, vol. 51. Applied Mathematical Sciences (1997)
178. Manià, B.: Sopra un essempio di Lavrentieff. Boll. Un. Mat. Ital. **13**, 147–153 (1934)
179. Marcellini, P.: Periodic solution and homogenization of nonlinear variational problems. Ann. Mat. Pura. Appl. **117**, 481–498 (1978)
180. Marcus, M., Mizel, V.J.: Transformations by functions in Sobolev spaces and lower semicontinuity for parametric variational problems. Bull. Am. Math. Soc. **79**, 790–795 (1973)
181. Massaccesi, A., Vittone, D.: An elementary proof of the rank-one theorem for BV functions. J. Eur. Math. Soc. (2016). To appear, arXiv:1601.02903
182. Matos, J.P.: Young measures and the absence of fine microstructures in a class of phase transitions. Eur. J. Appl. Math. **3**, 31–54 (1992)
183. Mattila, P.: Geometry of Sets and Measures in Euclidean Spaces. Cambridge Studies in Advanced Mathematics, vol. 44. Cambridge University Press, Cambridge (1995)
184. McShane, E.J.: Generalized curves. Duke Math. J. **6**, 513–536 (1940)
185. McShane, E.J.: Necessary conditions in generalized-curve problems of the calculus of variations. Duke Math. J. **7**, 1–27 (1940)
186. McShane, E.J.: Sufficient conditions for a weak relative minimum in the problem of Bolza. Trans. Am. Math. Soc. **52**, 344–379 (1942)
187. Milton, G.W.: The Theory of Composites. Cambridge Monographs on Applied and Computational Mathematics, vol. 6. Cambridge University Press, Cambridge (2002)
188. Mingione, G.: Regularity of minima: an invitation to the dark side of the calculus of variations. Appl. Math. **51**, 355–426 (2006)
189. Modica, L.: The gradient theory of phase transitions and the minimal interface criterion. Arch. Ration. Mech. Anal. **98**, 123–142 (1987)
190. Modica, L., Mortola, S.: The Γ-convergence of some functionals. Preprint 77-7, Instituto Matematico "Leonida Tonelli", University of Pisa (1977)
191. Mooney, C., Savin, O.: Some singular minimizers in low dimensions in the calculus of variations. Arch. Ration. Mech. Anal. **221**, 1–22 (2016)
192. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation. I – Basic Theory. Grundlehren der mathematischen Wissenschaften, vol. 330. Springer, Berlin (2006)
193. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation. II – Applications. Grundlehren der mathematischen Wissenschaften, vol. 331. Springer, Berlin (2006)
194. Morrey Jr., C.B.: On the solutions of quasi-linear elliptic partial differential equations. Trans. Am. Math. Soc. **43**, 126–166 (1938)
195. Morrey Jr., C.B.: Quasi-convexity and the lower semicontinuity of multiple integrals. Pac. J. Math. **2**, 25–53 (1952)
196. Morrey, Jr., C.B.: Multiple Integrals in the Calculus of Variations. Grundlehren der mathematischen Wissenschaften, vol. 130. Springer, Berlin (1966)
197. Moser, J.: A new proof of De Giorgi's theorem concerning the regularity problem for elliptic differential equations. Commun. Pure Appl. Math. **13**, 457–468 (1960)
198. Moser, J.: On Harnack's theorem for elliptic differential equations. Commun. Pure Appl. Math. **14**, 577–591 (1961)
199. Müller, S.: Homogenization of nonconvex integral functionals and cellular elastic materials. Arch. Ration. Mech. Anal. **99**, 189–212 (1987)
200. Müller, S.: On quasiconvex functions which are homogeneous of degree 1. Indiana Univ. Math. J. **41**, 295–301 (1992)

201. Müller, S.: Rank-one convexity implies quasiconvexity on diagonal matrices. Internat. Math. Res. Notices **20**, 1087–1095 (1999)
202. Müller, S.: A sharp version of Zhang's theorem on truncating sequences of gradients. Trans. Am. Math. Soc. **351**, 4585–4597 (1999)
203. Müller, S.: Variational models for microstructure and phase transitions. Calculus of Variations and Geometric Evolution Problems (Cetraro, 1996). Lecture Notes in Mathematics, vol. 1713, pp. 85–210. Springer, Berlin (1999)
204. Müller, S., Spector, S.J.: An existence theory for nonlinear elasticity that allows for cavitation. Arch. Ration. Mech. Anal. **131**, 1–66 (1995)
205. Müller, S., Šverák, V.: Attainment results for the two-well problem by convex integration. In: Geometric Analysis and the Calculus of Variations, pp. 239–251. International Press, Somerville (1996)
206. Müller, S., Šverák, V.: Convex integration with constraints and applications to phase transitions and partial differential equations. J. Eur. Math. Soc. **1**, 393–422 (1999)
207. Müller, S., Šverák, V.: Convex integration for Lipschitz mappings and counterexamples to regularity. Ann. Math. **157**, 715–742 (2003)
208. Müller, S., Sychev, M.A.: Optimal existence theorems for nonhomogeneous differential inclusions. J. Funct. Anal. **181**, 447–475 (2001)
209. Murat, F.: Compacité par compensation. Ann. Sc. Norm. Super. Pisa Cl. Sci. **5**, 489–507 (1978)
210. Murat, F.: Compacité par compensation. II. In: Proceedings of the International Meeting on Recent Methods in Nonlinear Analysis (Rome, 1978), pp. 245–256. Pitagora Editrice Bologna (1979)
211. Murat, F.: Compacité par compensation: condition nécessaire et suffisante de continuité faible sous une hypothèse de rang constant. Ann. Sc. Norm. Super. Pisa Cl. Sci. **8**, 69–102 (1981)
212. Nash, J.: C^1 isometric imbeddings. Ann. Math. **60**, 383–396 (1954)
213. Nash, J.: Continuity of solutions of parabolic and elliptic equations. Am. J. Math. **80**, 931–954 (1958)
214. Nečas, J.: On regularity of solutions to nonlinear variational inequalities for second-order elliptic systems. Rend. Mat. **8**, 481–498 (1975)
215. Nesi, V., Milton, G.W.: Polycrystalline configurations that maximize electrical resistivity. J. Mech. Phys. Solids **39**, 525–542 (1991)
216. Nguetseng, G.: A general convergence result for a functional related to the theory of homogenization. SIAM J. Math. Anal. **20**, 608–623 (1989)
217. Noether, E.: Invariante Variationsprobleme, pp. 235–257. Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse (1918)
218. Olver, P.J.: Applications of Lie Groups to Differential Equations. Graduate Texts in Mathematics, vol. 107, 2nd edn. Springer, Berlin (1993)
219. O'Neil, T.: A measure with a large set of tangent measures. Proc. Am. Math. Soc. **123**, 2217–2220 (1995)
220. Ornstein, D.: A non-inequality for differential operators in the L_1 norm. Arch. Ration. Mech. Anal. **11**, 40–49 (1962)
221. Pedregal, P.: Laminates and microstructure. Eur. J. Appl. Math. **4**, 121–149 (1993)
222. Pedregal, P.: Parametrized Measures and Variational Principles. Progress in Nonlinear Differential Equations and their Applications, vol. 30. Birkhäuser (1997)
223. Pompe, W.: The quasiconvex hull for the five-gradient problem. Calc. Var. Partial Differ. Equ. **37**, 461–473 (2010)
224. Preiss, D.: Geometry of measures in R^n: distribution, rectifiability, and densities. Ann. Math. **125**, 537–643 (1987)
225. Reshetnyak, Y.G.: Liouville's conformal mapping theorem under minimal regularity hypotheses. Sibirsk. Mat. Ž. **8**, 835–840 (1967)
226. Reshetnyak, Y.G.: The weak convergence of completely additive vector-valued set functions. Sibirsk. Mat. Ž. **9**, 1386–1394 (1968)

227. Rindler, F.: Lower semicontinuity for integral functionals in the space of functions of bounded deformation via rigidity and Young measures. Arch. Ration. Mech. Anal. **202**, 63–113 (2011)
228. Rindler, F.: Lower semicontinuity and Young measures in BV without Alberti's Rank-One theorem. Adv. Calc. Var. **5**, 127–159 (2012)
229. Rindler, F.: A local proof for the characterization of Young measures generated by sequences in BV. J. Funct. Anal. **266**, 6335–6371 (2014)
230. Rindler, F., Shaw, G.: Liftings, Young Measures, and Lower Semicontinuity. arXiv:1708.04165
231. Rindler, F., Shaw, G.: Strictly continuous extensions of functionals with linear growth to the space BV. Q. J. Math. **66**, 953–978 (2015)
232. Rockafellar, R.T.: Convex Analysis. Princeton Mathematical Series, vol. 28. Princeton University Press, Princeton (1970)
233. Rockafellar, R.T., Wets, R.J.B.: Variational Analysis. Grundlehren der mathematischen Wissenschaften, vol. 317. Springer, Berlin (1998)
234. Rogers, L.C.G., Williams, D.: Diffusions, Markov Processes, and Martingales. vol. 1: Foundations. Cambridge Mathematical Library. Cambridge University Press, Cambridge (2000). Reprint of the second edition (1994)
235. Roubíček, T.: Relaxation in Optimization Theory and Variational Calculus. De Gruyter Series in Nonlinear Analysis and Applications, vol. 4. De Gruyter (1997)
236. Rudin, W.: Real and Complex Analysis, 3rd edn. McGraw-Hill, New York (1986)
237. Schauder, J.: über lineare elliptische Differentialgleichungen zweiter Ordnung. Math. Z. **38**, 257–282 (1934)
238. Schauder, J.: Numerische Abschätzungen in elliptischen linearen Differentialgleichungen. Studia Math. **5**, 34–42 (1937)
239. Scheffer, V.: Regularity and Irregularity of Solutions to Nonlinear Second Order Elliptic Regularity and Irregularity of Solutions to Nonlinear Second Order Elliptic Systems and Inequalities. Ph.D. thesis, Princeton University (1974)
240. Scorza Dragoni, G.: Un teorema sulle funzioni continue rispetto ad una e misurabili rispetto ad un'altra variabile. Rend. Sem. Mat. Univ. Padova **17**, 102–106 (1948)
241. Seiner, H., Glatz, O., Landa, M.: A finite element analysis of the morphology of the twinned-to-detwinned interface observed in microstructure of the Cu-Al-Ni shape memory alloy. Int. J. Solids Struct. **48**, 2005–2014 (2011)
242. Serrin, J.: On the definition and properties of certain variational integrals. Trans. Am. Math. Soc. **101**, 139–167 (1961)
243. Shnirelman, A.: On the nonuniqueness of weak solution of the Euler equation. Commun. Pure Appl. Math. **50**, 1261–1286 (1997)
244. Simon, L.: Schauder estimates by scaling. Calc. Var. Partial Differ. Equ. **5**, 391–407 (1997)
245. Spadaro, E.N.: Non-uniqueness of minimizers for strictly polyconvex functionals. Arch. Ration. Mech. Anal. **193**, 659–678 (2009)
246. Stein, E.M.: Singular Integrals and Differentiability Properties of Functions. Princeton University Press, Princeton (1970)
247. Stein, E.M.: Harmonic Analysis. Princeton University Press, Princeton (1993)
248. Šverák, V.: Regularity properties of deformations with finite energy. Arch. Ration. Mech. Anal. **100**, 105–127 (1988)
249. Šverák, V.: On regularity for the Monge-Ampère equation without convexity assumptions, Technical report. Heriot-Watt University (1991)
250. Šverák, V.: Quasiconvex functions with subquadratic growth. Proc. R. Soc. Lond. Ser. A **433**, 723–725 (1991)
251. Šverák, V.: New examples of quasiconvex functions. Arch. Ration. Mech. Anal. **119**, 293–300 (1992)
252. Šverák, V.: Rank-one convexity does not imply quasiconvexity. Proc. R. Soc. Edinb. Sect. A **120**, 185–189 (1992)
253. Šverák, V.: On Tartar's conjecture. Ann. Inst. H. Poincaré Anal. Non Linéaire **10**, 405–412 (1993)

254. Šverák, V.: On the problem of two wells. Microstructure and Phase Transition. The IMA Volumes in Mathematics and its Applications, vol. 54, pp. 183–189. Springer, Berlin (1993)
255. Šverák, V., Yan, X.: A singular minimizer of a smooth strongly convex functional in three dimensions. Calc. Var. Partial Differ. Equ. **10**, 213–221 (2000)
256. Šverák, V., Yan, X.: Non-Lipschitz minimizers of smooth uniformly convex functionals. Proc. Natl. Acad. Sci. USA **99**, 15269–15276 (2002)
257. Sychev, M.A.: Comparing Various Methods of Resolving Nonconvex Variational Problems. Preprint 66, SISSA (1998)
258. Sychev, M.A.: A new approach to Young measure theory, relaxation and convergence in energy. Ann. Inst. H. Poincaré Anal. Non Linéaire **16**, 773–812 (1999)
259. Sychev, M.A.: Characterization of homogeneous gradient Young measures in case of arbitrary integrands. Ann. Sc. Norm. Super. Pisa Cl. Sci. **29**, 531–548 (2000)
260. Sychev, M.A.: Comparing two methods of resolving homogeneous differential inclusions. Calc. Var. Partial Differ. Equ. **13**, 213–229 (2001)
261. Sychev, M.A.: A few remarks on differential inclusions. Proc. R. Soc. Edinb. Sect. A **136**, 649–668 (2006)
262. Sychev, M.A.: Comparing various methods of resolving differential inclusions. J. Convex Anal. **18**, 1025–1045 (2011)
263. Székelyhidi Jr., L.: From isometric embeddings to turbulence. In: HCDTE Lecture Notes. Part II. Nonlinear Hyperbolic PDEs, Dispersive and Transport Equations. AIMS Series on Applied Mathematics, vol. 7, pp. 195–255. American Institute of Mathematical Sciences (AIMS) (2013)
264. Székelyhidi Jr., L.: The regularity of critical points of polyconvex functionals. Arch. Ration. Mech. Anal. **172**, 133–152 (2004)
265. Székelyhidi Jr., L., Wiedemann, E.: Young measures generated by ideal incompressible fluid flows. Arch. Ration. Mech. Anal. **206**, 333–366 (2012)
266. Tang, Q.: Almost-everywhere injectivity in nonlinear elasticity. Proc. R. Soc. Edinb. Sect. A **109**, 79–95 (1988)
267. Tartar, L.: Compensated compactness and applications to partial differential equations. In: Nonlinear Analysis and Mechanics: Heriot-Watt Symposium, Vol. IV. Research Notes in Mathematics, vol. 39, pp. 136–212. Pitman (1979)
268. Tartar, L.: The compensated compactness method applied to systems of conservation laws. In: Systems of Nonlinear Partial Differential Equations (Oxford, 1982). NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., vol. 111, pp. 263–285. Reidel (1983)
269. Tartar, L.: H-measures, a new approach for studying homogenisation, oscillations and concentration effects in partial differential equations. Proc. R. Soc. Edinb. Sect. A **115**, 193–230 (1990)
270. Tartar, L.: On mathematical tools for studying partial differential equations of continuum physics: H-measures and Young measures. In: Developments in Partial Differential Equations and Applications to Mathematical Physics (Ferrara, 1991), pp. 201–217. Plenum (1992)
271. Tartar, L.: Some remarks on separately convex functions. Microstructure and Phase Transition. The IMA Volumes in Mathematics and its Applications, vol. 54, pp. 191–204. Springer, Berlin (1993)
272. Tartar, L.: Beyond Young measures. Meccanica **30**, 505–526 (1995)
273. Temam, R.: Mathematical Problems in Plasticity. Gauthier-Villars (1985)
274. Temam, R., Strang, G.: Functions of bounded deformation. Arch. Ration. Mech. Anal. **75**, 7–21 (1980)
275. Tonelli, L.: Sur un méthode directe du calcul des variations. Rend. Circ. Mat. Palermo **39**, 233–264 (1915)
276. Tonelli, L.: La semicontinuità nel calcolo delle variazioni. Rend. Circ. Mat. Palermo **44**, 167–249 (1920)
277. Tonelli, L.: Opere scelte. Vol II: Calcolo delle variazioni. Edizioni Cremonese, Rome (1961)
278. Van der Vorst, R.C.A.M.: Variational identities and applications to differential systems. Arch. Ration. Mech. Anal. **116**, 375–398 (1992)

279. Wigner, E.P.: The unreasonable effectiveness of mathematics in the natural sciences. Commun. Pure Appl. Math. **13**, 1–14 (1960). Richard Courant Lecture in Mathematical Sciences Delivered at New York University, May 11, 1959
280. Young, L.C.: Generalized curves and the existence of an attained absolute minimum in the calculus of variations. C. R. Soc. Sci. Lett. Varsovie, Cl. III **30**, 212–234 (1937)
281. Young, L.C.: Generalized surfaces in the calculus of variations. Ann. Math. **43**, 84–103 (1942)
282. Young, L.C.: Generalized surfaces in the calculus of variations. II. Ann. Math. **43**, 530–544 (1942)
283. Young, L.C.: Lectures on the Calculus of Variations and Optimal Control Theory, 2nd edn. Chelsea (1980)
284. Zhang, K.: A construction of quasiconvex functions with linear growth at infinity. Ann. Sc. Norm. Super. Pisa Cl. Sci. **19**, 313–326 (1992)
285. Ziemer, W.P.: Weakly Differentiable Functions. Graduate Texts in Mathematics, vol. 120. Springer, Berlin (1989)

Index

F. Rindler, *Calculus of Variations*, Universitext,
https://doi.org/10.1007/978-3-319-77637-8