Lower Semicontinuity and Young Measures for Integral Functionals with Linear Growth

Johan Filip Rindler
Worcester College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy in Mathematics

Trinity Term 2011

In memory of

Rudolf Lange
(⋆15 August 1931 †18 January 1994)

and

Jörg Kraftmeier
(⋆4 August 1942 †12 June 2005).

Abstract

The first contribution of this thesis is a new proof of sequential weak* lower semicontinuity in $\mathrm{BV}(\Omega;\mathbb{R}^m)$ for integral functionals of the form

$$\begin{aligned}\mathcal{F}(u) := &\int_\Omega f(x,\nabla u)\,\mathrm{d}x + \int_\Omega f^\infty\left(x,\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}\right)\mathrm{d}|D^s u| \\ &+ \int_{\partial\Omega} f^\infty\big(x, u|_{\partial\Omega}\otimes n_\Omega\big)\,\mathrm{d}\mathcal{H}^{d-1}, \qquad\qquad u\in \mathrm{BV}(\Omega;\mathbb{R}^m),\end{aligned}$$

where $f\colon \overline{\Omega}\times\mathbb{R}^{m\times d}\to\mathbb{R}$ is a quasiconvex Carathéodory integrand with linear growth at infinity, i.e. $|f(x,A)|\le M(1+|A|)$ for some $M>0$, and f is such that the recession function $f^\infty(x,A) := \lim_{x'\to x, t\to\infty} t^{-1}f(x',tA)$ exists and is (jointly) continuous. In contrast to the classical proofs by Ambrosio & Dal Maso [*J. Funct. Anal.* 109 (1992), 76–97] and Fonseca & Müller [*Arch. Ration. Mech. Anal.* 123 (1993), 1–49], the one presented here does not use Alberti's Rank One Theorem [*Proc. Roy. Soc. Edinburgh Sect. A* 123 (1993), 239–274], but combines the usual blow-up method with a rigidity result for gradients.

The second main result is the first general weak* lower semicontinuity theorem in the space $\mathrm{BD}(\Omega)$ of functions of bounded deformation for functionals of the form

$$\begin{aligned}\mathcal{F}(u) := &\int_\Omega f(x,\mathcal{E}u)\,\mathrm{d}x + \int_\Omega f^\infty\left(x,\frac{\mathrm{d}E^s u}{\mathrm{d}|E^s u|}\right)\mathrm{d}|E^s u| \\ &+ \int_{\partial\Omega} f^\infty\big(x, u|_{\partial\Omega}\odot n_\Omega\big)\,\mathrm{d}\mathcal{H}^{d-1}, \qquad\qquad u\in \mathrm{BD}(\Omega),\end{aligned}$$

where f satisfies similar assumptions as before, in particular it is assumed symmetric-quasiconvex. The main novelty is that we allow for non-vanishing Cantor-parts in the symmetrized derivative Eu. The proof hinges on a construction of good blow-ups, which in turn is based on local rigidity arguments for some differential inclusions involving symmetrized gradients and an iteration of the blow-up construction. This strategy allows us to establish the lower semicontinuity result without an Alberti-type theorem in $\mathrm{BD}(\Omega)$, which is not available at present.

This dissertation also presents an improved functional-analytic framework for generalized Young measures (with linear growth) as introduced by DiPerna & Majda [*Comm. Math. Phys.* 108 (1987), 667–689]. Localization principles and Jensen-type inequalities for such Young measures form the technical backbone of the lower semicontinuity proofs.

MSC (2010): 49J45 (primary); 26B30, 35J50, 28B05, 49Q20, 74B05, 74C10.
Keywords: Bounded variation, BV, bounded deformation, BD, lower semicontinuity, Young measure, rigidity, differential inclusion, Alberti's Rank One Theorem.

Acknowledgments

First and foremost, it is a great pleasure to express deep gratitude and admiration to my supervisor Professor Jan Kristensen. His enthusiasm and commitment to research has been an immense source of inspiration and I will always fondly remember my time as his student.

I also wish to extend many thanks to my undergraduate thesis supervisor and mentor Professor Alexander Mielke, who set me on the path to work in the Calculus of Variations and who has always greatly supported my research interests. In this regard, I also owe much to my friend Dr. Günter von Häfen, who from my teenage years onward has taught me so much about science. Further, I thank Professor John M. Ball for the OxPDE scholarship that supported me during the work on this thesis (through the EPSRC Science and Innovation award EP/E035027/1 to OxPDE) and Professor Irene Fonseca for hosting me for three months at Carnegie Mellon University in Pittsburgh, USA.

I am very grateful to the many people with whom I have had the pleasure to discuss some of my mathematical ideas, including (in alphabetical order) Professor Giovanni Alberti, Professor John M. Ball, Dr. Yves Capdeboscq, Professor Gui-Qiang Chen, Professor Etienne Emmrich, Professor Irene Fonseca, Dr. Günter von Häfen, Dr. Duvan Henao, Professor Bernd Kirchheim, Dr. Carolin Kreisbeck, Dr. habil. Martin Kružík, Professor Giovanni Leoni, Professor Alexander Mielke, Professor Guiseppe Rosario Mingione, Professor Stefan Müller, Dr. Richard Nickl, Professor Harald Rindler, Professor Tomáš Roubíček, Professor Gregory Seregin, Professor Endre Süli, Professor László Székelyhidi, Jr., Professor Petra Wittbold, and also my fellow students and friends Michael Helmers, Moritz Schmitt, Parth Soneji and Emil Wiedemann.

It was a great pleasure to be invited to the workshop "Rate-independent systems: Modeling, Analysis, and Computations" at the Banff International Research Station (BIRS), Canada; for this opportunity I am indebted to Professor Guiseppe Savaré and Professor Ulisse Stefanelli. Moreover, I greatly enjoyed a one-week stay at the Hausdorff Center for Mathematics in Bonn, Germany, for which I thank Professor László Székelyhidi, Jr. and Emil Wiedemann. Furthermore, special thanks go to Professor Robert V. Kohn for sending me a hard copy of his PhD thesis and to all the OxPDE, Worcester College and Oxford University Admin Staff for much support.

Finally, I cannot express with words the gratitude I feel toward my mother Karin, my girlfriend Laura, my grandmother Helene, my uncle Harald, Kaye and Prakash, as well as toward my friends. Their understanding, support and encouragement was the basis from which I drew the strength to carry out this work.

Contents

Chapter 1

Introduction

In the wake of Morrey's groundbreaking work on quasiconvexity [93, 94] in the 1950s and 1960s, the theory of the vector-valued Calculus of Variations has sustained a steady growth for the last 60 years, branching out into many different sub-fields and enabling a variety of new applications. Of course, the "modern" theory of the Calculus of Variations is much older and probably started with the formulation of Hilbert's 19^{th}, 20^{th} and 23^{rd} problems in 1900 [72]. Many exciting developments have taken place recently and the Calculus of Variations retains its status as a prospering field with many connections to other areas of Partial Differential Equations, Analysis, and far beyond.

This work aims to advance the variational theory of integral functionals with integrands depending on the gradient or symmetrized gradient (the symmetric part of the gradient) of a vector-valued function and with linear growth at infinity. After a general discussion of the basic underlying principles, the present chapter will present the two pivotal theorems of this work in their respective contexts and briefly comment on the ideas of proof. We postpone precise definitions and references to subsequent chapters.

1.1 Linear-growth functionals depending on the gradient

The first main result of this thesis concerns variational principles of the following form:

$$\text{Minimize} \quad \mathcal{F}(u) := \int_\Omega f\bigl(x, \nabla u(x)\bigr) \,\mathrm{d}x \quad \text{over a class } \mathcal{U} \text{ of functions } u\colon \Omega \to \mathbb{R}^m.$$

Here, $\Omega \subset \mathbb{R}^d$ ($d \in \mathbb{N}$, usually $d > 1$) is a bounded open Lipschitz domain and $f\colon \Omega \times \mathbb{R}^{m\times d} \to \mathbb{R}$ is a Carathéodory integrand (i.e. measurable in its first and continuous in its second argument). This variational principle, however, is incomplete as long as we do not specify the class $\mathcal{U}$ of candidate functions u. To find a sensible class of such candidate functions, one needs to examine the growth and coercivity properties of the integrand f; we will do so below.

The overarching philosophy here is that of the so-called *Direct Method* of the Calculus of Variations, meaning that we construct a minimizer by taking a *minimizing sequence*

$(u_j) \subset \mathcal{U}$, that is

$$\mathcal{F}(u_j) \to \inf_{\mathcal{U}} \mathcal{F} \qquad \text{as } j \to \infty.$$

Then, by a *compactness* argument, we aim to pass to a limit $u_j \rightsquigarrow u_\infty$ (of a subsequence if necessary) with respect to a suitable convergence "$\rightsquigarrow$". In this work, we work directly with a notion of convergence for sequences, general topology tools only play a minor role. Under certain *continuity* assumptions on $\mathcal{F}$, this limit u_∞ indeed constitutes a (global) minimizer to our problem. It should be emphasized that the choice of convergence is not a-priori dictated by the problem, but is made purely on mathematical grounds and needs to balance the counteracting requirements of compactness and continuity. Moreover, the compactness available in practice might force us to consider a notion of convergence so weak that in general $u_\infty \notin \mathcal{U}$. In this case we need to *relax* our original problem to the closure of $\mathcal{U}$ with respect to the chosen convergence and look for minimizers there.

To carry out this program rigorously, the first step is to specify the growth and coercivity assumptions on the integrand f. Throughout this work we impose a hypothesis of *linear growth* on f, i.e. we suppose that there exists a constant $M > 0$ such that

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

where $|A|$ is the Frobenius norm of A (the Euclidean norm of A when considered as a vector in $\mathbb{R}^{md}$). Hence, the above integral functional is finite if u belongs to the Sobolev space $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$. If we furthermore consider a *coercivity* assumption of the form

$$C|A| - C^{-1} \leq f(x, A) \qquad \text{for all } x \in \Omega,\ A \in \mathbb{R}^{m \times d},$$

where $C > 0$ is another constant, it is not hard to see that for a minimizing sequence $(u_j) \subset \mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$ we have

$$\limsup_{j \to \infty} \|\nabla u_j\|_{\mathrm{L}^1(\Omega; \mathbb{R}^{m \times d})} \leq \frac{1}{C} \left(\inf_{\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)} \mathcal{F} \right) + \frac{|\Omega|}{C^2} < \infty.$$

So, assuming some sort of boundary conditions on the admissible u, we get by the Poincaré–Friedrichs inequality that

$$\limsup_{j \to \infty} \|u_j\|_{\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)} < \infty.$$

This suggests using a type of weak convergence to get a convergent subsequence. However, unlike in the Sobolev classes $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^m)$ with $p > 1$, in the linear-growth space $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$ a uniform norm-bound is *not* enough to deduce weak compactness due to the non-reflexivity of $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$. Thus, the usual compactness result, the Banach–Alaoglu–Bourbaki Theorem, is not applicable here. In fact, it is easy to give an example of a sequence bounded in the $\mathrm{W}^{1,1}$-norm that contains no weakly convergent subsequence: 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 in every reasonable sense to $u_\infty = \mathbb{1}_{(0,1)}$, which is clearly not in $\mathrm{W}^{1,1}(-1,1)$. The decisive feature of this example is that the gradients of the u_j *concentrate*, meaning that $u_j' \, \mathcal{L}^d \llcorner (-1,1)$ converges 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.

From this discussion we may conclude that the space $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ is indeed too small to serve as the set of candidate functions for our variational principle. Clearly, we need a space containing $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ such that a uniform norm-bound implies relative compactness for a suitable (weak) convergence. Moreover, in order to retain the connection to our original problem, it is imperative 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 continuous.

All these requirements are fulfilled by the space $\mathrm{BV}(\Omega;\mathbb{R}^m)$ of functions of bounded variation, see [10,50,130] for details about this important class of functions, applications and a historical perspective. In brief, $\mathrm{BV}(\Omega;\mathbb{R}^m)$ is defined to contain all $\mathrm{L}^1(\Omega;\mathbb{R}^m)$-functions whose distributional derivative can be represented as a $\mathbb{R}^{m\times d}$-valued finite Radon measure Du on Ω. In this space we use the weak* convergence $u_j \overset{*}{\rightharpoonup} u$ meaning that $u_j \to u$ (strongly) in $\mathrm{L}^1(\Omega;\mathbb{R}^m)$ and $Du_j \overset{*}{\rightharpoonup} Du$ in the sense of measures. Most important in the theory of these functions is the Lebesgue–Radon–Nikodým decomposition of the measure Du,

$$Du = \nabla u \, \mathcal{L}^d \llcorner \Omega + D^s u = \nabla u \, \mathcal{L}^d \llcorner \Omega + \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|} \, |D^s u|,$$

where ∇u denotes the density of Du with respect to the d-dimensional Lebesgue measure $\mathcal{L}^d$ (strictly speaking, this is an abuse of notation since this density ∇u is not necessarily curl-free), $D^s u$ is the singular part of Du (with respect to Lebesgue measure) and $\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}$ is the Radon–Nikodým derivative of $D^s u$ with respect to its total variation measure $|D^s u|$. Further, $D^s u$ can be split as

$$D^s u = D^j u + D^c u,$$

with the *jump part* $D^j u$ and the *Cantor-part* $D^c u$, respectively. Details about this splitting can be found in [10]. We only record that the measure $D^j u$ has the form $a \otimes n \, \mathcal{H}^{d-1} \llcorner J_u$, where J_u is the $\mathcal{H}^{d-1}$-rectifiable *jump set*, $n\colon J_u \to \mathbb{S}^{d-1}$ is its normal (for a chosen orientation), and $a\colon J_u \to \mathbb{R}^m$ denotes the height of the jump for each component of u in direction n. The measure $D^c u$ on the other hand contains all remaining parts of $D^s u$. While there are several restrictions on the structure of $D^c u$, this part is in general non-empty and might contain fractal measures. For example, the Cantor function is in BV and has as its derivative the (fractal) Cantor measure.

So far we have identified $\mathrm{BV}(\Omega;\mathbb{R}^m)$ as a possible extended space of candidate functions for the above variational principle. To continue the investigation, we now need a way of extending the functional $\mathcal{F}$ to a functional $\bar{\mathcal{F}}$ defined on $\mathrm{BV}(\Omega;\mathbb{R}^m)$ and 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 $\mathrm{BV}(\Omega;\mathbb{R}^m)$. This extension is given by

$$\bar{\mathcal{F}}(u) := \int_\Omega f\big(x, \nabla u(x)\big)\,\mathrm{d}x + \int_\Omega f^\infty\biggl(x, \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\biggr)\,\mathrm{d}|D^s u|(x),$$

where f^∞ is the (strong) *recession function*

$$f^\infty(x, A) := \lim_{\substack{x' \to x\\ A' \to A\\ t \to \infty}} \frac{f(x', tA')}{t}, \qquad x \in \overline{\Omega},\ A \in \mathbb{R}^{m\times d},$$

which we assume to exist as a *continuous* function. It turns out that the so-defined functional $\bar{\mathcal{F}}$ is continuous with respect to the $\langle\,\boldsymbol{\cdot}\,\rangle$-strict convergence, in which $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ is dense in $\mathrm{BV}(\Omega;\mathbb{R}^m)$. A precise definition of this notion of convergence can be found in the next chapter. We only remark that it is weaker than the convergence in norm, but stronger than the weak* convergence. In particular, from the density and continuity of $\bar{\mathcal{F}}$ we conclude that the infimum of $\mathcal{F}$ over $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ and the infimum of $\bar{\mathcal{F}}$ over $\mathrm{BV}(\Omega;\mathbb{R}^m)$ are the same. It is common to call $\bar{\mathcal{F}}$ the *relaxation* of $\mathcal{F}$.

Using the uniform $\mathrm{W}^{1,1}$-norm bound, we may invoke the compactness theorem for the weak* convergence in $\mathrm{BV}(\Omega;\mathbb{R}^m)$ to select a (non-relabeled) subsequence of our minimizing sequence $(u_j) \subset \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ with $u_j \overset{*}{\rightharpoonup} u_\infty \in \mathrm{BV}(\Omega;\mathbb{R}^m)$. Notice that our original minimizing sequence (u_j) from the space $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ is still minimizing for the relaxed problem (we could even choose $(u_j) \subset \mathrm{C}^\infty(\Omega;\mathbb{R}^m)$ by a mollification argument).

The final step in our program of showing the existence of a minimizer via the Direct Method, is to establish the *lower semicontinuity* property

$$\bar{\mathcal{F}}(u_\infty) \le \liminf_{j\to\infty} \mathcal{F}(u_j),$$

by which we may conclude that indeed u_∞ is a minimizer (of the relaxed problem).

This work is exclusively concerned with the question of lower semicontinuity. In the easiest case, lower semicontinuity can be established for convex integrands (see for instance Theorem 5.27 in [55]). However, convexity is only sufficient, but far from necessary for vector-valued problems ($m > 1$, $d > 1$) and generally not available in applications. Therefore, we need to look for alternative and more general notions of convexity. The "correct" one here turns out to be Morrey's quasiconvexity [93, 94]: A locally bounded Borel function $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ is called *quasiconvex* if

$$h(A) \le \fint_\omega h\big(A + \nabla\psi(z)\big)\,\mathrm{d}z \qquad \text{for all } A \in \mathbb{R}^{m\times d} \text{ and all } \psi \in \mathrm{W}^{1,\infty}_0(\omega;\mathbb{R}^m),$$

where $\omega \subset \mathbb{R}^d$ is an arbitrary bounded open Lipschitz domain (it can be shown that this definition does not depend on the choice of ω). Quasiconvexity is both necessary and sufficient for weak* lower semicontinuity in $\mathrm{BV}(\Omega;\mathbb{R}^m)$. For the necessity see Lemma 3.18 in [34]. Regarding the sufficiency, the main general lower semicontinuity theorem in our situation is the following (this is Theorem 4.1):

Theorem (Ambrosio & Dal Maso '92 and Fonseca & Müller '93). *Let $\Omega \subset \mathbb{R}^d$ be a bounded open Lipschitz domain and let $f\colon \overline{\Omega} \times \mathbb{R}^{m\times d} \to \mathbb{R}$ satisfy the following assumptions:*

(i) f is a Carathéodory function,

(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 and is (jointly) continuous on $\overline{\Omega} \times \mathbb{R}^{m\times d}$,

(iv) $f(x,\bullet)$ is quasiconvex for all $x \in \overline{\Omega}$.

Then, the functional

$$\mathcal{F}(u) := \int_\Omega f\bigl(x, \nabla u(x)\bigr) \,\mathrm{d}x + \int_\Omega f^\infty\biggl(x, \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\biggr) \,\mathrm{d}|D^s u|(x)$$
$$+ \int_{\partial\Omega} f^\infty\bigl(x, u|_{\partial\Omega}(x) \otimes n_\Omega(x)\bigr) \,\mathrm{d}\mathcal{H}^{d-1}(x), \qquad u \in \mathrm{BV}(\Omega;\mathbb{R}^m),$$

is lower semicontinuous with respect to weak convergence in the space* $\mathrm{BV}(\Omega;\mathbb{R}^m)$.

In the above definition of $\mathcal{F}$, the function $u|_{\partial\Omega} \in \mathrm{L}^1(\partial\Omega, \mathcal{H}^{d-1};\mathbb{R}^m)$ is the (inner) boundary trace of u onto $\partial\Omega$, whereas $n_\Omega\colon \partial\Omega \to \mathbb{S}^{d-1}$ is the boundary unit inner normal. If the boundary values of any admissible weakly* converging sequence are the same as the boundary values of the limit, then the boundary term may be omitted. The same is true if $f \ge 0$ since then we can only lose mass in the limit.

This lower semicontinuity theorem was first established by Ambrosio & Dal Maso [8] and Fonseca & Müller [57], also see [56], which introduced the blow-up method employed in the proof, [27] for a systematic approach to related relaxation problems, and [83] for the recent extension to signed integrands. Notice, however, that in contrast to the original works [8, 57] we stated this theorem with a stronger requirement on the existence of a recession function (not just the limes superior, but a proper limit). Our method of proof requires this strengthening, but at the same time generalizes the result slightly, allowing to treat signed Carathéodory integrands as well. More information on the comparison of different types of recession function can be found in Remark 4.14.

Traditionally, the proof of the above result crucially employs Alberti's Rank One Theorem (see Theorem 2.25), originally proved in [1]. This fundamental result asserts that for a function $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$,

$$\mathrm{rank}\biggl(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x_0)\biggr) \le 1 \qquad \text{for } |D^s u|\text{-almost every } x_0 \in \Omega.$$

Chapter 4 will give a proof of the lower semicontinuity theorem in BV that does not use Alberti's Rank One Theorem, but instead combines the usual blow-up arguments with a rigidity lemma. "Rigidity" here means that all (exact) solutions to certain *differential inclusions* involving the gradient have additional structure. The decisive point is to realize that in the currently known proof of lower semicontinuity, Alberti's Theorem is employed only as a rigidity result: Almost everywhere, blowing-up a function $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ around a singular point $x_0 \in \Omega$ yields a (local) BV-function $v \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^m)$ with constant polar function of the derivative, i.e.

$$Dv = P(x_0)|Dv|, \qquad \text{where} \qquad P(x_0) = \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x_0).$$

Alberti's Theorem now tells us that for $|D^s u|$-almost every $x_0 \in \Omega$, $P(x_0) = a \otimes \xi$ for some $a \in \mathbb{R}^m$, $\xi \in \mathbb{S}^{d-1}$, and hence we may infer that v can be written as

$$v(x) = v_0 + \psi(x \cdot \xi)a \qquad \text{for some } \psi \in \mathrm{BV}(\mathbb{R}),\ v_0 \in \mathbb{R}^m. \tag{1.1}$$

The key observation here is that a weaker statement can be proved much more easily: Suppose a function $v \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^m)$ satisfies

$$Dv = P|Dv| \qquad \text{for a fixed matrix } P \in \mathbb{R}^{m\times d} \text{ with } |P| = 1.$$

Then, if $\operatorname{rank} P \leq 1$, again (1.1) holds (with $P = a \otimes \xi$), whereas if $\operatorname{rank} P \geq 2$, then v must be affine. This simple rigidity result traces its origins to Hadamard's jump condition and Proposition 2 in [18]. For our purposes, however, we need a stronger statement than in the latter reference, but the proof is still elementary and based only on the fact that BV-derivatives must be curl-free, which translates into an algebraic condition on P, and the fact that gradients are always orthogonal to level sets. To the best of the author's knowledge, rigidity results seem not to have been employed explicitly in lower semicontinuity theory before (except the aforementioned use of Alberti's Theorem of course).

For a sequence $u_j \overset{*}{\rightharpoonup} u$ in the BV-lower semicontinuity theorem we distinguish several different types of blow-ups, depending on whether x_0 is a "regular" point for Du, or a "singular" point, and in the latter case also depending on whether $\operatorname{rank} P(x_0) \leq 1$ or $\operatorname{rank} P(x_0) \geq 2$. At regular points ($\mathcal{L}^d$-a.e.), we have a *regular blow-up*, that is an affine blow-up limit, and we can apply quasiconvexity directly. At singular points $x_0 \in \Omega$ with $\operatorname{rank} P(x_0) \leq 1$ we have a *fully singular blow-up*, meaning that by the above rigidity result we get a one-directional function in the blow-up limit and we need an averaging procedure before we can apply quasiconvexity (just as in the usual proof). On the other hand, if at a singular point x_0 it holds that $\operatorname{rank} P(x_0) \geq 2$, we call this a *semi-regular blow-up*, we get an affine function in the blow-up again, which can then be treated by just slightly adapting the procedure for regular blow-ups. Of course, from Alberti's Theorem we know that this case occurs only on a $|D^s u|$-negligible set, but the main objective of the new proof is to avoid using this deep result.

Combining these blow-ups with the machinery of Young measures explained below, we are then able to conclude the proof of the lower semicontinuity theorem in BV.

1.2 Functionals depending on the symmetrized gradient

We now turn our attention to a different minimization problem:

$$\text{Minimize} \quad \mathcal{F}(u) := \int_\Omega f\bigl(x, \mathcal{E}u(x)\bigr) \,\mathrm{d}x \quad \text{over a class of functions } u\colon \Omega \to \mathbb{R}^d,$$

where

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

is the *symmetrized gradient* of u and again $f\colon \Omega \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}$ is a Carathédory integrand. After imposing linear growth and coercivity assumptions on f as before, we are led to first consider candidate functions from the space

$$\mathrm{LD}(\Omega) := \bigl\{\, u \in \mathrm{L}^1(\Omega;\mathbb{R}^d) \;\;\mathbf{:}\;\; \mathcal{E}u \in \mathrm{L}^1(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}}) \,\bigr\},$$

where in this definition $\mathcal{E}u$ is understood in the generalized (distributional) sense. The space $\mathrm{LD}(\Omega)$ takes a role parallel to that of $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ for linear-growth functionals depending on the gradient as in the last section and one can show that these two spaces are *not* the same. This is contrary to the situation for the version of $\mathrm{LD}(\Omega)$ with L^1 replaced by L^p, $p > 1$, which turns out to be nothing else than $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^d)$. Indeed, for $p > 1$, Korn's inequality

$$\|\nabla u\|_{\mathrm{L}^p(\Omega;\mathbb{R}^{d\times d})} \le C\bigl(\|u\|_{\mathrm{L}^p(\Omega;\mathbb{R}^d)} + \|\mathcal{E}u\|_{\mathrm{L}^p(\Omega;\mathbb{R}^{d\times d})}\bigr)$$

holds (see Theorem 6.3-3 in [31] and also [73]), but the analogous statement for $p = 1$ is *false*, a result that is often called "Ornstein's Non-Inequality", see [32, 103]. It follows that $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^d)$ is a proper subset of $\mathrm{LD}(\Omega)$. Nevertheless, the space $\mathrm{LD}(\Omega)$ suffers from the same deficiency as $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ in that it is non-reflexive and hence the weak convergence is not suitable for our problem.

Similarly to the gradient case, however, we can define the space $\mathrm{BD}(\Omega)$ of functions of bounded deformation as the space of all functions $u \in \mathrm{L}^1(\Omega;\mathbb{R}^d)$ such that the distributional symmetrized derivative (defined by duality with the symmetrized gradient) is a $\mathbb{R}^{d\times d}_{\mathrm{sym}}$-valued finite Radon measure Eu on Ω. This space was introduced in the works [88, 111, 112] in order to treat variational problems from the mathematical theory of plasticity and has been investigated by various authors, see for example [7, 61, 71, 79, 80, 122, 123]. As before, we have the analogous notion of weak* convergence and the Lebesgue–Radon–Nikodým decomposition

$$Eu = \mathcal{E}u\, \mathcal{L}^d \llcorner \Omega + E^s u = \mathcal{E}u\, \mathcal{L}^d \llcorner \Omega + \frac{\mathrm{d}E^s u}{\mathrm{d}|E^s u|}\, |E^s u|.$$

Moreover, we can again (continuously with respect to the $\langle\,\boldsymbol{\cdot}\,\rangle$-strict convergence) extend the functional $\mathcal{F}$ to functions $u \in \mathrm{BD}(\Omega)$ via

$$\bar{\mathcal{F}}(u) := \int_\Omega f\bigl(x, \mathcal{E}u(x)\bigr) \,\mathrm{d}x + \int_\Omega f^\infty\biggl(x, \frac{\mathrm{d}E^s u}{\mathrm{d}|E^s u|}(x)\biggr) \,\mathrm{d}|E^s u|(x).$$

Several lower semicontinuity theorems in the space $\mathrm{BD}(\Omega)$ are available, see for example [22, 23, 47, 63], but they are all restricted to *special* functions of bounded deformation, i.e. those $u \in \mathrm{BD}(\Omega)$ for which the singular part $E^s u$ originates from jumps only and does not contain Cantor-type measures.

The main new contribution of this thesis is the following general lower semicontinuity theorem (this is Theorem 5.1), which is completely analogous to the corresponding result in BV from the last section.

Theorem. *Let $\Omega \subset \mathbb{R}^d$ be a bounded open Lipschitz domain and let $f\colon \overline{\Omega} \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}$ satisfy the following assumptions:*

(i) f is a Carathéodory function,

(ii) $|f(x,A)| \le M(1+|A|)$ for some $M > 0$ and all $x \in \overline{\Omega}$, $A \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$,

(iii) the (strong) recession function f^∞ exists and is (jointly) continuous on $\overline{\Omega} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$,

(iv) $f(x,\,\cdot\,)$ is symmetric-quasiconvex for all $x \in \overline{\Omega}$, that is,

$$f(x,A) \le \fint_\omega f\bigl(x, A + \mathcal{E}\psi(z)\bigr)\,\mathrm{d}z \qquad \text{for all } A \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \text{ and all } \psi \in \mathrm{W}^{1,\infty}_0(\omega;\mathbb{R}^d),$$

where $\omega \subset \mathbb{R}^d$ is an arbitrary bounded open Lipschitz domain.

Then, the functional

$$\begin{aligned}\mathcal{F}(u) := &\int_\Omega f\bigl(x, \mathcal{E}u(x)\bigr)\,\mathrm{d}x + \int_\Omega f^\infty\left(x, \frac{\mathrm{d}E^s u}{\mathrm{d}|E^s u|}(x)\right)\mathrm{d}|E^s u|(x)\\ &+ \int_{\partial\Omega} f^\infty\bigl(x, u|_{\partial\Omega}(x) \odot n_\Omega(x)\bigr)\,\mathrm{d}\mathcal{H}^{d-1}(x), \qquad u \in \mathrm{BD}(\Omega),\end{aligned}$$

is lower semicontinuous with respect to weak-convergence in the space* $\mathrm{BD}(\Omega)$.

The strategy for the proof hinges on a similar idea to the one outlined in the previous section for the new proof of the classical lower semicontinuity theorem in BV. While in BV this strategy merely provides a different proof of a known result, in BD we do not have an Alberti-type theorem at our disposal and so we need to rely on this new approach in order to prove a general lower semicontinuity theorem.

When trying to implement the method in BD, however, one is faced with the additional complication that the rigidity is much weaker: The natural distinction is whether $\frac{\mathrm{d}E^s u}{\mathrm{d}|E^s u|}(x_0)$ can be written as a symmetric tensor product $a \odot b := (a \otimes b + b \otimes a)/2$ for some $a, b \in \mathbb{R}^d$ or not. But in contrast to the gradient case it turns out that for $v \in \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$ the relation

$$Ev = P|Ev| \qquad \text{for a fixed matrix } P \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \text{ with } |P| = 1, \tag{1.2}$$

where $P \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is such that it cannot be written in the form $a \odot b$, does *not* imply that v is affine (not even for v smooth, see Example 5.26). In particular, blow-ups v of $u \in \mathrm{BD}(\Omega)$

(which always satisfy (1.2)) at points x_0 where $P(x_0) := \frac{\mathrm{d}E^s u}{\mathrm{d}|E^s u|}(x_0) \neq a \odot b$ for any $a, b \in \mathbb{R}^d$, might not necessarily be affine, at least this it is currently unknown. Using Fourier Analysis and an ellipticity argument, it is however possible to show that for any v solving (1.2), Ev is *absolutely continuous* with respect to Lebesgue measure and as regards blow-ups "an $\mathcal{L}^d$-absolutely continuous measure is as good as a constant multiple of $\mathcal{L}^d$". This is so because we may take a blow-up of the blow-up, which still is a blow-up of the original function (this will be used in the form that tangent measures to tangent measures are tangent measures), and this particular blow-up now indeed has a constant multiple of Lebesgue measure as its symmetrized derivative, hence it is affine.

On the other hand, at points $x_0 \in \Omega$ where $P(x_0) = a \odot b$ for some $a, b \in \mathbb{R}^d \setminus \{0\}$ with $a \neq b$, it turns out that the symmetrized derivative of any blow-up is the sum of a measure invariant under translations orthogonal to both a and b, and possibly an absolutely continuous part with *linear* density. If the linear part is non-zero, we can use the same "iterated blow-up trick" mentioned before to get an affine blow-up. If the linear part is zero, by virtue of the theory of sections of BD-functions we can show that the blow-up limit is the sum of two one-directional functions (depending only on $x \cdot a$ and $x \cdot b$, respectively), and so as before we have a well-behaved blow-up limit at our disposal, which may then be averaged (using parallelotopes with face normals a and b instead of the usual cubes) to get an affine function. The case $P(x_0) = a \odot a$ for some $a \in \mathbb{R}^d \setminus \{0\}$ is somewhat degenerate, but can also be treated with essentially the same methods (in this case, the remainder is not necessarily linear, but still vanishes in a second blow-up). The pivotal Theorem 5.6 details the construction of good blow-ups and can be considered the core of the argument. Having thus arrived at an affine function in all of the above cases, we can apply the symmetric-quasiconvexity locally.

1.3 Young measures

As the technical backbone for our results we heavily employ Young measure theory. This section aims to give a brief overview of their role in the proofs of the lower semicontinuity theorems. We confine ourselves to the situation in the space BV, but of course a similar route is also adopted in BD. A brief history of Young measures can be found in the appendix.

To motivate Young measures, let us first consider the question of $\mathrm{W}^{1,1}$-weak lower semicontinuity for the integral functional

$$\mathcal{F}(u) := \int_\Omega f\bigl(x, \nabla u(x)\bigr) \,\mathrm{d}x, \qquad u \in \mathrm{W}^{1,1}(\Omega; \mathbb{R}^m),$$

where as before $f \colon \Omega \times \mathbb{R}^{m \times d}$ is a Carathéodory integrand with linear growth. So, for a sequence $u_j \rightharpoonup u$ weakly in $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^m)$ (in particular, the sequence of gradients (∇u_j) is equiintegrable), we want to show

$$\liminf_{j \to \infty} \mathcal{F}(u_j) \geq \mathcal{F}(u).$$

The main feature of Young measure theory is that it allows us to pass to a limit in the expression $\mathcal{F}(u_j)$. Indeed, the so-called Fundamental Theorem asserts that there exists a family $(\nu_x)_{x\in\Omega}$ of probability measures carried by the matrix space $\mathbb{R}^{m\times d}$ that is (up to a non-relabeled subsequence) "generated by" the sequence of gradients (∇u_j), which entails in particular that

$$\lim_{j\to\infty} \mathcal{F}(u_j) = \int_\Omega \int_{\mathbb{R}^{m\times d}} f(x,A) \,\mathrm{d}\nu_x(A) \,\mathrm{d}x =: \int_\Omega \bigl\langle f(x,\cdot), \nu_x \bigr\rangle \,\mathrm{d}x.$$

In fact, (ν_x) does not depend on f and the above representation holds for all Carathéodory integrands $f\colon \Omega\times\mathbb{R}^{m\times d}\to\mathbb{R}$ with linear growth at infinity. Moreover, the barycenter of ν_x is $\nabla u(x)$, i.e. $\langle \mathrm{id}, \nu_x\rangle = \nabla u(x)$ almost everywhere. Then, lower semicontinuity follows if we can show the *Jensen-type inequality*

$$\begin{aligned}\int_{\mathbb{R}^{m\times d}} f(x,A)\,\mathrm{d}\nu_x(A) = \bigl\langle f(x,\cdot),\nu_x\bigr\rangle \geq f\bigl(x,\langle \mathrm{id},\nu_x\rangle\bigr) &= f\biggl(x, \int_{\mathbb{R}^{m\times d}} A \,\mathrm{d}\nu_x(A)\biggr) \\ &= f\bigl(x,\nabla u(x)\bigr)\end{aligned}$$

for almost every $x\in\Omega$. Notice that by introducing the intermediate step of passing to the Young measure limit, we have reduced the problem of lower semicontinuity to a question of *convexity* for "frozen" x. Indeed, the above Jensen-type inequality is clearly valid if $f(x,\cdot)$ is convex for almost every $x\in\Omega$. However, it turns out that for Young measures generated by a sequence of gradients, quasiconvexity suffices.

Unfortunately, the above reasoning is not satisfactory for the lower semicontinuity theorem in BV because we had to assume equiintegrability of the sequence (∇u_j), and indeed the Young measure representation above is not valid when assuming mere norm-boundedness. Once more the reason turns out to be the appearance of concentration phenomena, which we already encountered in the previous sections. Take for example the sequence of gradients $u_j' = j\mathbb{1}_{(0,1/j)}$ on the domain $\Omega = (-1,1)$. Then it can be shown that the generated Young measure (more precisely, the "biting Young measure") is $\nu_x = \delta_0$ a.e., essentially because the sequence converges to zero in measure. But clearly, for $f(x,A) := |A|$ we have $\mathcal{F}(u_j) = 1 \not\to 0 = \int_{-1}^1 \langle|\cdot|,\nu_x\rangle \,\mathrm{d}x$.

Thus, we need to extend our notion of Young measure to allow for the representation of concentration effects. This is indeed possible and was first carried out by DiPerna and Majda in connection to questions in fluid mechanics [43] and later refined by Alibert and Bouchitté [3] in a special but interesting case. In our terminology (see Chapter 2, which is based on the framework developed in [82]), the new Fundamental Theorem asserts that, only assuming a norm-bound on (∇u_j), there exist

(i) a family of probability measures $(\nu_x)_{x\in\Omega}$ carried by $\mathbb{R}^{m\times d}$,

(ii) a positive finite measure λ_ν on $\overline{\Omega}$, and

(iii) a family of probability measures $(\nu_x^\infty)_{x\in\overline{\Omega}}$ carried by the unit sphere $\partial\mathbb{B}^{m\times d}$ of $\mathbb{R}^{m\times d}$, such that, up to a subsequence,

$$\lim_{j\to\infty} \mathcal{F}(u_j) = \int_\Omega \bigl\langle f(x,\,\cdot\,), \nu_x \bigr\rangle \,\mathrm{d}x + \int_{\overline{\Omega}} \bigl\langle f^\infty(x,\,\cdot\,), \nu_x^\infty \bigr\rangle \,\mathrm{d}\lambda_\nu(x) =: \bigl\langle\!\bigl\langle f, \nu \bigr\rangle\!\bigr\rangle$$

for all Carathéodory integrands $f \colon \overline{\Omega} \times \mathbb{R}^{m\times d}$ that possess a (jointly) continuous (strong) recession function f^∞; notice the similarity to the definition of the relaxed functional $\bar{\mathcal{F}}$ in the previous section. In a more measure-theoretic fashion, this convergence can also be expressed as

$$f\bigl(x, \nabla u_j(x)\bigr)\, \mathcal{L}^d_x \llcorner \Omega \quad \overset{*}{\rightharpoonup} \quad \underbrace{\bigl\langle f(x,\,\cdot\,), \nu_x \bigr\rangle\, \mathcal{L}^d_x \llcorner \Omega}_{\text{oscillation part}} \;+\; \underbrace{\bigl\langle f^\infty(x,\,\cdot\,), \nu_x^\infty \bigr\rangle\, \lambda_\nu(x)}_{\text{concentration part}}.$$

Notice that this entails that the *barycenter*

$$[\nu] := \bigl\langle \mathrm{id}, \nu_x \bigr\rangle\, \mathcal{L}^d_x + \bigl\langle \mathrm{id}, \nu_x^\infty \bigr\rangle\, \lambda_\nu$$

of ν satisfies $[\nu] \llcorner \Omega = Du$ if $u_j \overset{*}{\rightharpoonup} u$ in $\mathrm{BV}(\Omega;\mathbb{R}^m)$.

For our example $u_j' = j\mathbb{1}_{(0,1/j)}$ from above, one can calculate $\nu_x = \delta_0$ a.e., $\lambda_\nu = \delta_0$, $\nu_0^\infty = \delta_{+1}$ (undefined for other x). Indeed, this is a precise formulation of the intuitively clear statement that the sequence (u_j') converges to zero in measure ($\nu_x = \delta_0$ a.e.), concentrates at 0 ($\lambda_\nu = \delta_0$), and does so in positive direction ($\nu_0^\infty = \delta_{+1}$, notice $\partial\mathbb{B}^{1\times 1} = \{-1,+1\}$). Therefore, the family $(\nu_x)_x$ is called the *oscillation measure*, λ_ν is the *concentration measure*, and the family $(\nu_x^\infty)_x$ is the *concentration-angle measure*.

To show lower semicontinuity in an analogous fashion to before, we now have to establish *two* Jensen-type inequalities, namely

$$\bigl\langle f(x,\,\cdot\,), \nu_x \bigr\rangle + \bigl\langle f^\infty, \nu_x^\infty \bigr\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x) \geq f\biggl(x, \bigl\langle \mathrm{id}, \nu_x \bigr\rangle + \bigl\langle \mathrm{id}, \nu_x^\infty \bigr\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x)\biggr)$$

for $\mathcal{L}^d$-almost every $x \in \Omega$, and

$$\bigl\langle f^\infty(x,\,\cdot\,), \nu_x^\infty \bigr\rangle \geq f^\infty\bigl(x, \bigl\langle \mathrm{id}, \nu_x^\infty \bigr\rangle\bigr)$$

for λ_ν^s-almost every $x \in \Omega$. If these inequalities hold, which will of course again depend on convexity properties of f, then indeed

$$\liminf_{j\to\infty} \mathcal{F}(u_j) \geq \bar{\mathcal{F}}(u),$$

which is the desired lower semicontinuity assertion.

In order to establish the above Jensen-type inequalities, we will combine the rigidity results mentioned in the first section of this introduction with localization principles for Young measures in terms of so-called regular and singular *tangent Young measures*, which encapsulate the blow-up process and contain local information about the Young measure under investigation at the blow-up point (see Propositions 3.5, 3.6). Essentially, we have to reduce to a situation where the quasiconvexity is applicable and for this we need to look at the local structure of the limit function (or generated Young measure).

One reason for choosing this Young measure approach to lower semicontinuity is that it provides a very conceptual and clean organization of the lower semicontinuity proofs, which split into an "abstract" Young measure part and the "concrete" verification of the Jensen-type inequalities. In fact, only through this point of view did it became apparent to the author how to argue without Alberti's result in the BV-lower semicontinuity theorem and also how to establish lower semicontinuity in BD without an Alberti-type theorem at our disposal.

In the opinion of the author, the most important feature of Young measure theory is that the Young measure convergence has the same good compactness properties as the weak* convergence of functions or measures (the barycenters), but passing to a Young measure limit as opposed to a classical weak* limit retains much more information. Roughly, Young measures do not average out oscillations and do not mix diffuse concentrations with parts of the sequence that converge in measure. In particular, the tangent Young measures mentioned above exist wherever tangent measures exist (which is almost everywhere), but contain more information about the blow-up sequence.

We close this section with the remark that it is also possible to completely eliminate Young measures from the new proof of the BV-lower semicontinuity theorem (see Remark 4.16 for more details). However, for the theorem in BD, the iterated blow-up technique relies heavily on the use of Young measures and they cannot be removed easily (see Remark 5.19).

1.4 Contributions of this work

The main contributions of this work are the following:

- Chapter 5 establishes the first full weak* lower semicontinuity theorem in the space BD of functions of bounded deformation.

- In Chapter 4 a new proof of the classical lower semicontinuity theorem in the space BV is presented that does not use Alberti's Rank One Theorem.

- Both proofs hinge on the apparently new idea (in the context of lower semicontinuity) to combine the classical blow-up technique with rigidity arguments.

- On a technical side, the proofs are based on Jensen-type inequalities for generalized Young measures with respect to (symmetric-)quasiconvex functions. Whereas this technique is well-known in the superlinear-growth case, it apparently has not been used in the linear-growth (BV/BD) case.

- Chapter 2 develops an improved functional-analytic framework for generalized Young measures, which is more heavily based on duality arguments; this also leads to a new proof of the Fundamental Theorem.

- Localization principles for generalized Young measures are established in Chapter 3. The resulting tangent Young measures parallel classical tangent measures, but contain more information about blow-ups.

- An important step in the proof of lower semicontinuity in BD is the construction of *good* blow-ups, which is achieved by first using rigidity arguments to prove preliminary properties of blow-up limits and then iterating the blow-up construction. This construction uses the fact that "blow-ups of blow-ups are blow-ups" and is carried out in the framework of tangent Young measures.

- We show a few new rigidity results for symmetrized gradients (the corresponding results for gradients are most likely known).

Most of the material in this work has already been published in the form of journal papers, which also form the text base for this dissertation:

- F. Rindler, *Lower semicontinuity for integral functionals in the space of functions of bounded deformation via rigidity and Young measures*, Arch. Ration. Mech. Anal. **202** (2011), 63–113:

  Lower semicontinuity in BD (Chapter 5) and localization principles for Young measures (Chapter 3).

- F. Rindler, *Lower semicontinuity and Young measures in BV without Alberti's Rank One Theorem*, Adv. Calc. Var., to appear (Ahead of Print: DOI 10.1515/ACV.2011.008; submitted February 2010):

  Lower semicontinuity in BV (Chapter 4).

- J. Kristensen and F. Rindler, *Characterization of generalized gradient Young measures generated by sequences in $W^{1,1}$ and BV*, Arch. Ration. Mech. Anal. **197** (2010), 539–598:

  Functional-analytic framework for Young measures (Chapter 2); the main characterization result of that work as well as the averaging, shifting and approximation theorems for generalized Young measures are not part of this thesis.

- J. Kristensen and F. Rindler, *Relaxation of signed integral functionals in BV*, Calc. Var. Partial Differential Equations **37** (2010), 29–62:

  A few minor auxiliary results; the main lower semicontinuity result from that work is not part of this thesis.

Chapter 2

Young measures and gradient Young measures

In the introduction we have already seen an outline of how Young measures can be used to prove lower semicontinuity theorems. This chapter develops the necessary theory to make these arguments precise. After presenting some preliminaries on measure theory and parametrized measures in the first sections, we will formally define generalized Young measures within the improved functional analytic framework from [82], placing a heavy emphasis on duality methods. Then, we investigate some of their main properties and connections to different convergence notions for sequences. Finally, we treat in some detail the important case of Young measures generated by sequences of $\mathrm{W}^{1,1}$-gradients or BV-derivatives.

2.1 Preliminaries

In this section we introduce notation and collect a few preliminary results.

2.1.1 General notation

The symbol $\mathbb{B}^N$ stands for the open unit ball in $\mathbb{R}^N$, $\mathbb{S}^{N-1}$ or $\partial\mathbb{B}^N$ denote the corresponding unit sphere, and $B(x_0, r)$ is the open ball centered at x_0 with radius $r > 0$. The matrix space $\mathbb{R}^{m\times d}$ will always be equipped with the Frobenius norm $|A| := (\sum_{i=1}^m \sum_{j=1}^d A^i_j)^{1/2}$ (i and j are the row and column indices, respectively); $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ and $\mathbb{R}^{d\times d}_{\mathrm{skew}}$ are the subspaces of symmetric and skew-symmetric matrices, respectively. The tensor product $a \otimes b$, where $a \in \mathbb{R}^m$, $b \in \mathbb{R}^d$, designates the $(m \times d)$-matrix ab^T, whereas the symmetric tensor product is defined for two vectors $a, b \in \mathbb{R}^d$ as $a \odot b := (a \otimes b + b \otimes a)/2$.

If X is a Borel subset of $\mathbb{R}^d$, let $\mathrm{C}_c(X)$ be the space of continuous functions with compact support in X and let $\mathrm{C}_0(X)$ be its completion with respect to the $\|\bullet\|_\infty$-norm. The space $\mathrm{C}_0(X)$, equipped with this norm, is a (separable) Banach space (note that if X is compact, this does not imply "zero boundary values"). By $\mathrm{C}_c(X;\mathbb{R}^N)$ and $\mathrm{C}_0(X;\mathbb{R}^N)$ we denote the corresponding spaces of functions with values in $\mathbb{R}^N$. The tensor product $f \otimes g$ of two real-valued functions f and g is defined through $(f \otimes g)(x,y) := f(x)g(y)$. By $\mathbb{1}$ and $\mathbb{1}_U$ we understand the constant 1-function and the characteristic function of U, respectively.

We use the Lebesgue spaces $\mathrm{L}^p(\Omega)$, $\mathrm{L}^p(\Omega;\mathbb{R}^m)$, $\mathrm{L}^p_{\mathrm{loc}}(\Omega)$, $\ldots$, where $p \in [1,\infty]$ and $\Omega \subset \mathbb{R}^d$ is an open set, and the Sobolev spaces $\mathrm{W}^{1,p}(\Omega)$, $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^m)$, $\mathrm{W}^{1,p}_0(\Omega)$, $\mathrm{W}^{1,p}_g(\Omega)$, $\mathrm{W}^{1,p}_{\mathrm{loc}}(\Omega), \ldots$ with their usual meanings.

2.1.2 Measure theory

In the following, we briefly recall some of the notions from measure theory employed in this work. More information about these topics can be found in [10, 55, 89].

We denote the Borel σ-algebra on a set $E \subset \mathbb{R}^d$ by $\mathcal{B}(E)$. The space $\mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$ contains all $\mathbb{R}^N$-valued set-functions that are defined on the relatively compact Borel subsets and that are σ-additive and finite when restricted to the Borel σ-algebra on a compact subset of $\mathbb{R}^d$. We call its elements vector-valued **local (Radon) measures**. The space $\mathbf{M}(\mathbb{R}^d;\mathbb{R}^N)$ contains all $\mathbb{R}^N$-valued **finite (Radon) measures** on the Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$ on $\mathbb{R}^d$ with values in $\mathbb{R}^N$. Positive measures are contained in the analogous spaces $\mathbf{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ and $\mathbf{M}^+(\mathbb{R}^d)$, respectively. A probability measure is a positive measure $\mu \in \mathbf{M}^+(\mathbb{R}^d)$ with $\mu(\mathbb{R}^d) = 1$, we write $\mu \in \mathbf{M}^1(\mathbb{R}^d)$. We will also employ the spaces $\mathbf{M}(X;\mathbb{R}^N)$, $\mathbf{M}_{\mathrm{loc}}(X;\mathbb{R}^N)$, $\mathbf{M}^+(X;\mathbb{R}^N)$, $\mathbf{M}^1(X;\mathbb{R}^N)$ with a Borel set $X \subset \mathbb{R}^d$ replacing $\mathbb{R}^d$; all of the following statements, with the appropriate adjustments, also hold for these spaces.

For every local measure $\mu \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$, we denote by $|\mu| \in \mathbf{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ its **total variation measure**. The restriction of a (local) measure $\mu \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$ to a Borel set $A \subset \mathbb{R}^d$ is written as $\mu \llcorner A$ and defined by $(\mu \llcorner A)(B) := \mu(B \cap A)$ for all relatively compact Borel sets $B \subset \mathbb{R}^d$. For a positive measure $\mu \in \mathbf{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$, the support $\operatorname{supp}\mu$ is the set of all $x \in \mathbb{R}^d$ such that $\mu(B(x,r)) > 0$ for all $r > 0$, which is always a closed set. For a vector measure $\mu \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$, the support of μ is simply set to be the support of $|\mu|$.

Lebesgue measure in $\mathbb{R}^d$ is denoted by $\mathcal{L}^d$, sometimes augmented to $\mathcal{L}^d_x$ to give a name to the integration variable. For a Lebesgue-measurable set $A \subset \mathbb{R}^d$, we will often simply write $|A|$ instead of $\mathcal{L}^d(A)$; ω_d denotes the volume of the unit ball in $\mathbb{R}^d$. The symbol $\mathcal{H}^k$ stands for the k-dimensional Hausdorff outer measure, $k \in [0,\infty)$. When restricted to a $\mathcal{H}^k$-rectifiable set $S \subset \mathbb{R}^d$ (see Section 2.9 of [10], we only need the fact that Lipschitz boundaries of open sets in $\mathbb{R}^d$ are $\mathcal{H}^{d-1}$-rectifiable), $\mathcal{H}^k \llcorner S$ is a local Radon measure.

The pairing $\langle f, \mu \rangle$ between a Borel-measurable function $f \colon \mathbb{R}^d \to \mathbb{R}^N$ and a positive measure $\mu \in \mathbf{M}^+(\mathbb{R}^d)$, or, if f has compact support also with $\mu \in \mathbf{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$, is defined as

$$\langle f, \mu \rangle := \int f \, \mathrm{d}\mu \quad (\in \mathbb{R}^N),$$

provided this integral exists in the sense of Lebesgue.

Every measure $\mu \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$ has an (essentially unique) **Lebesgue–Radon–Nikodým decomposition** $\mu = \frac{\mathrm{d}\mu}{\mathrm{d}\lambda}\lambda + \mu^s$ with respect to a positive measure $\lambda \in \mathbf{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$, i.e.

$$\mu(A) = \int_A \frac{\mathrm{d}\mu}{\mathrm{d}\lambda} \, \mathrm{d}\lambda + \mu^s(A)$$

for all relatively compact Borel sets $A \subset \mathbb{R}^d$. In this decomposition, μ^s and λ are mutually singular, i.e. concentrated on mutually negligible sets. The function $\frac{\mathrm{d}\mu}{\mathrm{d}\lambda} \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d, \lambda; \mathbb{R}^N)$ (the Lebesgue space with respect to the measure λ) is called the **Radon–Nikodým derivative** (or **density**) of μ with respect to λ and may be computed by

$$\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}(x_0) = \lim_{r \downarrow 0} \frac{\mu(B(x_0, r))}{\lambda(B(x_0, r))} \qquad \text{for } \lambda\text{-a.e. } x_0 \in \operatorname{supp} \lambda.$$

If not otherwise specified, μ^s will always mean the singular part of the measure μ with respect to Lebesgue measure. The function $\frac{\mathrm{d}\mu}{\mathrm{d}|\mu|} \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d, |\mu|; \mathbb{R}^N)$ is called the **polar function** of μ and satisfies $\left|\frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x)\right| = 1$ at $|\mu|$-almost every $x \in \mathbb{R}^d$.

Several times we will employ the **pushforward** $T_*\mu := \mu \circ T^{-1} \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$ of a local measure $\mu \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$ under an affine map $T\colon \mathbb{R}^d \to \mathbb{R}^d$, $x \mapsto x_0 + Lx$, where $x_0 \in \mathbb{R}^d$ and $L \in \mathbb{R}^{d \times d}$ is an invertible matrix (of course, pushforwards are defined for more general T, but we will not need those). For a measurable function $f\colon \mathbb{R}^d \to \mathbb{R}^N$ and $\mu \in \mathbf{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ we have the transformation rule

$$\bigl\langle f, T_*\mu \bigr\rangle = \int f \,\mathrm{d}(T_*\mu) = \int f \circ T \,\mathrm{d}\mu = \bigl\langle f \circ T, \mu \bigr\rangle$$

provided one, hence both, of these integrals are well-defined. Also, with $\det T := \det L$, we have the following formulas for densities:

$$\frac{\mathrm{d}T_*\mu}{\mathrm{d}\mathcal{L}^d} = |\det T|^{-1} \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^d} \circ T^{-1}, \qquad \frac{\mathrm{d}T_*\mu}{\mathrm{d}|T_*\mu|} = \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|} \circ T^{-1}. \tag{2.1}$$

Mostly, we will use pushforwards under the blow-up transformation $T^{(x_0,r)}(x) := (x - x_0)/r$, where $x_0 \in \mathbb{R}^d$ and $r > 0$. For this particular transformation we have $|\det T^{(x_0,r)}|^{-1} = r^d$.

By the Riesz Representation Theorem, we may consider $\mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$ as the dual space to the locally convex space $\mathrm{C}_c(\mathbb{R}^d; \mathbb{R}^N)$ and $\mathbf{M}(X; \mathbb{R}^N)$ as the dual space to the Banach space $\mathrm{C}_0(X; \mathbb{R}^N)$, where as before X is a Borel subset of $\mathbb{R}^d$. These dualities induce the **(local) weak* convergence** $\mu_j \overset{*}{\rightharpoonup} \mu$ in $\mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$ defined as $\langle \psi, \mu_j \rangle \to \langle \psi, \mu \rangle$ (in $\mathbb{R}^N$) for all $\psi \in \mathrm{C}_c(\mathbb{R}^d)$ as well as the **weak* convergence** $\mu_j \overset{*}{\rightharpoonup} \mu$ in $\mathbf{M}(X; \mathbb{R}^N)$ meaning $\langle \psi, \mu_j \rangle \to \langle \psi, \mu \rangle$ for all $\psi \in \mathrm{C}_0(X)$. Both convergences (we only work with convergences here, not with topologies) have good compactness properties. In particular, every sequence $(\mu_j) \subset \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$ satisfying $\sup_j |\mu_j|(K) < \infty$ for all compact $K \subset \mathbb{R}^d$ has a (locally) weakly* converging subsequence. Likewise, if for a sequence $(\mu_j) \subset \mathbf{M}(X; \mathbb{R}^N)$ we have $\sup_j |\mu_j|(X) < \infty$, then this sequence is (sequentially) weakly* relatively compact.

Since the **strong (norm) convergence** of a sequence $(\mu_j) \subset \mathbf{M}(\mathbb{R}^d; \mathbb{R}^N)$ to $\mu \in \mathbf{M}(\mathbb{R}^d; \mathbb{R}^N)$, i.e. $|\mu_j - \mu|(\mathbb{R}^d) \to 0$, is too strong for most applications, one introduces two intermediate convergences, which are often useful: The first is the so-called **strict convergence** and requires that $\mu_j \overset{*}{\rightharpoonup} \mu$ and $|\mu_j|(\mathbb{R}^d) \to |\mu|(\mathbb{R}^d)$. With the **area functional** $\langle \cdot \rangle(\mathbb{R}^d)\colon \mathbf{M}(\mathbb{R}^d; \mathbb{R}^N) \to \mathbb{R}$,

$$\langle \mu \rangle(\mathbb{R}^d) := \int \sqrt{1 + \left|\frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^d}(x)\right|^2} \,\mathrm{d}x + |\mu^s|(\mathbb{R}^d), \qquad \mu \in \mathbf{M}(\mathbb{R}^d; \mathbb{R}^N),$$

we furthermore define the **$\langle\bullet\rangle$-strict (area-strict) convergence** in $\mathbf{M}(\mathbb{R}^d;\mathbb{R}^N)$ to comprise $\mu_j \overset{*}{\rightharpoonup} \mu$ and $\langle\mu_j\rangle(\mathbb{R}^d) \to \langle\mu\rangle(\mathbb{R}^d)$, see [39, 82, 83] and also Counterexample 2.4 below for a discussion why the $\langle\bullet\rangle$-strict convergence is important in our situation. It can be shown (by mollification) that smooth measures are dense in $\mathbf{M}(\mathbb{R}^d;\mathbb{R}^N)$ with respect to the $\langle\bullet\rangle$-strict convergence. Notice that by Reshetnyak's Continuity Theorem 2.3 below, $\langle\bullet\rangle$-strict convergence is stronger than strict convergence.

2.1.3 Parametrized measures

For Borel sets $E \subset \mathbb{R}^d$, $X \subset \mathbb{R}^N$, a **parametrized measure** $(\nu_x)_{x\in E} \subset \mathbf{M}(X)$ is a mapping from E to the set $\mathbf{M}(X)$ of finite measures on X. It is said to be **weakly* μ-measurable**, where $\mu \in \mathbf{M}^+(E)$, if $x \mapsto \nu_x(B)$ is μ-measurable for all Borel sets $B \in \mathcal{B}(X)$ (it suffices to check this for all relatively open sets). Here, μ-measurability is understood to mean measurability with respect to the μ-completion of the Borel σ-algebra $\mathcal{B}(E)$ on E. Equivalently, $(\nu_x)_{x\in E}$ is weakly* μ-measurable if the function $x \mapsto \int_X f(x,y)\,\mathrm{d}\nu_x(y)$ is μ-measurable for every bounded Borel function $f\colon E\times X \to \mathbb{R}$, see Proposition 2.26 in [10] for this equivalence. We denote by $\mathrm{L}^\infty_{w*}(E,\mu;\mathbf{M}(X))$ the set of all weakly* μ-measurable parametrized measures $(\nu_x)_{x\in E} \subset \mathbf{M}(X)$ with the property that $\operatorname{ess\,sup}_{x\in E} |\nu_x|(X) < \infty$, the essential supremum taken with respect to μ. We write $\mathrm{L}^\infty_{w*}(E;\mathbf{M}(X))$ if $\mu = \mathcal{L}^d$. See [41, 48] for details on these spaces.

The **generalized product**, denoted by $\mu\otimes\nu_x$, of a positive Radon measure $\mu \in \mathbf{M}^+(E)$ and a weakly* μ-measurable parametrized measure $(\nu_x)_{x\in E} \subset \mathbf{M}(X)$ is defined through

$$(\mu\otimes\nu_x)(U) := \int_E\int_X \mathbb{1}_U(x,y)\,\mathrm{d}\nu_x(y)\,\mathrm{d}\mu(x), \qquad U \in \mathcal{B}(E\times X). \tag{2.2}$$

By approximation with simple functions, the following integration formula can be proved for all Borel-measurable $f \in \mathrm{L}^1(E\times X, \mu\otimes\nu_x)$:

$$\int_{E\times X} f\,\mathrm{d}(\mu\otimes\nu_x) = \int_E\int_X f(x,y)\,\mathrm{d}\nu_x(y)\,\mathrm{d}\mu(x).$$

We will later need the following special case of a well-known disintegration theorem (also called slicing or layerwise decomposition):

Theorem 2.1 (Disintegration of measures). *Let $\nu \in \mathbf{M}^+(E\times X)$, where $E \subset \mathbb{R}^d$, $X \subset \mathbb{R}^N$ are Borel sets, and define $\kappa \in \mathbf{M}^+(E)$ through $\kappa(B) := \nu(B\times X)$ for all $B \in \mathcal{B}(E)$. Then, there exists a κ-essentially unique, weakly* κ-measurable family $(\eta_x)_{x\in E} \subset \mathbf{M}^1(X)$ of probability measures on X such that $\nu = \kappa\otimes\eta_x$.*

See for example Theorem 2.28 in [10] for a proof.

2.2 Integrands

In the following, we deal with two classes of integrands. The first class $\mathbf{E}(\Omega;\mathbb{R}^N)$ will turn out to be the natural choice for the functional analytic setup for Young measures, while the second class, denoted $\mathbf{R}(\Omega;\mathbb{R}^N)$, is larger and contains integrands for which we will be able to compute limits.

Let again $\Omega \subset \mathbb{R}^d$ be a bounded open Lipschitz domain. In the first class $\mathbf{E}(\Omega;\mathbb{R}^N)$, we assume our integrands $f \in \mathrm{C}(\overline{\Omega} \times \mathbb{R}^N)$ to be continuous. To state the crucial growth property required of f, we need the following linear transformations mapping $\mathrm{C}(\overline{\Omega} \times \mathbb{R}^N)$ to $\mathrm{C}(\overline{\Omega} \times \mathbb{B}^N)$ and back: For $f \in \mathrm{C}(\overline{\Omega} \times \mathbb{R}^N)$ and $g \in \mathrm{C}(\overline{\Omega} \times \mathbb{B}^N)$ define

$$\begin{aligned} (Sf)(x,\hat{A}) &:= (1-|\hat{A}|)\, f\biggl(x, \frac{\hat{A}}{1-|\hat{A}|}\biggr), && x \in \overline{\Omega},\ \hat{A} \in \mathbb{B}^N, \quad \text{and} \\ (S^{-1}g)(x,A) &:= (1+|A|)\, g\biggl(x, \frac{A}{1+|A|}\biggr), && x \in \overline{\Omega},\ A \in \mathbb{R}^N. \end{aligned} \tag{2.3}$$

It is an easy calculation to verify $S^{-1}Sf = f$ and $SS^{-1}g = g$. Now, the property we require of f is the following:

$$Sf \text{ extends into a bounded continuous function on } \overline{\Omega \times \mathbb{B}^N}. \tag{2.4}$$

In particular, this entails that f has **linear growth at infinity**, i.e. there exists a constant $M \geq 0$ (in fact, $M = \|Sf \colon \overline{\Omega \times \mathbb{B}^N}\|_\infty$ is the smallest such constant) such that

$$|f(x,A)| \leq M(1+|A|) \qquad \text{for all } x \in \overline{\Omega},\ A \in \mathbb{R}^N. \tag{2.5}$$

We collect all such integrands into the set

$$\mathbf{E}(\Omega;\mathbb{R}^N) := \bigl\{\, f \in \mathrm{C}(\overline{\Omega} \times \mathbb{R}^N) \ \ \mathbf{:}\ \ f \text{ satisfies (2.4)} \,\bigr\}.$$

Under the natural norm

$$\|f\|_{\mathbf{E}(\Omega;\mathbb{R}^N)} := \|Sf \colon \overline{\Omega \times \mathbb{B}^N}\|_\infty = \sup_{(x,\hat{A}) \in \overline{\Omega \times \mathbb{B}^N}} |Sf(x,\hat{A})|, \qquad f \in \mathbf{E}(\Omega;\mathbb{R}^N),$$

the space $\mathbf{E}(\Omega;\mathbb{R}^N)$ 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 when $\mathrm{C}(\overline{\Omega \times \mathbb{B}^N})$ is equipped with the supremum norm. In particular, it follows that $\mathbf{E}(\Omega;\mathbb{R}^N)$ is separable.

For all $f \in \mathbf{E}(\Omega;\mathbb{R}^N)$, the **(strong) recession function** $f^\infty \colon \overline{\Omega} \times \mathbb{R}^N \to \mathbb{R}$,

$$f^\infty(x,A) := \lim_{\substack{x' \to x \\ A' \to A \\ t \to \infty}} \frac{f(x',tA')}{t}, \qquad x \in \overline{\Omega},\ A \in \mathbb{R}^N, \tag{2.6}$$

exists and takes finite values. Moreover, it follows easily that the function f^∞ is **positively 1-homogeneous** in A (i.e. $f^\infty(x, \theta A) = \theta f^\infty(x, A)$ for all $\theta \geq 0$) and agrees with Sf on $\overline{\Omega} \times \mathbb{S}^{N-1}$ (to see the latter, substitute $t = s/(1-s)$, $s \in (0,1)$ and let $s \to 1$). We refer to Counterexample 2.19 for the necessity of including the convergence $x' \to x$ in the definition of f^∞.

Membership in $\mathbf{E}(\Omega; \mathbb{R}^N)$ can equivalently be expressed by requiring that f^∞ exists in the sense of (2.6). Also, it is not difficult to see that every continuous function that is either uniformly bounded or positively 1-homogeneous in its second argument belongs to $\mathbf{E}(\Omega; \mathbb{R}^N)$, cf. Lemma 2.2 in [3].

The space $\mathbf{E}_c(\mathbb{R}^d; \mathbb{R}^N)$ is defined similarly to $\mathbf{E}(\mathbb{R}^d; \mathbb{R}^N)$, but additionally we require that for each element $f \in \mathbf{E}_c(\mathbb{R}^d; \mathbb{R}^N) \subset \mathrm{C}(\mathbb{R}^d \times \mathbb{R}^N)$ there exists a compact set $K \subset \mathbb{R}^d$ such that $\operatorname{supp} f(\,\cdot\,, A) \subset K$ for all $A \in \mathbb{R}^N$.

The second class of integrands is larger than $\mathbf{E}(\Omega; \mathbb{R}^N)$ and will (partially) dispense with the continuity in the x-variable. We recall that a **Carathéodory function** is an $\mathcal{L}^d \times \mathcal{B}(\mathbb{R}^N)$-measurable function $f \colon \overline{\Omega} \times \mathbb{R}^N \to \mathbb{R}$ such that $A \mapsto f(x, A)$ is continuous for almost every $x \in \overline{\Omega}$. In fact, it can be shown that it suffices to check measurability of $x \mapsto f(x, A)$ for all fixed $A \in \mathbb{R}^N$ (see for example Proposition 5.6 in [10]). An important property of these functions f is that the composed function $x \mapsto f(x, v(x))$ is Lebesgue-measurable whenever $v \colon \Omega \to \mathbb{R}^N$ is. With this notion, the **representation integrands** are defined as follows:

$$\mathbf{R}(\Omega; \mathbb{R}^N) := \Big\{ f \colon \overline{\Omega} \times \mathbb{R}^N \to \mathbb{R} \;\;\mathbf{:}\;\; \begin{array}{l} f \text{ Carathéodory with linear growth at infinity} \\ \text{and it exists } f^\infty \in \mathrm{C}(\overline{\Omega} \times \mathbb{R}^N) \end{array} \Big\}.$$

Here, the existence of f^∞ is understood in the sense of (2.6). Note that we do *not* identify integrands that are equal almost everywhere.

If we only have $f \in \mathrm{C}(\overline{\Omega} \times \mathbb{R}^N)$ with linear growth at infinity, the recession function f^∞ does not necessarily exist (not even for *quasiconvex* $f \in \mathrm{C}(\mathbb{R}^N)$, see Theorem 2 of [95]). But for such functions f we can always define the **generalized recession function** $f^\# \colon \overline{\Omega} \times \mathbb{R}^N \to \mathbb{R}$ by

$$f^\#(x, A) := \limsup_{\substack{x' \to x \\ A' \to A \\ t \to \infty}} \frac{f(x', tA')}{t}, \qquad x \in \overline{\Omega},\ A \in \mathbb{R}^N, \tag{2.7}$$

which again is always positively 1-homogeneous and the linear growth at infinity of f suffices for $f^\#$ to take only real values. In other works, $f^\#$ is usually just called the “recession function”, but here the distinction to f^∞ is important. It is elementary to show that $f^\#(x, \,\cdot\,)$ is always upper semicontinuous. For a convex function f, always $f^\# = f^\infty$. We refer to [12], in particular Section 2.5, for a more systematic approach to recession functions and their associated cones.

If f is Lipschitz continuous in its second argument (with uniform Lipschitz constant), then the definitions of f^∞ and $f^\#$ simplify to

$$\begin{aligned} f^\infty(x, A) &= \lim_{\substack{x' \to x \\ t \to \infty}} \frac{f(x', tA)}{t}, \\ f^\#(x, A) &= \limsup_{\substack{x' \to x \\ t \to \infty}} \frac{f(x', tA)}{t}, \qquad x \in \overline{\Omega},\ A \in \mathbb{R}^N. \end{aligned} \tag{2.8}$$

Later, we will need the following approximation lemma, first proved in Lemma 2.3 of [3]:

Lemma 2.2. *For every function $f \in \mathrm{C}(\overline{\Omega} \times \mathbb{R}^N)$ with linear growth at infinity there exists 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, the linear growth constants of the f_n can be chosen to be the same as the linear growth constant of f.

Proof. This proof is adapted from the one in the appendix of [83].

Let $(Sf)^{\mathrm{usc}} \colon \overline{\Omega \times \mathbb{B}^N} \to \mathbb{R}$ be the upper semicontinuous extension of $Sf \colon \overline{\Omega} \times \mathbb{B}^N \to \mathbb{R}$. If $|f(x, A)| \le M(1 + |A|)$ for some $M \ge 0$ and all $x \in \overline{\Omega}$ and $A \in \mathbb{R}^N$, then $|Sf| \le M$, and so $|(Sf)^{\mathrm{usc}}| \le M$ as well. Since Sf is upper semicontinuous (even continuous) on $\overline{\Omega} \times \mathbb{B}^N$, $(Sf)^{\mathrm{usc}}|_{\overline{\Omega}\times\mathbb{B}^N} = Sf$.

For any sequences $x_n \to x$ in $\overline{\Omega}$, $(\hat{A}_n) \subset \mathbb{B}^N$, $\hat{A}_n \to \hat{A} \in \partial\mathbb{B}^N$, and $t_n \to \infty$, we observe

$$\limsup_{n\to\infty} \frac{f(x_n, t_n\hat{A}_n)}{t_n} = \limsup_{n\to\infty} (t_n^{-1} + |\hat{A}_n|) Sf\Big(x_n, \frac{\hat{A}_n}{t_n^{-1} + |\hat{A}_n|}\Big).$$

Hence, $f^\# = (Sf)^{\mathrm{usc}}|_{\overline{\Omega}\times\mathbb{S}^{N-1}}$. More precisely, "$\le$" follows from the upper semicontinuity of $(Sf)^{\mathrm{usc}}$ and "$\ge$" by taking a (maximizing) 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 employing $t_n := (1 - |\hat{A}_n|)^{-1}$.

Now take a decreasing sequence $(g_k) \subset \mathrm{C}(\overline{\Omega \times \mathbb{B}^N})$ satisfying $|g_k| \le M$ and $g_k \downarrow (Sf)^{\mathrm{usc}}$ in $\overline{\Omega \times \mathbb{B}^N}$ and set $f_k := S^{-1}g_k$. Then, $f^k \downarrow f$, f_k^∞ exists in the sense of (2.6) and we have $f_k^\infty = g_k|_{\overline{\Omega}\times\mathbb{S}^{N-1}} \downarrow (Sf)^{\mathrm{usc}}|_{\overline{\Omega}\times\mathbb{S}^{N-1}} = f^\#$. $\square$

Finally, we record a version of Reshetnyak's Continuity Theorem.

Theorem 2.3 (Reshetnyak Continuity Theorem). *For a sequence $(\mu_j) \subset \mathbf{M}(\mathbb{R}^d; \mathbb{R}^N)$ and $\mu \in \mathbf{M}(\mathbb{R}^d; \mathbb{R}^N)$ assume $\mu_j \to \mu$ with respect to the $\langle\boldsymbol{\cdot}\rangle$-strict convergence. Then,*

$$\begin{aligned} &\int f\Big(x, \frac{\mathrm{d}\mu_j}{\mathrm{d}\mathcal{L}^d}(x)\Big)\,\mathrm{d}x + \int f^\infty\Big(x, \frac{\mathrm{d}\mu_j^s}{\mathrm{d}|\mu_j^s|}(x)\Big)\,\mathrm{d}|\mu_j^s|(x) \\ &\qquad \to \int f\Big(x, \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^d}(x)\Big)\,\mathrm{d}x + \int f^\infty\Big(x, \frac{\mathrm{d}\mu^s}{\mathrm{d}|\mu^s|}(x)\Big)\,\mathrm{d}|\mu^s|(x) \end{aligned}$$

for all $f \in \mathbf{E}(\mathbb{R}^d; \mathbb{R}^N)$.

Our version differs from Reshetnyak's original statement in Theorem 3 of [106] (also see Theorem 2.39 in [10]) in that the integrands considered by Reshetnyak were positively 1-homogeneous and he only needed to require strict convergence of the μ_js. We show how our statement can be reduced to that of Reshetnyak; this proof is taken from the appendix of [83].

Proof that Theorem 2.3 follows from [106]. For an integrand $f \in \mathbf{E}(\mathbb{R}^d; \mathbb{R}^N)$ define the **perspective integrand** $\tilde{f} \colon \mathbb{R}^d \times \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}$ by

$$\tilde{f}(x, A, s) := \begin{cases} sf\bigl(x, s^{-1}A\bigr) & \text{if } s \neq 0, \\ f^\infty(x, A) & \text{if } s = 0, \end{cases}$$

where $x \in \mathbb{R}^d$, $A \in \mathbb{R}^N$ and $s \in \mathbb{R}$. Note that by assumption $\tilde{f}$ is continuous and $(A, s) \mapsto f(x, A, s)$ is positively 1-homogeneous (jointly in (A, s)) for each fixed $x \in \mathbb{R}^d$. Next, for $\mu \in \mathbf{M}(\mathbb{R}^d; \mathbb{R}^N)$ we let $\tilde{\mu} \in \mathbf{M}(\mathbb{R}^d; \mathbb{R}^N \times \mathbb{R})$ be defined as

$$\tilde{\mu}(B) := \bigl(\mu(B), \mathcal{L}^d(B)\bigr) \qquad \text{for Borel sets } B \in \mathcal{B}(\mathbb{R}^d).$$

If $\mu = a\mathcal{L}^d + \mu^s$ is the Lebesgue–Radon–Nikodým decomposition of μ, then

$$\tilde{\mu} = (a, 1)\,\mathcal{L}^d + (1, 0)\,\mu^s$$

is the Lebesgue–Radon–Nikodým decomposition of $\tilde{\mu}$. The decisive fact is now that

$$|\tilde{\mu}|(\mathbb{R}^d) = \int |(a, 1)| \,\mathrm{d}x + \mu^s(\mathbb{R}^d) = \langle \mu \rangle(\mathbb{R}^d)$$

and hence $\langle\boldsymbol{\cdot}\rangle$-strict convergence of a sequence $(\mu_j) \subset \mathbf{M}(\mathbb{R}^d; \mathbb{R}^N)$ implies strict convergence of the corresponding $(\tilde{\mu}_j)$. Thus, by Reshetnyak's original result, Theorem 3 in [106] (or Theorem 2.39 in [10]), the mapping

$$\mu \mapsto \int f\biggl(x, \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^d}(x)\biggr) \,\mathrm{d}x + \int f^\infty\biggl(x, \frac{\mathrm{d}\mu^s}{\mathrm{d}|\mu^s|}(x)\biggr) \,\mathrm{d}|\mu^s|(x) = \int \tilde{f}\biggl(x, \frac{\mathrm{d}\tilde{\mu}}{\mathrm{d}|\tilde{\mu}|}(x)\biggr) \,\mathrm{d}|\tilde{\mu}|(x)$$

is $\langle\boldsymbol{\cdot}\rangle$-strictly continuous. □

Counterexample 2.4. That indeed we need $\langle\boldsymbol{\cdot}\rangle$-strict convergence instead of mere strict convergence for the extension of Reshetnyak's Theorem to hold can be seen by considering the sequence $\mu_j := [1 + \sin(2\pi j x)]\,\mathcal{L}^d \llcorner (0, 1)$, which converges strictly, but not $\langle\boldsymbol{\cdot}\rangle$-strictly, and so for this sequence (μ_j) the conclusion of Reshetnyak's Theorem fails in particular for the integrand $f(A) := \sqrt{1 + |A|^2}$.

2.3 Generalized Young measures

The framework for generalized Young measures presented here is more functional-analytic in spirit than [3, 43] and its key point is that (generalized) Young measures are just special representations of certain functionals from the dual space to $\mathbf{E}(\Omega; \mathbb{R}^N)$, we refer to [35] for a similar treatment. Moreover, we are inspired by the theory of classical Young measures in the papers [25, 114] (note in particular Theorem 3.6 in [114]) in the sense that we first prove compactness, which in our situation is a straightforward consequence of a transformation of Young measures to the space $\mathrm{C}(\overline{\Omega \times \mathbb{B}^N})$ and the Banach–Alaoglu–Bourbaki Theorem, and then establish the Fundamental Theorem on generation of Young measures as a simple corollary.

We work in a relatively concrete setting with $\mathbb{R}^N$ as target space, which is then compactified by the sphere compactification, but we remark that more general target spaces and compactifications can be used too. For example, instead of $\mathbb{R}^N$ one could take a Banach space X with the analytic Radon–Nikodým property, i.e. the validity of the Radon–Nikodým Theorem for vector measures with values in X; in fact, even more general spaces are possible, cf. [13]. Even when sticking to $\mathbb{R}^N$ as target space, larger compactifications are conceivable, in particular those generated by separable complete rings of continuous bounded functions, see Section 4.8 in [53] and also [30], or even the Stone–Čech compactification $\beta\mathbb{R}^N$. With these abstract compactifications, however, the Young measure representation is much less concrete. In some applications it is also useful to take a general finite measure space instead of Ω with Lebesgue measure. We refer to [3, 13, 29, 43, 85] for further information on such generalizations.

2.3.1 Basic definitions

A **(generalized) Young measure** on the open set $\Omega \subset \mathbb{R}^d$ and with values in $\mathbb{R}^N$ is a triple $(\nu_x, \lambda_\nu, \nu_x^\infty)$ consisting of

(i) a parametrized family of probability measures $(\nu_x)_{x\in\Omega} \subset \mathbf{M}^1(\mathbb{R}^N)$,

(ii) a positive finite measure $\lambda_\nu \in \mathbf{M}^+(\overline{\Omega})$ and

(iii) a parametrized family of probability measures $(\nu_x^\infty)_{x\in\overline{\Omega}} \subset \mathbf{M}^1(\mathbb{S}^{N-1})$.

Moreover, we require that

(iv) the map $x \mapsto \nu_x$ is weakly* measurable with respect to $\mathcal{L}^d$, i.e. the function $x \mapsto \langle f(x, \cdot), \nu_x\rangle$ is $\mathcal{L}^d$-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, and

(vi) $x \mapsto \langle |\cdot|, \nu_x\rangle \in \mathrm{L}^1(\Omega)$.

The set $\mathbf{Y}(\Omega;\mathbb{R}^N)$ contains all these Young measures. We will see that it suffices to define the maps $x \mapsto \nu_x$ and $x \mapsto \nu_x^\infty$ only up to $\mathcal{L}^d$- and λ_ν-negligible sets, respectively. So, equivalently to (i)–(vi), we might require

$$
\begin{aligned}
&(\nu_x) \in \mathrm{L}^\infty_{w*}(\Omega;\mathbf{M}^1(\mathbb{R}^N)), && \lambda_\nu \in \mathbf{M}^+(\overline{\Omega}),\\
&(\nu_x^\infty) \in \mathrm{L}^\infty_{w*}(\overline{\Omega},\lambda_\nu;\mathbf{M}^1(\mathbb{S}^{N-1})), && x \mapsto \big\langle |\cdot|, \nu_x \big\rangle \in \mathrm{L}^1(\Omega).
\end{aligned}
$$

The parametrized measure (ν_x) is called the **oscillation measure**, the measure λ_ν is the **concentration measure**, and (ν_x^∞) is the **concentration-angle measure**. We will later see that indeed the concentration measure quantifies the location and magnitude of concentration while the concentration-angle measure describes its direction. Often, we will simply write ν when in fact we mean the triple $(\nu_x, \lambda_\nu, \nu_x^\infty)$.

Similarly, we define the space $\mathbf{Y}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$, but with λ_ν only a local measure and $x \mapsto \langle |\cdot|, \nu_x\rangle \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$. However, in most of the following we restrict ourselves to presenting the theory for $\mathbf{Y}(\Omega;\mathbb{R}^N)$ only. This goes with the understanding that all the results carry over to $\mathbf{Y}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$ with the appropriate changes.

There is a different view on a Young measure $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$ in the spirit of Berliocchi and Lasry [25]: For

$$
\nu := (\mathcal{L}^d \llcorner \Omega) \otimes \nu_x \qquad \text{and} \qquad \nu^\infty := \lambda_\nu \otimes \nu_x^\infty,
$$

where these generalized products are to be understood in the sense of (2.2), we have that

$$
\nu(\omega \times \mathbb{R}^N) = |\Omega \cap \omega| \qquad \text{and} \qquad \nu^\infty(\omega \times \mathbb{S}^{N-1}) = \lambda_\nu(\omega) \qquad \text{for all } \omega \in \mathcal{B}(\overline{\Omega}).
$$

In the language of Berliocchi and Lasry this means that ν is a Young measure with respect to $\mathcal{L}^d \llcorner \Omega$ and target space $\mathbb{R}^N$ and that ν^∞ is a Young measure with respect to λ_ν and target space $\mathbb{S}^{N-1}$.

Under the **duality product**

$$
\begin{aligned}
\langle\!\langle f, \nu \rangle\!\rangle &:= \int_\Omega \big\langle f(x,\cdot), \nu_x \big\rangle \,\mathrm{d}x + \int_{\overline{\Omega}} \big\langle f^\infty(x,\cdot), \nu_x^\infty \big\rangle \,\mathrm{d}\lambda_\nu(x)\\
&= \int_\Omega \int_{\mathbb{R}^N} f(x,A) \,\mathrm{d}\nu_x(A) \,\mathrm{d}x + \int_{\overline{\Omega}} \int_{\mathbb{S}^{N-1}} f^\infty(x,A) \,\mathrm{d}\nu_x^\infty(A) \,\mathrm{d}\lambda_\nu(x),
\end{aligned}
$$

where $f \in \mathbf{E}(\Omega;\mathbb{R}^N)$ and $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$, the space $\mathbf{Y}(\Omega;\mathbb{R}^N)$ of Young measures can be considered a part of the dual space $\mathbf{E}(\Omega;\mathbb{R}^N)^*$ (the boundedness of $\langle\!\langle \cdot, \nu \rangle\!\rangle$ can be checked by the growth estimate (2.5) and the requirements on Young measures). The inclusion is strict since for instance $f \mapsto \alpha \int_\Omega f(x,0) \,\mathrm{d}x$ lies in $\mathbf{E}(\Omega;\mathbb{R}^N)^* \setminus \mathbf{Y}(\Omega;\mathbb{R}^N)$ whenever $\alpha \neq 1$. We say that a sequence of Young measures $(\nu_j) \subset \mathbf{Y}(\Omega;\mathbb{R}^N)$ **converges weakly*** to $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$, in symbols $\nu_j \overset{*}{\rightharpoondown} \nu$, if for every $f \in \mathbf{E}(\Omega;\mathbb{R}^N)$ it holds that $\langle\!\langle f, \nu_j \rangle\!\rangle \to \langle\!\langle f, \nu \rangle\!\rangle$.

In $\mathbf{Y}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$, we use an analogous duality product, but now between the spaces $\mathbf{E}_c(\mathbb{R}^d;\mathbb{R}^N)$ and $\mathbf{Y}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$. Accordingly, the **(local) weak* convergence** $\nu_j \overset{*}{\rightharpoondown} \nu$ in $\mathbf{Y}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$ is defined relative to $\mathbf{E}_c(\mathbb{R}^d;\mathbb{R}^N)$, i.e. $\nu_j \overset{*}{\rightharpoondown} \nu$ if $\langle\!\langle f, \nu_j \rangle\!\rangle \to \langle\!\langle f, \nu \rangle\!\rangle$ for all $f \in \mathbf{E}_c(\mathbb{R}^d;\mathbb{R}^N)$.

$$\begin{tikzcd}
\mathbf{E}(\Omega;\mathbb{R}^N) \ni f \arrow[r, "S"] \arrow[d] & Sf \in \mathrm{C}(\overline{\Omega \times \mathbb{B}^N}) \arrow[d] \\
\langle\!\langle \cdot, \cdot \rangle\!\rangle & \langle \cdot, \cdot \rangle \\
\mathbf{Y}(\Omega;\mathbb{R}^N) \ni \nu \arrow[u] \arrow[r, "(S^{-1})^*"] & (S^{-1})^*\nu \in \mathbf{M}^+(\overline{\Omega \times \mathbb{B}^N}) \text{ with (2.9)} \arrow[u]
\end{tikzcd}$$

Figure 2.1: Duality relationships.

For every Young measure $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$ we define its **barycenter** as the measure $[\nu] \in \mathbf{M}(\overline{\Omega};\mathbb{R}^N)$ given by

$$[\nu] := \big\langle \mathrm{id}, \nu_x \big\rangle \, \mathcal{L}^d_x \llcorner \Omega + \big\langle \mathrm{id}, \nu_x^\infty \big\rangle \, \lambda_\nu(x),$$

where we understand $\langle \mathrm{id}, \mu \rangle$ as denoting the center of mass (or barycenter) of a probability measure μ. For $\nu \in \mathbf{Y}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$ of course we only get $[\nu] \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$.

2.3.2 Duality and compactness

Based on duality considerations, we will now establish the central compactness result of the theory.

Since $S \colon \mathbf{E}(\Omega;\mathbb{R}^N) \to \mathrm{C}(\overline{\Omega \times \mathbb{B}^N})$ as defined in (2.3) is a continuous isomorphism, so is its dual (or adjoint) operator $S^* \colon \mathbf{M}(\overline{\Omega \times \mathbb{B}^N}) \to \mathbf{E}(\Omega;\mathbb{R}^N)^*$ (in view of the Riesz Representation Theorem, we always identify $\mathrm{C}(\overline{\Omega \times \mathbb{B}^N})^*$ with $\mathbf{M}(\overline{\Omega \times \mathbb{B}^N})$); similarly for $(S^{-1})^* = (S^*)^{-1}$. Therefore, for any $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N) \subset \mathbf{E}(\Omega;\mathbb{R}^N)^*$, we have $(S^{-1})^*\nu = (S^*)^{-1}\nu \in \mathbf{M}(\overline{\Omega \times \mathbb{B}^N})$. By the definition of the dual operator,

$$\big\langle Sf, (S^{-1})^*\nu \big\rangle = \big\langle\!\big\langle f, \nu \big\rangle\!\big\rangle \qquad \text{for all } f \in \mathbf{E}(\Omega;\mathbb{R}^N).$$

We can also write out the definition of $(S^{-1})^*\nu$ explicitly:

$$\big\langle \Phi, (S^{-1})^*\nu \big\rangle = \big\langle\!\big\langle S^{-1}\Phi, \nu \big\rangle\!\big\rangle = \int_\Omega \big\langle S^{-1}\Phi(x, \cdot\,), \nu_x \big\rangle \,\mathrm{d}x + \int_{\overline{\Omega}} \big\langle \Phi(x, \cdot\,), \nu_x^\infty \big\rangle \,\mathrm{d}\lambda_\nu(x)$$

for any $\Phi \in \mathrm{C}(\overline{\Omega \times \mathbb{B}^N})$ (notice that $(S^{-1}\Phi)^\infty|_{\overline{\Omega} \times \mathbb{S}^{N-1}} = \Phi|_{\overline{\Omega} \times \mathbb{S}^{N-1}}$). See Figure 2.1 for a diagram of the duality relationships.

The next lemma characterizes the image of $\mathbf{Y}(\Omega;\mathbb{R}^N)$ in $\mathbf{M}(\overline{\Omega \times \mathbb{B}^N})$ under the map $(S^{-1})^*$. This will later allow us to infer many useful properties of $\mathbf{Y}(\Omega;\mathbb{R}^N)$.

Lemma 2.5. *The set $(S^{-1})^*\mathbf{Y}(\Omega;\mathbb{R}^N) \subset \mathbf{M}(\overline{\Omega \times \mathbb{B}^N})$ consists precisely of the measures $\mu \in \mathbf{M}(\overline{\Omega \times \mathbb{B}^N})$ that are positive and that satisfy*

$$\int_{\overline{\Omega \times \mathbb{B}^N}} \varphi(x)(1 - |A|) \,\mathrm{d}\mu(x, A) = \int_\Omega \varphi(x) \,\mathrm{d}x \qquad \text{for all } \varphi \in \mathrm{C}(\overline{\Omega}). \tag{2.9}$$

Proof. The above explicit representation of $(S^{-1})^*\nu$, $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$, gives for $\Phi = \mathbb{1}$

$$((S^{-1})^*\nu)(\overline{\Omega \times \mathbb{B}^N}) = \big\langle \mathbb{1}, (S^{-1})^*\nu \big\rangle = \int_\Omega \big\langle 1 + |\cdot|, \nu_x \big\rangle \,\mathrm{d}x + \lambda_\nu(\overline{\Omega}),$$

which is finite by the assumptions on ν. Moreover, for $\varphi \in \mathrm{C}(\overline{\Omega})$ we have with $\Psi(x,A) := \varphi(x)(1-|A|)$ (notice $S^{-1}\Psi = \varphi$, $\Psi|_{\overline{\Omega}\times\mathbb{S}^{N-1}} \equiv 0$)

$$\int_{\overline{\Omega\times\mathbb{B}^N}} \varphi(x)(1-|A|) \,\mathrm{d}((S^{-1})^*\nu)(x,A) = \big\langle \Psi, (S^{-1})^*\nu \big\rangle = \int_\Omega \varphi(x) \,\mathrm{d}x,$$

establishing (2.9). The positivity of $(S^{-1})^*\nu$ is clear because whenever $\Phi \geq 0$, then also $\langle \Phi, (S^{-1})^*\nu \rangle = \langle\!\langle S^{-1}\Phi, \nu \rangle\!\rangle \geq 0$.

For the other direction we need to find for every $\mu \in \mathbf{M}^+(\overline{\Omega \times \mathbb{B}^N})$ satisfying (2.9) a Young measure $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$ with $\mu = (S^{-1})^*\nu$ or, equivalently,

$$\big\langle Sf, \mu \big\rangle = \big\langle Sf, (S^{-1})^*\nu \big\rangle = \big\langle\!\big\langle f, \nu \big\rangle\!\big\rangle \qquad \text{for all } f \in \mathbf{E}(\Omega;\mathbb{R}^N).$$

So take $\mu \in \mathbf{M}^+(\overline{\Omega \times \mathbb{B}^N})$ with (2.9). By the Disintegration Theorem 2.1, there exists a measure $\kappa \in \mathbf{M}^+(\overline{\Omega})$ and a weakly* κ-measurable family of probability measures $(\eta_x)_{x\in\overline{\Omega}} \subset \mathbf{M}^1(\overline{\mathbb{B}^N})$ such that $\mu = \kappa \otimes \eta_x$. Let $\kappa = p\mathcal{L}^d + \kappa^s$ be the Lebesgue–Radon–Nikodým decomposition of κ with respect to Lebesgue measure ($p \in \mathrm{L}^1(\Omega)$). From (2.9) with φ approximating characteristic functions of open sets and differentiating the integral with respect to the domain (see for example Chapter 7 in [110]), we get $\langle 1 - |\cdot|, \eta_x \rangle \,\kappa^s = 0$, whence η_x is concentrated in $\mathbb{S}^{N-1}$ for κ^s-a.e. $x \in \overline{\Omega}$.

Therefore, we may decompose ($f \in \mathbf{E}(\Omega;\mathbb{R}^N)$, so $Sf = f^\infty$ on $\overline{\Omega} \times \mathbb{S}^{N-1}$)

$$\begin{aligned}
\big\langle Sf(x,\cdot), \eta_x \big\rangle \,\kappa &= \left(\int_{\overline{\mathbb{B}^N}} Sf(x,\cdot) \,\mathrm{d}\eta_x \right) \kappa \\
&= \left(p(x) \int_{\mathbb{B}^N} Sf(x,\cdot) \,\mathrm{d}\eta_x \right) \mathcal{L}^d \llcorner \Omega \\
&\quad + \left(\eta_x(\mathbb{S}^{N-1}) p(x) \,⨍_{\partial\mathbb{B}^N} f^\infty(x,\cdot) \,\mathrm{d}\eta_x \right) \mathcal{L}^d \llcorner \Omega \\
&\quad + \left(\eta_x(\mathbb{S}^{N-1}) \,⨍_{\partial\mathbb{B}^N} f^\infty(x,\cdot) \,\mathrm{d}\eta_x \right) \kappa^s.
\end{aligned}$$

If we define the measures $\nu_x \in \mathbf{M}^+(\mathbb{R}^N)$, $x \in \Omega$, through

$$\big\langle h, \nu_x \big\rangle := p(x) \int_{\mathbb{B}^N} Sh \,\mathrm{d}\eta_x, \qquad h \in \mathrm{C}_0(\mathbb{R}^N),$$

the measures $\nu_x^\infty \in \mathbf{M}^+(\mathbb{S}^{N-1})$, $x \in \overline{\Omega}$, through

$$\big\langle h^\infty, \nu_x^\infty \big\rangle := ⨍_{\mathbb{S}^{N-1}} h^\infty \,\mathrm{d}\eta_x, \qquad h^\infty \in \mathrm{C}(\mathbb{S}^{N-1}),$$

and finally (see below for finiteness)

$$\lambda_\nu := \eta_x(\mathbb{S}^{N-1})\,\kappa = \eta_x(\mathbb{S}^{N-1})\big[p\,\mathcal{L}^d \llcorner \Omega + \kappa^s\big] \in \mathbf{M}^+(\overline{\Omega}),$$

then the above decomposition becomes

$$\big\langle Sf(x,\bullet),\eta_x\big\rangle\,\kappa = \big\langle f(x,\bullet),\nu_x\big\rangle\,\mathcal{L}^d \llcorner \Omega + \big\langle f^\infty(x,\bullet),\nu_x^\infty\big\rangle\,\lambda_\nu. \tag{2.10}$$

We next prove that ν_x and ν_x^∞ are indeed probability measures. Since ν_x^∞ is defined through averaging, the second claim is trivial. To prove that $\langle \mathbb{1},\nu_x\rangle = 1$, for $\mathcal{L}^d$-almost every $x \in \Omega$, observe that by (2.10) with $f(x,A) := \varphi(x)$ and (2.9)

$$\begin{aligned}\int_\Omega \varphi(x)\big\langle \mathbb{1},\nu_x\big\rangle \,\mathrm{d}x &= \int_\Omega \varphi(x)\big\langle 1-|\bullet|,\eta_x\big\rangle \,\mathrm{d}\kappa(x)\\ &= \int_{\Omega\times\mathbb{B}^N} \varphi(x)(1-|A|)\,\mathrm{d}\mu(x,A) = \int_\Omega \varphi(x)\,\mathrm{d}x\end{aligned}$$

for every $\varphi \in \mathrm{C}(\overline{\Omega})$.

The only thing left to show is that $x \mapsto \langle|\bullet|,\nu_x\rangle \in \mathrm{L}^1(\Omega)$ and $\lambda_\nu(\overline{\Omega}) < \infty$. This follows from (2.10) with $f(x,A) := 1 + |A|$:

$$\int_\Omega \big\langle 1+|\bullet|,\nu_x\big\rangle \,\mathrm{d}x + \lambda_\nu(\overline{\Omega}) = \big\langle \mathbb{1},\mu\big\rangle = \mu(\overline{\Omega\times\mathbb{B}^N}) < \infty,$$

since $Sf(x,A) = \mathbb{1}$. □

Remark 2.6. In the above proof we (more or less) explicitly constructed the dual operator $S^*\colon \mathbf{M}(\overline{\Omega\times\mathbb{B}^N}) \to \mathbf{Y}(\Omega;\mathbb{R}^N)$ as $\mu \mapsto \nu = (\nu_x,\lambda_\nu,\nu_x^\infty)$ whenever μ satisfies the assumptions of the lemma.

Because essentially we are working in a dual space, compactness is now easy to establish. For this, recall that $(S^{-1})^*$ is a weak*-isomorphism and hence all considerations on $\mathbf{Y}(\Omega;\mathbb{R}^N)$ regarding this topology may be investigated by means of the image of $\mathbf{Y}(\Omega;\mathbb{R}^N)$ in $\mathbf{M}(\overline{\Omega\times\mathbb{B}^N})$ under $(S^{-1})^*$.

Corollary 2.7 (Closedness). *The set $\mathbf{Y}(\Omega;\mathbb{R}^N)$ is weakly* closed (considered as a subset of $\mathbf{E}(\Omega;\mathbb{R}^N)^*$).*

Proof. It suffices to realize that $\mathbf{M}^+(\overline{\Omega\times\mathbb{B}^N})$ and condition (2.9) are weakly* closed. The first assertion is trivial and the second one follows since $\varphi(x)(1-|A|)$, $\varphi \in \mathrm{C}(\overline{\Omega})$, is an admissible test function for the weak* topology on $\mathbf{M}(\overline{\Omega\times\mathbb{B}^N}) \cong \mathrm{C}(\overline{\Omega\times\mathbb{B}^N})^*$. □

Corollary 2.8 (Compactness). *Let $(\nu_j) \subset \mathbf{Y}(\Omega;\mathbb{R}^N)$ be a sequence of Young measures such that*

$$\sup_j \big\langle\!\big\langle \mathbb{1}\otimes|\bullet|,\nu_j\big\rangle\!\big\rangle < \infty,$$

or, equivalently,

(i) the functions $x \mapsto \langle |\cdot|, (\nu_j)_x \rangle$ *are uniformly bounded in* $\mathrm{L}^1(\Omega)$ *and*

(ii) the sequence $(\lambda_{\nu_j}(\overline{\Omega}))$ *is uniformly bounded.*

Then, (ν_j) *is weakly* sequentially relatively compact in* $\mathbf{Y}(\Omega;\mathbb{R}^N)$*, i.e. there exists a subsequence (not relabeled) such that* $\nu_j \overset{*}{\rightharpoondown} \nu$ *and* $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$*.*

Proof. By the Banach–Alaoglu–Bourbaki Theorem and the fact that $(S^{-1})^*$ is a weak*-isomorphism between $\mathbf{Y}(\Omega;\mathbb{R}^N)$ and a weakly* closed subset of $\mathbf{M}^+(\overline{\Omega \times \mathbb{B}^N})$, it suffices to show that the measures $(S^{-1})^*\nu_j$ are uniformly bounded as functionals on $\mathrm{C}(\overline{\Omega \times \mathbb{B}^N})$.

Let $\Phi \in \mathrm{C}(\overline{\Omega \times \mathbb{B}^N})$ with $\|\Phi\|_\infty \le 1$. Writing out the definition of S^{-1}, we get

$$\begin{aligned}
\bigl|\bigl\langle \Phi, (S^{-1})^*\nu_j \bigr\rangle\bigr| &= \bigl|\bigl\langle\!\bigl\langle S^{-1}\Phi, \nu_j \bigr\rangle\!\bigr\rangle\bigr| \\
&\le \int_\Omega \int_{\mathbb{R}^N} (1+|A|) \Bigl|\Phi\Bigl(x, \frac{A}{1+|A|}\Bigr)\Bigr| \;\mathrm{d}(\nu_j)_x(A)\;\mathrm{d}x \\
&\qquad + \int_{\overline{\Omega}} \int_{\mathbb{S}^{N-1}} |\Phi(x,A)| \;\mathrm{d}(\nu_j)_x^\infty(A)\;\mathrm{d}\lambda_\nu(x) \\
&\le \sup\nolimits_j \biggl(\int_\Omega \bigl\langle 1+|\cdot|, (\nu_j)_x \bigr\rangle \;\mathrm{d}x + \lambda_{\nu_j}(\overline{\Omega}) \biggr) < \infty.
\end{aligned}$$

Now the Banach–Alaoglu–Bourbaki Theorem applies and by Corollary 2.7 the limit is again representable as a Young measure. The *sequential* relative compactness follows from the separability of $\mathbf{E}(\Omega;\mathbb{R}^N)$ in a standard way, we omit the details here. □

By a localization technique, we analogously get the following compactness result in $\mathbf{Y}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$:

Corollary 2.9 (Compactness in $\mathbf{Y}_{\mathrm{loc}}$). *Let* $(\nu_j) \subset \mathbf{Y}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$ *be a sequence of Young measures satisfying*

$$\sup\nolimits_j \bigl\langle\!\bigl\langle \varphi \otimes |\cdot|, \nu_j \bigr\rangle\!\bigr\rangle < \infty \qquad \text{for all } \varphi \in \mathrm{C}_c(\mathbb{R}^d).$$

Then, there exists a subsequence (not relabeled) of (ν_j) *with* $\nu_j \overset{*}{\rightharpoondown} \nu$ *in* $\mathbf{Y}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$*.*

2.3.3 Generation

To every Radon measure $\gamma \in \mathbf{M}(\overline{\Omega};\mathbb{R}^N)$ with Lebesgue–Radon–Nikodým decomposition $\gamma = a\,\mathcal{L}^d \llcorner \Omega + \gamma^s$, we associate an **elementary Young measure** $\varepsilon_\gamma \in \mathbf{Y}(\Omega;\mathbb{R}^N)$ via

$$(\varepsilon_\gamma)_x := \delta_{a(x)} \quad \mathcal{L}^d\text{-a.e.}, \qquad \lambda_{\varepsilon_\gamma} := |\gamma^s|, \qquad (\varepsilon_\gamma)_x^\infty := \delta_{p(x)} \quad |\gamma^s|\text{-a.e.},$$

where

$$p := \frac{\mathrm{d}\gamma^s}{\mathrm{d}|\gamma^s|} \in \mathrm{L}^1(\Omega, |\gamma^s|; \mathbb{S}^{N-1}).$$

Let $(\gamma_j) \subset \mathbf{M}(\overline{\Omega};\mathbb{R}^N)$ be a sequence of such measures with $\sup_j |\gamma_j|(\overline{\Omega}) < \infty$. Then, we say that the γ_j **generate** the Young measure $\nu = (\nu_x, \lambda_\nu, \nu_x^\infty) \in \mathbf{Y}(\Omega;\mathbb{R}^N)$, in symbols $\gamma_j \overset{\mathbf{Y}}{\to} \nu$, if for all $f \in \mathbf{E}(\Omega;\mathbb{R}^N)$ we have

$$\begin{aligned} f\biggl(x, \frac{\mathrm{d}\gamma_j}{\mathrm{d}\mathcal{L}^d}(x)\biggr)\, \mathcal{L}_x^d \llcorner \Omega \;+\; f^\infty\biggl(x, \frac{\mathrm{d}\gamma_j^s}{\mathrm{d}|\gamma_j^s|}(x)\biggr)\, |\gamma_j^s|(x) \\ \overset{*}{\rightharpoondown}\;\; \bigl\langle f(x,\,\boldsymbol{\cdot}\,), \nu_x \bigr\rangle\, \mathcal{L}_x^d \llcorner \Omega \;+\; \bigl\langle f^\infty(x,\,\boldsymbol{\cdot}\,), \nu_x^\infty \bigr\rangle\, \lambda_\nu(x) \qquad \text{in } \mathbf{M}(\overline{\Omega}). \end{aligned}$$

In fact, with the above notation for elementary Young measures, we could also express generation equivalently as

$$\varepsilon_{\gamma_j} \overset{*}{\rightharpoondown} \nu \quad \text{in } \mathbf{Y}(\Omega;\mathbb{R}^N), \qquad \text{i.e.} \qquad \bigl\langle\!\bigl\langle f, \varepsilon_{\gamma_j} \bigr\rangle\!\bigr\rangle \to \bigl\langle\!\bigl\langle f, \nu \bigr\rangle\!\bigr\rangle \quad \text{for all } f \in \mathbf{E}(\Omega;\mathbb{R}^N).$$

We will, however, prefer the notation $\gamma_j \overset{\mathbf{Y}}{\to} \nu$ in order to avoid confusion with weak* convergence in $\mathbf{M}(\overline{\Omega};\mathbb{R}^N)$. If a Young measure $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$ is generated by a sequence of functions $(u_j) \subset \mathrm{L}^1(\Omega;\mathbb{R}^N)$ in the sense that $u_j\, \mathcal{L}^d \llcorner \Omega \overset{\mathbf{Y}}{\to} \nu$, we usually simply write $u_j \overset{\mathbf{Y}}{\to} \nu$.

If $\mu_j \overset{*}{\rightharpoondown} \mu$ in $\mathbf{M}(\overline{\Omega};\mathbb{R}^N)$ (note the closure) and $\mu_j \overset{\mathbf{Y}}{\to} \nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$, then the barycenter $[\nu]$ of ν equals μ. Let us further explicitly note that the concentration measure λ_ν equals the weak* limit of the total variation measures $|\gamma_j|$ minus $\langle |\,\boldsymbol{\cdot}\,|, \nu_x \rangle\, \mathcal{L}^d \llcorner \Omega$.

Do all (bounded) sequences of Radon measures generate some Young measure up to a subsequence? This is indeed the case and often called the Fundamental Theorem for Young measures. In this context it was first established in [43] and [3]. For us it is an easy corollary to the general compactness result, Corollary 2.8.

Theorem 2.10 (Fundamental Theorem). *Let $(\gamma_j) \subset \mathbf{M}(\overline{\Omega};\mathbb{R}^N)$ be a sequence of Radon measures that is uniformly bounded in the total variation norm (i.e. $\sup_j |\gamma_j|(\overline{\Omega}) < \infty$). Then, there exists a subsequence (not relabeled) such that $\gamma_j \overset{\mathbf{Y}}{\to} \nu$ for some Young measure $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$.*

Proof. Just apply the compactness result, Corollary 2.8, using $\nu_j := \varepsilon_{\gamma_j}$, the elementary Young measure associated with γ_j. The hypotheses of Corollary 2.8 follow from the boundedness of the sequence (γ_j) in the total variation norm because $\langle\!\langle \mathbb{1} \otimes |\,\boldsymbol{\cdot}\,|, \varepsilon_{\gamma_j} \rangle\!\rangle = |\gamma_j|(\overline{\Omega})$ is uniformly bounded. □

Remark 2.11. The assumptions of the Fundamental Theorem exactly correspond to the hypotheses needed to pass to a weakly* converging subsequence of measures, but the Young measure limit retains much more information, see the examples below and in Section 2.6.3.

By means of Reshetnyak's Continuity Theorem 2.3, we immediately get:

Proposition 2.12. *If $\gamma_j \to \gamma$ $\langle\,\boldsymbol{\cdot}\,\rangle$-strictly in $\mathbf{M}(\overline{\Omega};\mathbb{R}^N)$, then $\gamma_j \overset{\mathbf{Y}}{\to} \varepsilon_\gamma$.*

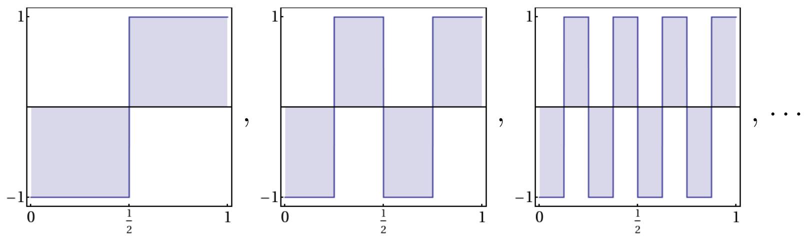

Figure 2.2: Young measure generation in the pure oscillation case.

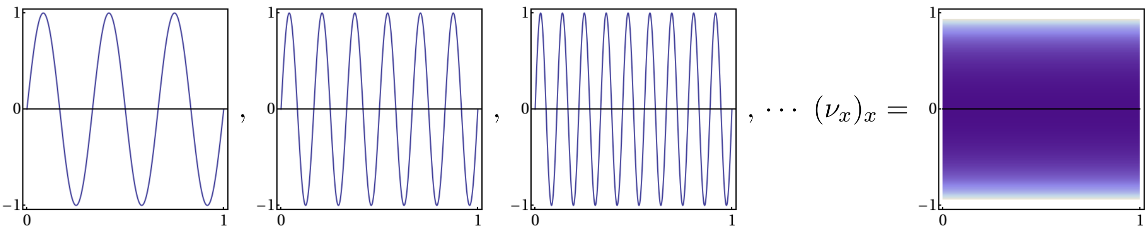

Figure 2.3: A second oscillation example.

Similar statements as before also hold for generation in $\mathbf{Y}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$ by (locally) weakly* relatively compact sequences of measures in $\mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$. If $\varepsilon_{\mu_j} \overset{*}{\rightharpoonup} \nu$ in $\mathbf{Y}_{\mathrm{loc}}(\Omega;\mathbb{R}^N)$, then we again write $\mu_j \overset{\mathbf{Y}}{\to} \nu$, the ambient space being clear from the context.

We end this section by exhibiting three simple examples for the generation of Young measures.

Example 2.13. In $\Omega := (0,1)$ define $u := \mathbb{1}_{(1/2,1)} - \mathbb{1}_{(0,1/2)}$ 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 2.2) generate the homogeneous Young measure $\nu \in \mathbf{GY}((0,1);\mathbb{R})$ with

$$\nu_x = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1} \quad \text{a.e.}, \qquad \lambda_\nu = 0.$$

Example 2.14. Take $\Omega := (0,1)$ and let $u_j(x) = \sin(2\pi jx)$ for $j \in \mathbb{N}$ (see Figure 2.3). The sequence (u_j) generates the homogeneous Young measure $\nu \in \mathbf{GY}((0,1);\mathbb{R})$ with

$$\nu_x = \frac{1}{\pi\sqrt{1-y^2}}\,\mathcal{L}^1_y \llcorner (-1,1) \quad \text{a.e.}, \qquad \lambda_\nu = 0.$$

Example 2.15. Take again $\Omega := (0,1)$ and define (see Figure 2.4)

$$u_j := 2^j \mathbb{1}_{\left(\frac{1}{2}-2^{-j-1},\frac{1}{2}+2^{-j-1}\right)}, \qquad j \in \mathbb{N}.$$

Then, $u_j \overset{\mathbf{Y}}{\to} \nu \in \mathbf{GY}((0,1);\mathbb{R})$ with

$$\nu_x = \delta_0 \quad \text{a.e.}, \qquad \lambda_\nu = \delta_{1/2}, \qquad \nu^\infty_{1/2} = \delta_{+1}.$$

More elaborate examples can be found in Section 2.6.3.

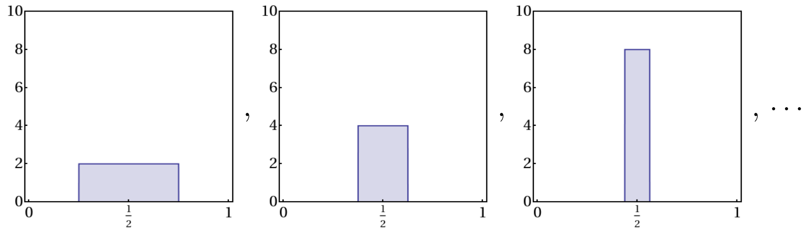

Figure 2.4: A pure concentration example.

2.4 Density and extended representation

The next result assures that we do not have to test weak* convergence in $\mathbf{Y}(\Omega;\mathbb{R}^N)$ with all $f \in \mathbf{E}(\Omega;\mathbb{R}^N)$, but only with a countable subset of tensor product functions.

Lemma 2.16. *There exists a countable set of functions* $\{f_k\} = \{\varphi_k \otimes h_k \ : \ k \in \mathbb{N}\} \subset \mathbf{E}(\Omega;\mathbb{R}^N)$*, where* $\varphi_k \in \mathrm{C}(\overline{\Omega})$ *and* $h_k \in \mathrm{C}(\mathbb{R}^N)$*, such that the knowledge of* $\langle\!\langle f_k, \nu \rangle\!\rangle$ *completely determines* $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$*. Moreover, all the* h_k *can be chosen to be Lipschitz continuous.*

More precisely, the lemma states that if for two Young measures $\mu, \nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$ it holds that $\langle\!\langle f_k, \mu \rangle\!\rangle = \langle\!\langle f_k, \nu \rangle\!\rangle$ for all $k \in \mathbb{N}$, then $\mu = \nu$, i.e. $\mu_x = \nu_x$ for $\mathcal{L}^d$-a.e. $x \in \Omega$, $\lambda_\mu = \lambda_\nu$, and $\mu_x^\infty = \nu_x^\infty$ for λ_μ-a.e. $x \in \overline{\Omega}$ (equivalently, λ_ν-a.e. $x \in \overline{\Omega}$). An immediate consequence is that to uniquely identify the limit in the weak* convergence $\nu_j \overset{*}{\rightharpoonup} \nu$ in $\mathbf{Y}(\Omega;\mathbb{R}^N)$, it suffices to test with the collection $\{f_k\}$, we say that the f_k "determine" the Young measure convergence.

Proof. Take countable sets $\mathcal{A} \subset \mathrm{C}(\overline{\Omega})$, $\mathcal{B} \subset \mathrm{C}_c^1(\mathbb{R}^N)$, which are dense in $\mathrm{C}(\overline{\Omega})$ and $\mathrm{C}_0(\mathbb{R}^N)$, respectively (dense with respect to the $\|\bullet\|_\infty$-norm), and also take $\mathcal{C} \subset \mathrm{C}_c^1(\mathbb{R}^N)$ such that the restrictions of the functions in $\mathcal{C}$ to $\mathbb{S}^{N-1}$ are dense in $\mathrm{C}(\mathbb{S}^{N-1})$; assume furthermore that $\mathcal{C}$ contains a function h^∞ with $h^\infty|_{\mathbb{S}^{N-1}} \equiv 1$. Define $(Gh^\infty)(A) := |A|h^\infty(A/|A|)$ $(A \in \mathbb{R}^N)$ for $h^\infty \in \mathcal{C}$. Then let

$$\{\varphi_k \otimes h_k\}_k := (\mathcal{A} \otimes \mathcal{B}) \cup (\mathcal{A} \otimes G(\mathcal{C})) \subset \mathbf{E}(\Omega;\mathbb{R}^N)$$

(the tensor products on the right hand side are understood to act pointwise and not to include some form of tensor product completion). By standard results of measure theory, the knowledge of

$$\langle\!\langle \varphi \otimes h, \nu \rangle\!\rangle = \int_\Omega \varphi(x) \langle h, \nu_x \rangle \,\mathrm{d}x \qquad \text{for all } \varphi \in \mathrm{C}(\overline{\Omega}),\ h \in \mathrm{C}_0(\mathbb{R}^N)$$

determines first the $\mathrm{L}^1(\Omega)$-function $x \mapsto \langle h, \nu_x \rangle$ and then the measures ν_x up to an $\mathcal{L}^d$-negligible set since $\mathbf{M}(\mathbb{R}^N) \cong \mathrm{C}_0(\mathbb{R}^N)^*$. By density, this identification also holds by only considering $\varphi \otimes h \in \mathcal{A} \otimes \mathcal{B}$. Now testing with $\mathcal{A} \otimes (G\mathbb{1})$ gives for all $\varphi \in \mathrm{C}(\overline{\Omega})$

$$\langle\!\langle \varphi \otimes G\mathbb{1}, \nu \rangle\!\rangle = \int_\Omega \varphi(x) \langle G\mathbb{1}, \nu_x \rangle \,\mathrm{d}x + \int_{\overline{\Omega}} \varphi(x) \langle \mathbb{1}, \nu_x^\infty \rangle \,\mathrm{d}\lambda_\nu(x).$$

Since the first integral is already identified and the second integral reduces to $\int_{\overline{\Omega}} \varphi \, \mathrm{d}\lambda_\nu$, similarly to before $\mathbf{M}(\overline{\Omega}) \cong \mathrm{C}(\overline{\Omega})^*$ implies that λ_ν is also uniquely determined. Finally, the identification of ν_x^∞ is analogous to that of ν_x, noticing that $(Gh^\infty)^\infty = h^\infty$ on $\mathbb{S}^{N-1}$ for all $h^\infty \in \mathcal{C}$ (of course, identification is possible only up to a λ_ν-negligible set).

Finally, we remark that each h_k, $k \in \mathbb{N}$, is Lipschitz continuous. If $h_k \in \mathcal{B}$, this is obviously true. If $h_k = G(h^\infty)$ with $h^\infty \in \mathcal{C}$, then there exists a constant $c > 0$ such that

$$\left| h^\infty\left(\frac{A}{|A|}\right) - h^\infty\left(\frac{B}{|B|}\right) \right| \le c \left| \frac{A}{|A|} - \frac{B}{|B|} \right| \qquad \text{for all } A, B \in \mathbb{R}^N \setminus \{0\}.$$

Hence,

$$\begin{aligned}
\big|h_k(A) - h_k(B)\big| &\le \left| h^\infty\left(\frac{A}{|A|}\right) - h^\infty\left(\frac{B}{|B|}\right) \right| |B| + \left| h^\infty\left(\frac{A}{|A|}\right) \right| |A - B| \\
&\le c \left| \frac{A}{|A|}|B| - B \right| + \Big(\max_{\mathbb{S}^{N-1}} |h^\infty| \Big) |A - B| \\
&\le \big(2c + \max_{\mathbb{S}^{N-1}} |h^\infty|\big) |A - B|,
\end{aligned}$$

so h_k is Lipschitz continuous. □

As an immediate corollary, we get that the weak* topology on $\mathbf{Y}(\Omega;\mathbb{R}^N)$ is metrizable on bounded sets.

In the first instance, $\nu_j \overset{*}{\rightharpoonup} \nu$ in $\mathbf{Y}(\Omega;\mathbb{R}^N)$ only implies the convergence $\langle\langle f, \nu_j \rangle\rangle \to \langle\langle f, \nu \rangle\rangle$ for $f \in \mathbf{E}(\Omega;\mathbb{R}^N)$, but the next result shows that we have convergence for the class of integrands $\mathbf{R}(\Omega;\mathbb{R}^N)$ as well. We can even relax the continuity of the recession function a little further if the integrand is compatible with the Young measure ν.

Proposition 2.17. *Let $\nu_j \overset{*}{\rightharpoonup} \nu$ in $\mathbf{Y}(\Omega;\mathbb{R}^N)$. Then, $\langle\langle f, \nu_j \rangle\rangle \to \langle\langle f, \nu \rangle\rangle$ holds provided one of the following conditions is satisfied:*

(i) $f \in \mathbf{R}(\Omega;\mathbb{R}^N)$,

(ii) $f(x, A) = \mathbb{1}_U(x) g(x, A)$, *where* $g \in \mathbf{E}(\Omega;\mathbb{R}^N)$ *and* $U \in \mathcal{B}(\overline{\Omega})$ *with* $(\mathcal{L}^d + \lambda_\nu)(\partial U) = 0$,

(iii) $f \colon \overline{\Omega} \times \mathbb{R}^N \to \mathbb{R}$ *is a Carathéodory integrand possessing a recession function* $f^\infty \colon \overline{\Omega} \times \mathbb{R}^N \to \mathbb{R}$ *in the sense of* (2.6) *for* $(x, A) \in (\overline{\Omega} \setminus N) \times \mathbb{R}^N$ *and* f^∞ *is jointly continuous on* $(\overline{\Omega} \setminus N) \times \mathbb{R}^N$, *where* $N \subset \overline{\Omega}$ *is a Borel set with* $(\mathcal{L}^d + \lambda_\nu)(N) = 0$.

Similar statements hold for $\mathbf{Y}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$.

Before we prove the representation result, note that for each integrand $f \in \mathbf{R}(\Omega;\mathbb{R}^N)$ and $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$ the expression $\langle\langle f, \nu \rangle\rangle$ is well-defined. For the concentration part this is clear since f^∞ is continuous. For the oscillation part, the assertion follows first under the additional hypothesis that f is a bounded Borel function by uniform approximation with simple functions and using the weak* measurability of ν_x. Then approximate a general Borel function f pointwise from below by bounded Borel functions. Finally, use the fact that a function is Lebesgue-measurable if and only if it is equal to a Borel function outside an $\mathcal{L}^d$-negligible set. The cases in the proposition, in which f^∞ is not continuous (possibly not even defined) still do not cause problems since we always assumed that the exceptional set is Borel-measurable and $(\mathcal{L}^d + \lambda_\nu)$-negligible.

Proof. We first show (iii), which immediately implies (i).

Step 1. Assume first that $f(x, \bullet)$ has uniformly bounded support, say $\operatorname{supp} f(x, \bullet) \subset B(0, R)$ for all $x \in \overline{\Omega}$ and some $R > 0$. In this case we only need to show

$$\int_\Omega \bigl\langle f(x, \bullet), (\nu_j)_x \bigr\rangle \,\mathrm{d}x \to \int_\Omega \bigl\langle f(x, \bullet), \nu_x \bigr\rangle \,\mathrm{d}x.$$

Take a countable dense family $\{\psi_k\}_k \subset \mathrm{C}_0(B(0, R/(1+R)))$ and define for fixed $\varepsilon > 0$

$$E_k := \bigl\{\, x \in \overline{\Omega} \;:\; \|Sf(x, \bullet) - \psi_k\|_\infty \le \varepsilon \,\bigr\}, \qquad k \in \mathbb{N}.$$

The sets E_k are measurable since by continuity of $f(x, \bullet)$ they can be written as

$$E_k = \bigcap_{q \in \mathbb{Q}^d} \bigl\{\, x \in \overline{\Omega} \;:\; |Sf(x, q) - \psi_k(q)| \le \varepsilon \,\bigr\}.$$

For every $x \in \overline{\Omega}$, $Sf(x, \bullet) \in \mathrm{C}_0(B(0, R/(1+R)))$ and hence $x \in E_k$ for some $k \in \mathbb{N}$. The sets $F_k := E_k \setminus \bigcup_{i<k} F_i$ form a measurable partition of $\overline{\Omega}$ and for

$$g_\varepsilon(x, A) = \sum_{k \in \mathbb{N}} \mathbb{1}_{F_k}(x) S^{-1}\psi_k(A), \qquad x \in \overline{\Omega},\ A \in \mathbb{R}^N,$$

we have $\|Sf - Sg_\varepsilon\|_\infty \le \varepsilon$, whereby $\|g_\varepsilon\|_\infty \le (1 + R)(\|Sf\|_\infty + \varepsilon)$.

Next, observe that from the weak* convergence $\nu_j \overset{*}{\rightharpoondown} \nu$ in $\mathbf{Y}(\Omega; \mathbb{R}^N)$ we can infer

$$\int_{F_k} \bigl\langle S^{-1}\psi_k, (\nu_j)_x \bigr\rangle \,\mathrm{d}x \to \int_{F_k} \bigl\langle S^{-1}\psi_k, \nu_x \bigr\rangle \,\mathrm{d}x$$

for each fixed $k \in \mathbb{N}$. This is so because $S^{-1}\psi_k$ is L^∞-bounded (see above) and because we can L^1-approximate $\mathbb{1}_{F_k}$ by smooth functions.

Using the Dominated Convergence Theorem (for sums, observe the uniform bound $(1 + R)(\|Sf\|_\infty + \varepsilon)$ on all the $S^{-1}\psi_k$), we get

$$\begin{aligned}
\lim_{j \to \infty} \int_\Omega \bigl\langle g_\varepsilon(x, \bullet), (\nu_j)_x \bigr\rangle \,\mathrm{d}x &= \lim_{j \to \infty} \sum_{k \in \mathbb{N}} \int_{F_k} \bigl\langle S^{-1}\psi_k, (\nu_j)_x \bigr\rangle \,\mathrm{d}x \\
&= \sum_{k \in \mathbb{N}} \lim_{j \to \infty} \int_{F_k} \bigl\langle S^{-1}\psi_k, (\nu_j)_x \bigr\rangle \,\mathrm{d}x = \sum_{k \in \mathbb{N}} \int_{F_k} \bigl\langle S^{-1}\psi_k, \nu_x \bigr\rangle \,\mathrm{d}x \\
&= \int_\Omega \bigl\langle g_\varepsilon(x, \bullet), \nu_x \bigr\rangle \,\mathrm{d}x.
\end{aligned}$$

From $\|Sf - Sg_\varepsilon\|_\infty \le \varepsilon$ we further get that uniformly for all $j \in \mathbb{N}$,

$$\left| \int_\Omega \bigl\langle g_\varepsilon(x, \bullet), (\nu_j)_x \bigr\rangle \,\mathrm{d}x - \int_\Omega \bigl\langle f(x, \bullet), (\nu_j)_x \bigr\rangle \,\mathrm{d}x \right| \le \varepsilon \int_\Omega \bigl\langle 1 + |\bullet|, (\nu_j)_x \bigr\rangle \,\mathrm{d}x \le \varepsilon C$$

for some j-independent constant $C > 0$ (by the Young measure convergence); an analogous estimate holds with ν in place of ν_j. This gives

$$\left| \lim_{j\to\infty} \langle\!\langle f, \nu_j \rangle\!\rangle - \langle\!\langle f, \nu \rangle\!\rangle \right| \le \left| \lim_{j\to\infty} \langle\!\langle g_\varepsilon, \nu_j \rangle\!\rangle - \langle\!\langle g_\varepsilon, \nu \rangle\!\rangle \right| + 2\varepsilon C,$$

and we may conclude the proof in the case where f has uniformly bounded support.

Step 2. We now prove the claim if $f^\infty \equiv 0$. For $\varepsilon > 0$ take $r \in (0,1)$ so large that $|Sf(x,A)| \le \varepsilon$ whenever $x \in \overline{\Omega}$ and $A \in \mathbb{B}^N$ with $|A| \ge r$. If such an r did not exist, we could find sequences $(x_n)_n \subset \overline{\Omega}$, $(A_n)_n \subset \mathbb{B}^N$ with $|A_n| \ge 1 - n^{-1}$ and $Sf(x_n, A_n) \ge \varepsilon$. By compactness, $x_n \to x \in \overline{\Omega}$, $A_n \to A \in \mathbb{S}^{N-1}$ and we would get a contradiction to $f^\infty \equiv 0$. So, if we pick $\psi \in \mathrm{C}_c(\mathbb{R}^N)$ with $0 \le \psi \le 1$ and $\psi \equiv 1$ on $B(0, r/(1-r))$, we get by the previous step that

$$\langle\!\langle f\psi, \nu_j \rangle\!\rangle \to \langle\!\langle f\psi, \nu \rangle\!\rangle.$$

On the other hand, for all $j \in \mathbb{N}$, we have

$$|\langle\!\langle f\psi - f, \nu_j \rangle\!\rangle| \le \varepsilon \left[\int_\Omega \bigl\langle 1 + |\,\boldsymbol{\cdot}\,|, (\nu_j)_x \bigr\rangle \,\mathrm{d}x + \lambda_{\nu_j}(\overline{\Omega}) \right] \le \varepsilon C$$

for some constant $C > 0$. This also holds for ν in place of ν_j. Hence, the conclusion for $f^\infty \equiv 0$ follows.

Step 3. If f^∞ is not identically zero, split f via

$$f = h + f^\infty, \qquad \text{where} \qquad h := f - f^\infty.$$

Obviously, $h^\infty \equiv 0$. Hence the last step applies and yields $\langle\!\langle h, \nu_j \rangle\!\rangle \to \langle\!\langle h, \nu \rangle\!\rangle$.

The measures $\mu_j := (S^{-1})^* \nu_j \in \mathbf{M}(\overline{\Omega \times \mathbb{B}^N})$ act on $\Phi \in \mathrm{C}(\overline{\Omega \times \mathbb{B}^N})$ via

$$\bigl\langle \Phi, \mu_j \bigr\rangle := \int_\Omega \bigl\langle S^{-1}\Phi(x, \boldsymbol{\cdot}), (\nu_j)_x \bigr\rangle \,\mathrm{d}x + \int_{\overline{\Omega}} \bigl\langle \Phi(x, \boldsymbol{\cdot}), (\nu_j)_x^\infty \bigr\rangle \,\mathrm{d}\lambda_{\nu_j}(x),$$

similarly for $\mu := (S^{-1})^*\nu$. Since $(S^{-1})^*$ is a weak*-isomorphism between $\mathbf{E}(\Omega; \mathbb{R}^N)^*$ and $\mathbf{M}(\overline{\Omega \times \mathbb{B}^N})$, we have $\mu_j \overset{*}{\rightharpoonup} \mu$ in $\mathbf{M}(\overline{\Omega \times \mathbb{B}^N})$.

Let $\gamma_j := \langle 1 + |\,\boldsymbol{\cdot}\,|, (\nu_j)_x \rangle \, \mathcal{L}^d \llcorner \Omega + \lambda_{\nu_j}$, which converges weakly* to $\langle 1 + |\,\boldsymbol{\cdot}\,|, \nu_x \rangle \, \mathcal{L}^d \llcorner \Omega + \lambda_\nu$. Taking Φ with $\|\Phi\|_\infty \le 1$, we observe that $|\mu_j|(U \times \overline{\mathbb{B}^N}) \le \gamma_j(U)$ for all Borel sets $U \subset \overline{\Omega}$. Hence, if we denote by Λ the weak* limit of the $|\mu_j|$, it follows that $\Lambda(U \times \overline{\mathbb{B}^N}) = 0$ if $(\mathcal{L}^d + \lambda_\nu)(U) = 0$ (the measures $\varphi \mapsto \langle \varphi, \gamma_j \rangle - \langle \varphi \otimes \mathbb{1}, |\mu_j| \rangle$, $\varphi \in \mathrm{C}(\overline{\Omega})$, are positive and hence stay positive in the limit).

By standard results in measure theory, $\langle \Psi, \mu_j \rangle \to \langle \Psi, \mu \rangle$ holds for any bounded Borel function Ψ on $\overline{\Omega \times \mathbb{B}^N}$ which has a Λ-negligible set of discontinuity points (see for example Proposition 1.62 (b) in [10]). In particular, we may take $\Psi := Sf^\infty$, which is continuous outside the set $N \times \overline{\mathbb{B}^N}$ and $(\mathcal{L}^d + \lambda_\nu)(N) = 0$. Hence,

$$\langle\!\langle f^\infty, \nu_j \rangle\!\rangle = \bigl\langle Sf^\infty, \mu_j \bigr\rangle \to \bigl\langle Sf^\infty, \mu \bigr\rangle = \langle\!\langle f^\infty, \nu \rangle\!\rangle, \tag{2.11}$$

which finally yields

$$\langle\!\langle f, \nu_j \rangle\!\rangle = \langle\!\langle h, \nu_j \rangle\!\rangle + \langle\!\langle f^\infty, \nu_j \rangle\!\rangle \to \langle\!\langle h, \nu \rangle\!\rangle + \langle\!\langle f^\infty, \nu \rangle\!\rangle = \langle\!\langle f, \nu \rangle\!\rangle$$

and we have proved (iii).

For (ii), similarly to the third step in the proof of (iii), we let $\Psi(x, A) := Sf(x, A) = \mathbb{1}_U(x) Sg(x, A)$ and conclude as in (2.11). $\square$

Counterexample 2.18. It is easy to see that the assumption on the continuity of f^∞ is necessary for the result to hold: Take $\Omega := (-1, 1)$, $f(x, A) := \mathbb{1}_{(0,1)}(x)|A|$, whence also $f^\infty(x, A) = \mathbb{1}_{(0,1)}(x)$. Further, let $\nu_j \in \mathbf{Y}(\Omega; \mathbb{R}^N)$ be defined through

$$(\nu_j)_x := \delta_0 \quad \text{a.e.}, \qquad \lambda_{\nu_j} := \delta_{1/j}, \qquad (\nu_j)^\infty_{1/j} := \delta_{+1}.$$

Then it is easily seen that $\nu_j \overset{*}{\rightharpoonup} \nu$ with

$$\nu_x = \delta_0 \quad \text{a.e.}, \qquad \lambda_\nu = \delta_0, \qquad \nu^\infty_0 = \delta_{+1},$$

but

$$\lim_{j \to \infty} \langle\!\langle f, \nu_j \rangle\!\rangle = \lim_{j \to \infty} \int_{(0,1)} \mathbb{1}_{(0,1)}(x) \, \mathrm{d}\delta_{1/j}(x) = 1 \neq 0 = \langle\!\langle f, \nu \rangle\!\rangle.$$

In accordance with (ii) of the preceding proposition, the discontinuity set $\{0, 1\}$ is not negligible with respect to $\mathcal{L}^d + \delta_0$.

We close this section with a counterexample showing that in the definition of the recession function f^∞, cf. (2.6), it is necessary to consider the limit over all $x' \to x$ and not merely for x fixed.

Counterexample 2.19. Let $\Omega := (-1, 1)$ and define

$$f(x, A) := \begin{cases} |A| & \text{if } x \neq 0 \text{ and } |A| \leq |x|^{-1}, \\ 0 & \text{otherwise.} \end{cases}, \qquad x \in [-1, 1], \, A \in \mathbb{R}.$$

Then, the sequence of derivatives

$$u_j'(x) := j^2 \mathbb{1}_{\left(-\frac{1}{j^2}, \frac{1}{j^2}\right)}(x), \qquad x \in (-1, 1),$$

satisfies $\|u_j'\|_{\mathrm{L}^1(-1,1)} = 2$ for all $j \in \mathbb{N}$ and generates the Young measure $\nu \in \mathbf{Y}((-1, 1); \mathbb{R})$ with

$$\nu_x = \delta_0 \quad \text{a.e.}, \qquad \lambda_\nu = 2\delta_0, \qquad \nu^\infty_0 = \delta_{+1}.$$

Now, if we had defined

$$f^\infty(x, A) := \lim_{\substack{A' \to A \\ t \to \infty}} \frac{f(x, tA')}{t}, \qquad x \in \Omega, \, A \in \mathbb{R}^N,$$

instead of (2.6) (without $x' \to x$), for our f above we would have $f^\infty \equiv 0$. But clearly,

$$\langle\!\langle f, \varepsilon_{u_j'} \rangle\!\rangle = \langle\!\langle |\cdot|, \varepsilon_{u_j'} \rangle\!\rangle = 2 \not\to 0 = \langle\!\langle f, \nu \rangle\!\rangle.$$

Notice that with the original definition (2.6) for f^∞, $f^\infty(0, \cdot)$ is not well-defined and since $\lambda_\nu = 2\delta_0$, this must be excluded.

2.5 Young measures and convergence of functions

For a Young measure $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$ that is generated by a sequence of functions $(u_j) \subset \mathrm{L}^1(\Omega;\mathbb{R}^N)$, we can examine how convergence and compactness properties of the sequence (u_j) are reflected in the generated Young measure ν. The first result in this direction shows that the concentration measure λ_ν of a Young measure $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$ indeed quantifies all the concentration. This proposition and its proof are adapted from Theorem 2.9 (ii) in [3].

Proposition 2.20. *Let $(u_j) \subset \mathrm{L}^1(\Omega;\mathbb{R}^N)$ with $u_j \overset{\mathbf{Y}}{\to} \nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$. Then, the sequence (u_j) is equiintegrable if and only if $\lambda_\nu = 0$. Moreover, for $f \in \mathbf{E}(\Omega;\mathbb{R}^N)$ let $F_j(x) := f(x, u_j(x))$ $(x \in \Omega)$. Then, the sequence (F_j) is equiintegrable if and only if $\langle |f^\infty(x, \cdot)|, \nu_x^\infty \rangle = 0$ for λ_ν-almost every $x \in \overline{\Omega}$.*

Proof. Clearly, the second statement implies the first (take $f(x, A) := |A|$). For $f \in \mathbf{E}(\Omega;\mathbb{R}^N)$ and $R > 0$ define

$$\eta_R := \limsup_{j\to\infty} \int_{\{|F_j|\geq R\}} |F_j| \,\mathrm{d}x.$$

The Dunford–Pettis Theorem says that the sequence (F_j) is equiintegrable if and only if $\eta_\infty := \lim_{R\uparrow\infty} \eta_R = 0$. We will show that

$$\eta_\infty = \bigl\langle |f^\infty(x, \cdot)|, \nu_x^\infty \bigr\rangle \,\lambda_\nu, \tag{2.12}$$

from which the claim follows.

To see (2.12), first let

$$h(t) := \begin{cases} 0 & \text{if } 0 < t < \frac{1}{2}, \\ 2t - 1 & \text{if } \frac{1}{2} \leq t \leq 1, \\ t & \text{if } t > 1, \end{cases} \qquad h_R(t) := Rh\biggl(\frac{t}{R}\biggr), \qquad t \geq 0,$$

and observe $h_{2R}(t) \leq t\mathbb{1}_{[R,\infty)}(t) \leq h_R(t)$ for all $t \geq 0$, as well as $(h_R \circ |f|)^\infty = |f^\infty|$. Thus,

$$\int_\Omega h_{2R}(|F_j(x)|) \,\mathrm{d}x \leq \eta_R \leq \int_\Omega h_R(|F_j(x)|) \,\mathrm{d}x$$

and passing to the Young measure limit as $j \to \infty$, we get

$$\begin{aligned} &\int_\Omega \bigl\langle h_{2R} \circ |f(x, \cdot)|, \nu_x \bigr\rangle \,\mathrm{d}x + \int_{\overline{\Omega}} \bigl\langle |f^\infty(x, \cdot)|, \nu_x^\infty \bigr\rangle \,\mathrm{d}\lambda_\nu(x) \\ &\qquad \leq \eta_R \leq \int_\Omega \bigl\langle h_R \circ |f(x, \cdot)|, \nu_x \bigr\rangle \,\mathrm{d}x + \int_{\overline{\Omega}} \bigl\langle |f^\infty(x, \cdot)|, \nu_x^\infty \bigr\rangle \,\mathrm{d}\lambda_\nu(x). \end{aligned}$$

Letting $R \uparrow \infty$ and invoking the Monotone Convergence Theorem, (2.12) follows and the proof is finished. □

Remark 2.21. The Fundamental Theorem for classical Young measures (see for example Theorem 3.1 in [96]) states that $f(x, u_j(x)) \rightharpoonup \bigl\langle f(x, \cdot), \nu_x \bigr\rangle$ if and only if the sequence $(f(x, u_j(x)))$ is equiintegrable. Using the preceding proposition, this result can now be recovered from our Fundamental Theorem 2.10.

Intuitively, the absence of oscillations makes a sequence converge in measure. This is made precise in the following proposition (see Remark 2.10 in [3]).

Proposition 2.22. *Let $(u_j) \subset \mathrm{L}^1(\Omega;\mathbb{R}^N)$ with $u_j \xrightarrow{\mathbf{Y}} \nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$. Then, the sequence (u_j) is converges in measure to $u \in \mathrm{L}^1(\Omega;\mathbb{R}^N)$ if and only if $\nu_x = \delta_{u(x)}$ a.e.*

Proof. Take the bounded, positive integrand

$$f(x,A) := \frac{|A-u(x)|}{1+|A-u(x)|}, \qquad x \in \Omega,\ A \in \mathbb{R}^N,$$

which lies in $\mathbf{R}(\Omega;\mathbb{R}^N)$ and satisfies $f^\infty \equiv 0$. Then, for all $\delta \in (0,1)$, Markov's Inequality and Proposition 2.17 (i) yield

$$\begin{aligned}\limsup_{j\to\infty} \bigl|\{\, x\in\Omega \ :\ f(x,u_j(x)) \geq \delta \,\}\bigr| &\leq \lim_{j\to\infty} \frac{1}{\delta}\int_\Omega f\bigl(x,u_j(x)\bigr)\,\mathrm{d}x \\ &= \frac{1}{\delta}\int_\Omega \bigl\langle f(x,\bullet),\nu_x\bigr\rangle\,\mathrm{d}x.\end{aligned}$$

On the other hand,

$$\begin{aligned}\int_\Omega \bigl\langle f(x,\bullet),\nu_x\bigr\rangle\,\mathrm{d}x &= \lim_{j\to\infty}\int_\Omega f\bigl(x,u_j(x)\bigr)\,\mathrm{d}x \\ &\leq \delta|\Omega| + \limsup_{j\to\infty}\bigl|\{\, x\in\Omega \ :\ f(x,u_j(x)) \geq \delta \,\}\bigr|.\end{aligned}$$

The preceding estimates show that $f(x,u_j(x))$ converges to zero in measure if and only if $\langle f(x,\bullet),\nu_x\rangle = 0$ for $\mathcal{L}^d$-almost every $x\in\Omega$, which is the case if and only if $\nu_x = \delta_{u(x)}$ almost everywhere. But clearly, $f(x,u_j(x))$ converges to zero in measure if and only if u_j converges to u in measure. □

By Vitali's Convergence Theorem we immediately get from Propositions 2.20, 2.22 the following result, which states that for L^1-bounded sequences strong convergence is equivalent to the absence of both oscillations and concentrations.

Corollary 2.23. *Let $(u_j) \subset \mathrm{L}^1(\Omega;\mathbb{R}^N)$ with $u_j \xrightarrow{\mathbf{Y}} \nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$. Then, the sequence (u_j) is converges strongly in $\mathrm{L}^1(\Omega;\mathbb{R}^m)$ to $u \in \mathrm{L}^1(\Omega;\mathbb{R}^N)$ if and only if $\nu_x = \delta_{u(x)}$ a.e. and $\lambda_\nu = 0$.*

2.6 Gradient Young measures

In this work, the Young measures in which we are mainly interested arise from sequences of gradients or symmetrized gradients. This section will introduce **gradient Young measures**, the corresponding notion of BD-Young measures will be treated in Section 5.1.

Right from the start we consider Young measures generated by derivatives of BV-functions and not only those generated by $\mathrm{W}^{1,1}$-gradients. The general framework of the preceding sections is now specialized to the case $\mathbb{R}^N = \mathbb{R}^{m\times d}$ (recall that we are identifying

$\mathbb{R}^{m\times d}$ with $\mathbb{R}^{md}$) and we use $\mathbb{B}^{m\times d}$ and $\partial\mathbb{B}^{m\times d}$ for the unit ball and unit sphere in $\mathbb{R}^{m\times d}$. Again, we can also develop the theory for the local versions of the space BV and for local gradient Young measures, but we do not carry this out explicitly, despite using these spaces later.

2.6.1 Functions of bounded variation

Let $\Omega \subset \mathbb{R}^d$ be a bounded open Lipschitz domain (some extensions to non-Lipschitz Ω are possible, but we omit details here for simplicity). A function $u \in \mathrm{L}^1(\Omega;\mathbb{R}^m)$ is a **function of bounded variation** (see [10, 50, 130] for all the following facts) if its distributional derivative Du is a finite $\mathbb{R}^{m\times d}$-valued Radon measure. This means that there exists a (unique) measure $Du \in \mathbf{M}(\Omega;\mathbb{R}^{m\times d})$ such that for all $\psi \in \mathrm{C}^1_c(\Omega)$ the integration-by-parts formula

$$\int_\Omega \frac{\partial\psi}{\partial x_j} u^i \,\mathrm{d}x = -\int_\Omega \psi \,\mathrm{d}Du^i_j, \qquad i = 1,\dots,m,\ j = 1,\dots d,$$

holds; denote by $\mathrm{BV}(\Omega;\mathbb{R}^m)$ the space of these functions. Clearly, we have that $u \in \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ precisely when the measure Du is absolutely continuous with respect to $\mathcal{L}^d$.

Under the norm

$$\|u\|_{\mathrm{BV}(\Omega;\mathbb{R}^m)} := \|u\|_{\mathrm{L}^1(\Omega;\mathbb{R}^m)} + |Du|(\Omega)$$

the linear space $\mathrm{BV}(\Omega;\mathbb{R}^m)$ becomes a Banach space. The **strong (norm) convergence** $u_j \to u$, however, is rarely used. Instead, a sequence $(u_j) \subset \mathrm{BV}(\Omega;\mathbb{R}^m)$ is said to converge **weakly*** to $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ if $u_j \to u$ strongly in $\mathrm{L}^1(\Omega;\mathbb{R}^m)$ and $Du_j \overset{*}{\rightharpoonup} Du$ in $\mathbf{M}(\Omega;\mathbb{R}^{m\times d})$. Uniformly norm-bounded sequences in $\mathrm{BV}(\Omega;\mathbb{R}^m)$ always have a weakly* converging subsequence. If additionally $|Du_j|(\Omega) \to |Du|(\Omega)$, then the u_j are said to converge **strictly**. If even $\langle Du_j\rangle(\Omega) \to \langle Du\rangle(\Omega)$, where $\langle\,\cdot\,\rangle$ denotes the area functional, then we speak of **$\langle\,\cdot\,\rangle$-strict (area-strict) convergence**. It is well-known that smooth functions are $\langle\,\cdot\,\rangle$-strictly dense in $\mathrm{BV}(\Omega;\mathbb{R}^m)$, see Lemma 2.24 below. For more about the notion of $\langle\,\cdot\,\rangle$-strict convergence and why it is necessary and useful, see the proof of Reshetnyak's Continuity Theorem 2.3 and Counterexample 2.4.

Since Ω is a bounded open Lipschitz domain, it is possible to define the **trace** $u|_{\partial\Omega} \in \mathrm{L}^1(\partial\Omega,\mathcal{H}^{d-1};\mathbb{R}^m)$ (the space of $\mathcal{H}^{d-1}$-integrable functions on $\partial\Omega$ with values in $\mathbb{R}^m$) of a function in $\mathrm{BV}(\Omega;\mathbb{R}^m)$ onto $\partial\Omega$, which coincides with the natural trace for all $u \in \mathrm{BV}(\Omega;\mathbb{R}^m) \cap \mathrm{C}(\overline{\Omega};\mathbb{R}^m)$, see Section 3.8 of [10] for details. Denote by $\mathrm{BV}_g(\Omega;\mathbb{R}^m)$ the space of $\mathrm{BV}(\Omega;\mathbb{R}^m)$-functions with trace g on $\partial\Omega$; similarly for $\mathrm{W}^{1,1}_g(\Omega;\mathbb{R}^m)$. Recall that according to Gagliardo's Trace Theorem [62] the trace operator maps both spaces $\mathrm{BV}(\Omega;\mathbb{R}^m)$ and $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ surjectively onto $\mathrm{L}^1(\partial\Omega,\mathcal{H}^{d-1};\mathbb{R}^m)$. It is shown in Theorem 3.88 of [10] that the trace operator is continuous between $\mathrm{BV}(\Omega;\mathbb{R}^m)$ with the strict convergence and $\mathrm{L}^1(\partial\Omega,\mathcal{H}^{d-1};\mathbb{R}^m)$ with the strong (norm) convergence.

The next lemma shows that we can approximate BV-functions by smooth ones with respect to the $\langle\,\cdot\,\rangle$-topology. This is slightly stronger than the usual approximation with respect to the strict topology (as for instance in Theorem 3.9 in [10]), but since we will need to apply Reshetnyak's Continuity Theorem 2.3 later, we need this stronger version. The proof, however, is essentially the same and proceeds by mollification.

Lemma 2.24. *Let $\Omega \subset \mathbb{R}^d$ be a bounded open Lipschitz domain. For each $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ there exists $(v_j) \subset (\mathrm{W}^{1,1}_u \cap \mathrm{C}^\infty)(\Omega;\mathbb{R}^m)$ such that $v_j \to u$ $\langle\cdot\rangle$-strictly on Ω. If $u \in \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$, then we can in addition arrange that $v_j \to u$ strongly in $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$.*

A proof of this theorem can be found in Lemma B.1 of [26]. The result also holds true for more general domains Ω (with no regularity assumption on $\partial\Omega$), see the appendix of [82].

In the Lebesgue–Radon–Nikodým decomposition

$$Du = \nabla u\, \mathcal{L}^d \llcorner \Omega + D^s u$$

of the derivative of $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ (with respect to Lebesgue measure) we call the function $\nabla u \in \mathrm{L}^1(\Omega;\mathbb{R}^{m\times d})$ the **approximate gradient** of u and the measure $D^s u \in \mathbf{M}(\Omega;\mathbb{R}^{m\times d})$ the **singular part** of the derivative. The singular part $D^s u$ can be further split 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 originates from jumps in the function u and has the form

$$D^j u = (u^+ - u^-) \otimes n_{J_u}\, \mathcal{H}^{d-1} \llcorner J_u,$$

where $J_u \subset \Omega$ is the $\mathcal{H}^{d-1}$-rectifiable **jump set**, n_{J_u} is its normal (for a fixed orientation), and $u^\pm$ are the traces of u onto J_u (in positive and negative n_{J_u}-direction respectively). The Cantor part contains the remainder of $D^j u$. In general, one can show that if $\mathcal{H}^{d-1}(S) = 0$, then also $D^c u(S) = 0$ (for the jump part this is obviously also true), but otherwise this part may contain fractal measures, e.g. a Cantor measure, see Example 2.38 below. More information on both the jump and the Cantor part can be found in Section 3.7 of [10].

A highly non-trivial result in the theory of BV-functions is the following result of Alberti [1], which settled a conjecture of Ambrosio & De Giorgi [9]:

Theorem 2.25 (Alberti's Rank One Theorem). *Let $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$. Then,*

$$\mathrm{rank}\biggl(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x_0)\biggr) \le 1 \qquad \text{for } |D^s u|\text{-almost every } x_0 \in \Omega.$$

This theorem has the following "physical" interpretation: Even at ($D^c u$-almost all) points $x_0 \in \Omega$ around which $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ has Cantor-type (e.g. fractal) structure, the "slope" of u has a well-defined *direction*, given by the vector $\xi \in \mathbb{S}^{d-1}$ from $\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x_0) = a \otimes \xi$.

Of course, Alberti's Rank One Theorem is trivially true at jump points ($D^j u$-almost everywhere), but for the Cantor-part the proof is rather involved and based on a "decomposition technique" together with the BV-coarea formula. Recently, the proof was made more accessible in [38] and there is also an announcement of a completely new argument in [2].

2.6.2 Generation

Given $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$, we associate to its derivative $Du \in \mathbf{M}(\Omega;\mathbb{R}^{m\times d})$ an elementary Young measure $\varepsilon_{Du} \in \mathbf{Y}(\Omega;\mathbb{R}^{m\times d})$ just as in Section 2.3.3, i.e.

$$(\varepsilon_{Du})_x := \delta_{\nabla u(x)} \quad \mathcal{L}^d\text{-a.e.}, \qquad \lambda_{\varepsilon_{Du}} := |D^s u|, \qquad (\varepsilon_{Du})_x^\infty := \delta_{p(x)} \quad |D^s u|\text{-a.e.},$$

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 a norm-bounded sequence $(u_j) \subset \mathrm{BV}(\Omega;\mathbb{R}^m)$ we say that the derivatives Du_j **generate** the Young measure $\nu = (\nu_x, \lambda_\nu, \nu_x^\infty) \in \mathbf{Y}(\Omega;\mathbb{R}^{m\times d})$, in symbols $Du_j \overset{\mathbf{Y}}{\to} \nu$, if for all $f \in \mathbf{E}(\Omega;\mathbb{R}^{m\times d})$ we have that

$$\begin{aligned} &f\bigl(x, \nabla u_j(x)\bigr)\, \mathcal{L}^d_x \llcorner \Omega \;+\; f^\infty\biggl(x, \frac{\mathrm{d}D^s u_j}{\mathrm{d}|D^s u_j|}(x)\biggr)\, |D^s u_j|(x) \\ &\quad \overset{*}{\rightharpoondown} \;\bigl\langle f(x, \cdot), \nu_x \bigr\rangle\, \mathcal{L}^d_x \llcorner \Omega \;+\; \bigl\langle f^\infty(x, \cdot), \nu_x^\infty \bigr\rangle\, \lambda_\nu(x) \qquad \text{in } \mathbf{M}(\overline{\Omega}), \end{aligned} \tag{2.13}$$

or, equivalently,

$$\varepsilon_{Du_j} \overset{*}{\rightharpoondown} \nu \quad \text{in } \mathbf{Y}(\Omega;\mathbb{R}^{m\times d}).$$

The set $\mathbf{GY}(\Omega;\mathbb{R}^{m\times d})$ of **gradient Young measures** is the class of those Young measures $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^{m\times d})$ for which there exists a norm-bounded sequence (u_j) in $\mathrm{BV}(\Omega;\mathbb{R}^m)$ with $Du_j \overset{\mathbf{Y}}{\to} \nu$.

The next proposition follows directly from Reshetnyak's Continuity Theorem 2.3:

Proposition 2.26. *If $u_j \to u$ $\langle\cdot\rangle$-strictly in $\mathrm{BV}(\Omega;\mathbb{R}^m)$, then $Du_j \overset{\mathbf{Y}}{\to} \varepsilon_{Du}$.*

This also justifies our choice of elementary Young measures for BV-functions: Take a sequence $(u_j) \subset (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(\Omega;\mathbb{R}^m)$ with $u_j \to u$ $\langle\cdot\rangle$-strictly by Lemma 2.24. Then, the preceding proposition implies that $Du_j \overset{\mathbf{Y}}{\to} \varepsilon_{Du}$ as $j \to \infty$.

The Fundamental Theorem 2.10 now takes the following form:

Theorem 2.27 (Gradient Young measure generation). *Let $(u_j) \subset \mathrm{BV}(\Omega;\mathbb{R}^m)$ be a uniformly norm-bounded sequence. Then, there exists a subsequence (not relabeled) such that $Du_j \overset{\mathbf{Y}}{\to} \nu$ for some Young measure $\nu \in \mathbf{GY}(\Omega;\mathbb{R}^{m\times d})$.*

What can one say about the barycenter $[\nu] \in \mathbf{M}(\overline{\Omega};\mathbb{R}^{m\times d})$ of a gradient Young measure $\nu \in \mathbf{GY}(\Omega;\mathbb{R}^{m\times d})$? For classical Young measures (say $(u_j) \subset \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ equiintegrable), barycenters are always $\mathrm{W}^{1,1}$-gradients, so in our more general BV-setting it is somewhat natural to expect that $\xi = Du$ for some $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$. This, however, is not entirely true:

Counterexample 2.28. Let $\Omega = (0,1)$ and define $u_j := -\mathbb{1}_{(0,1/j)}$. Then we have $u_j \overset{*}{\rightharpoondown} 0$ in $\mathrm{BV}(\Omega;\mathbb{R})$, $Du_j = \delta_{1/j}$ and $Du_j \overset{\mathbf{Y}}{\to} \nu$ with $\nu_x = \delta_0$ a.e., $\lambda_\nu = \delta_0$ and $\nu_0^\infty = \delta_{+1}$. But $[\nu] = \delta_0$ is not a BV-derivative on $(0,1)$.

Of course, the problem in this counterexample is that a sequence may concentrate at the boundary of Ω and this effect must be included in the concentration measure if our integrand is nonzero on the boundary of Ω, otherwise the limit representation of a functional through the Young measure may fail:

Counterexample 2.29. Take all ingredients from the preceding counterexample, but assume that we are not transforming to $\overline{\Omega \times \mathbb{B}^1}$, but only to $\Omega \times \overline{\mathbb{B}^1}$ in our Young measure framework. Then we have $\nu_j \overset{*}{\rightharpoonup} \tilde{\nu}$ in $\mathbf{Y}(\Omega;\mathbb{R})$ with $\tilde{\nu}_x = \delta_0$ a.e. and $\lambda_{\tilde{\nu}} = 0$. However, for the integrand $f(x, A) := |A|$ we get

$$\lim_{j\to\infty} \int_\Omega \left| \frac{\mathrm{d}Du_j}{\mathrm{d}|Du_j|} \right| \mathrm{d}|Du_j| = \lim_{j\to\infty} |Du_j|(\Omega) = 1 \neq 0 = \big\langle\!\big\langle |\cdot|, \tilde{\nu} \big\rangle\!\big\rangle,$$

so the Young measure representation fails. Notice that the representation holds with ν from Counterexample 2.28 instead of $\tilde{\nu}$.

True, however, is that the barycenter *restricted to* Ω is a BV-derivative. Indeed, for a sequence $u_j \overset{*}{\rightharpoonup} u$ in $\mathrm{BV}(\Omega;\mathbb{R}^m)$ and the integrand $f := \mathrm{id}$ (or more precisely componentwise), we get from (2.13)

$$Du_j \overset{*}{\rightharpoonup} \big\langle \mathrm{id}, \nu_x \big\rangle \, \mathcal{L}^d \llcorner \Omega + \big\langle \mathrm{id}, \nu_x^\infty \big\rangle \, \lambda_\nu = [\nu] \qquad \text{in } \mathbf{M}(\overline{\Omega};\mathbb{R}^{m\times d}).$$

But on the other hand

$$Du_j \overset{*}{\rightharpoonup} Du \qquad \text{in } \mathbf{M}(\Omega;\mathbb{R}^{m\times d}),$$

hence

$$Du = [\nu] \llcorner \Omega = \big\langle \mathrm{id}, \nu_x \big\rangle \, \mathcal{L}^d \llcorner \Omega + \big\langle \mathrm{id}, \nu_x^\infty \big\rangle \, \lambda_\nu \llcorner \Omega.$$

Note that λ_ν might be nontrivial even if the barycenter is a $\mathrm{W}^{1,1}$-gradient, see Example 2.36 below. We refer to any $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ such that $Du = [\nu]\llcorner\Omega$ as an **underlying deformation** for ν.

The above considerations allow us to transport two well-known statements about BV-derivatives to our gradient Young measures. First, $|[\nu]|\llcorner\Omega$ cannot charge sets $S \subset \Omega$ with $\mathcal{H}^{d-1}(S) = 0$, see for instance Lemma 3.76 in [10]. Furthermore,

$$D^s u = \big\langle \mathrm{id}, \nu_x^\infty \big\rangle \, \lambda_\nu^s \llcorner \Omega$$

and if $\lambda_\nu^s \llcorner \Omega = a|D^s u| + \lambda_\nu^*$ is the Lebesgue–Radon–Nikodým decomposition of $\lambda_\nu^s \llcorner \Omega$ with respect to $|D^s u|$, then

$$\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x) = \big\langle \mathrm{id}, \nu_x^\infty \big\rangle \, a(x) \qquad \text{for } |D^s u|\text{-a.e. } x \in \Omega$$

and

$$\big\langle \mathrm{id}, \nu_x^\infty \big\rangle = 0 \qquad \lambda_\nu^*\text{-a.e.}$$

Therefore, by Alberti's Rank One Theorem 2.25, the matrix $\langle \mathrm{id}, \nu_x^\infty \rangle$ has rank one for $|D^s u|$-a.e. $x \in \Omega$ (notice $a(x) \neq 0$ for $|D^s u|$-a.e. $x \in \Omega$).

In many technical arguments it is important to improve generating sequences as follows:

Proposition 2.30. *Let $\nu \in \mathbf{GY}(\Omega; \mathbb{R}^{m \times d})$.*

(i) There exists a generating sequence $(u_j) \subset (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(\Omega; \mathbb{R}^m)$ with $Du_j \xrightarrow{\mathbf{Y}} \nu$.

(ii) If additionally $\lambda_\nu(\partial\Omega) = 0$, then the u_j from (i) can be chosen to satisfy $u_j|_{\partial\Omega} = u|_{\partial\Omega}$, where $u \in \mathrm{BV}(\Omega; \mathbb{R}^m)$ is an arbitrary underlying deformation of ν.

Proof. For (i) take a sequence $(v_j) \subset \mathrm{BV}(\Omega; \mathbb{R}^m)$ with $Dv_j \xrightarrow{\mathbf{Y}} \nu$. Let $\{f_k\} \subset \mathbf{E}(\Omega; \mathbb{R}^{m \times d})$ be as in Lemma 2.16 and for each fixed $j \in \mathbb{N}$ take $u_j \in (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(\Omega; \mathbb{R}^m)$ satisfying

$$\left| \int_\Omega f_k\bigl(x, \nabla u_j(x)\bigr) \,\mathrm{d}x - \int_\Omega f_k\bigl(x, \nabla v_j(x)\bigr) \,\mathrm{d}x - \int_\Omega f_k^\infty\left(x, \frac{\mathrm{d}D^s v_j}{\mathrm{d}|D^s v_j|}(x)\right) \mathrm{d}|D^s v_j|(x) \right| \leq \frac{1}{j}$$

whenever $k \leq j$. This is possible by Proposition 2.26 since smooth functions are $\langle \cdot \rangle$-strictly dense in $\mathrm{BV}(\Omega; \mathbb{R}^m)$ by Lemma 2.24. Thus,

$$\int_\Omega f_k\bigl(x, \nabla u_j(x)\bigr) \,\mathrm{d}x \to \langle\!\langle f_k, \nu \rangle\!\rangle \qquad \text{as } j \to \infty \qquad \text{for all } k \in \mathbb{N},$$

and by Lemma 2.16 we conclude that $\nabla u_j \xrightarrow{\mathbf{Y}} \nu$.

For (ii), by virtue of (i) take a sequence $(v_j) \subset (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(\Omega; \mathbb{R}^m)$ such that $Dv_j \xrightarrow{\mathbf{Y}} \nu$. Adjusting the v_j and taking another subsequence if necessary, we may also require $v_j \to u$ in $\mathrm{L}^1(\Omega; \mathbb{R}^m)$. Let moreover $(\rho_n) \subset \mathrm{C}_c^\infty(\Omega; [0,1])$ be a sequence of cut-off functions with $\rho_n \uparrow \mathbb{1}_\Omega$ pointwise as $n \to \infty$ and use Lemma 2.24 to get a sequence $(w_j) \subset (\mathrm{W}_u^{1,1} \cap \mathrm{C}^\infty)(\Omega; \mathbb{R}^m)$ such that $w_j \to u$ $\langle \cdot \rangle$-strictly. For

$$u_{j,n} := \rho_n v_j + (1 - \rho_n) w_j \in \mathrm{W}_u^{1,1}(\Omega; \mathbb{R}^m)$$

we observe

$$\lim_{n \to \infty} \lim_{j \to \infty} \|u_{j,n} - u\|_{\mathrm{L}^1(\Omega; \mathbb{R}^m)} = 0.$$

For every $f \in \mathbf{E}(\Omega; \mathbb{R}^{m \times d})$ with linear growth constant $M > 0$ we have

$$\bigl|\langle\!\langle f, \varepsilon_{Du_{j,n}} \rangle\!\rangle - \langle\!\langle f, \varepsilon_{Dv_j} \rangle\!\rangle\bigr| \leq M \int_{\Omega \setminus K_n} 2 + 2|\nabla v_j| + |\nabla w_j| + |v_j - w_j||\nabla \rho_n| \,\mathrm{d}x,$$

where $K_n := \{\, x \in \Omega \;:\; \rho_n(x) = 1 \,\}$. Hence (we may without loss of generality assume $(\mathcal{L}^d + \lambda_\nu)(\partial K_n) = 0$),

$$\begin{aligned}
&\limsup_{n \to \infty} \limsup_{j \to \infty} \bigl|\langle\!\langle f, \varepsilon_{Du_{j,n}} \rangle\!\rangle - \langle\!\langle f, \varepsilon_{Dv_j} \rangle\!\rangle\bigr| \\
&\qquad \leq \limsup_{n \to \infty} 2M \biggl(\int_{\overline{\Omega} \setminus K_n} \langle 1 + |\cdot|, \nu_x \rangle \,\mathrm{d}x + \lambda_\nu(\overline{\Omega} \setminus K_n) + |Du|(\overline{\Omega} \setminus K_n) \biggr) \\
&\qquad = 2M \lambda_\nu(\partial\Omega) = 0.
\end{aligned}$$

Thus, we can use a diagonalization argument to finish the proof. $\square$

Remark 2.31. The assumption $\lambda_\nu(\partial\Omega) = 0$ is necessary for (ii) to hold. Take for example the sequence $u_j := \mathbb{1}_{(0,1/j)}$ on $\Omega := (0,1)$, whence $Du_j = -\delta_{1/j}$ and $u_j \overset{*}{\rightharpoonup} 0$ in $\mathrm{BV}(0,1)$, but any attempt to modify the u_j to have zero boundary values necessarily introduces an additional part in the derivative, which destroys the Young measure representation for the integrand $f := \mathbb{1}_{\overline{\Omega}} \otimes |\bullet|$.

We close this section by citing the following result, Proposition 5 in [82]:

Proposition 2.32. *For any $\lambda \in \mathbf{M}^+(\overline{\Omega})$, there exists $\nu \in \mathbf{GY}(\Omega; \mathbb{R}^{m\times d})$ with $[\nu] = 0$ and $\lambda_\nu = \lambda$.*

2.6.3 Examples

We now provide a few examples of gradient Young measures to illustrate the theory and to exhibit model cases of oscillation and concentration effects. Most of these examples are standard and taken from [3, 21].

Example 2.33. Take an open Lipschitz domain $\Omega \subset \mathbb{R}^2$, let $A, B \in \mathbb{R}^{2\times 2}$ with $B - A = a \otimes n$, where $a, n \in \mathbb{R}^2$, and define

$$u(x) := Ax + \left(\int_0^{x\cdot n} \chi(t) \,\mathrm{d}t \right) a, \qquad x \in \Omega,$$

where $\chi(t) := \mathbb{1}_{\bigcup_{z\in\mathbb{Z}}[z,z+1/2)}$. If we let $u_j(x) := u(jx)/j$, $x \in \Omega$, then ∇u_j generates the homogeneous Young measure $\nu \in \mathbf{GY}(\Omega; \mathbb{R}^2)$ with

$$\nu_x = \frac{1}{2}\delta_A + \frac{1}{2}\delta_B \quad \text{a.e.}, \qquad \lambda_\nu = 0.$$

This example can also be extended to include multiple scales, cf. [96].

Next we look at the classic concentration examples.

Example 2.34. Take $\Omega := (0,1)$ and $u_j(x) = jx\mathbb{1}_{(0,1/j)}(x) + \mathbb{1}_{(1/j,1)}(x)$, $x \in \Omega$. Hence, $u_j' = j\mathbb{1}_{(0,1/j)}$ and $u_j' \overset{\mathbf{Y}}{\to} \nu \in \mathbf{GY}((0,1); \mathbb{R})$ with

$$\nu_x = \delta_0 \quad \text{a.e.}, \qquad \lambda_\nu = \delta_0, \qquad \nu_0^\infty = \delta_{+1}.$$

Example 2.35. Take $\Omega := (-1,1)$ and set $v_j(x) := u_j(x) - u_j(-x)$, $x \in \Omega$, with the u_j from the previous example. Then $v_j' \overset{\mathbf{Y}}{\to} \nu \in \mathbf{GY}((0,1); \mathbb{R})$ with

$$\nu_x = \delta_0 \quad \text{a.e.}, \qquad \lambda_\nu = 2\delta_0, \qquad \nu_0^\infty = \frac{1}{2}\delta_{+1} + \frac{1}{2}\delta_{-1}.$$

The next example shows that also diffuse concentration may occur:

Example 2.36 (Diffuse concentration). In the domain $\Omega := (0,1)$ define the $\mathrm{L}^1(0,1)$-bounded sequence of derivatives

$$u_j' = \sum_{k=0}^{j-1} j\mathbb{1}_{\left(\frac{k}{j},\frac{k}{j}+\frac{1}{j^2}\right)}, \qquad j \in \mathbb{N}.$$

The u_j' converge to 0 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 $k \in \mathbb{N}$ (Chacon's biting lemma states that for every uniformly L^1-bounded sequence there always exists a subsequence converging in the biting sense, see for example Section 6.4 in [104]). On the other hand, the u_j' also converge to $\mathbb{1}$ weakly* in the sense of measures. The sequence is not equiintegrable, however, because

$$\lim_{M\uparrow\infty} \sup_{j\in\mathbb{N}} \int_{\{|u_j'|\geq M\}} |u_j'| \,\mathrm{d}x = 1,$$

and hence not weakly sequentially compact by the Dunford–Pettis Theorem. Note that mere non-equiintegrability in itself does not automatically imply that there is no L^1-weakly converging subsequence, cf. the subtleties in Section 4 of [21] (one can for example take the union of a non-equiintegrable sequence of functions and infinitely often insert one fixed function). Here, however, no L^1-weakly converging subsequence exists: 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 shown directly in [3] that $u_j' \xrightarrow{\mathbf{Y}} \nu$ with

$$\nu_x = \delta_0 \quad \text{a.e.}, \qquad \lambda_\nu = \mathcal{L}^1 \llcorner (0,1), \qquad \nu_x^\infty = \delta_{+1} \quad \text{a.e.}.$$

One can also get this result by an averaging procedure, see Section 5.2 in [82], in particular Example 7.

The fact that $\lambda_\nu \neq 0$ indicates that concentration occurs, which is agreement with Proposition 2.20 since the sequence (u_j') is not equiintegrable. Also, in accordance with point (i) of Theorem 2.9 in [3], the biting limit is 0. Observe that the barycenter is simply $Du = \mathcal{L}^1 \llcorner (0,1)$, so this example shows that the barycenter may be a $\mathrm{W}^{1,1}$-gradient even though the generating sequence concentrates. In other words, concentration does not always mean that the space $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ is left under weak* convergence (as was the case in Example 2.34).

The preceding example can also be adapted to make the "concentration occur in all directions", in the sense that the concentration-angle measure is absolutely continuous with respect to the $(m-1)$-dimensional Hausdorff measure on $\partial\mathbb{B}^{m\times 1} = \mathbb{S}^{m-1}$.

Example 2.37. Modifying the last example to (now $m = 2$)

$$u_j'(x) = \sum_{k=0}^{j-1} j\mathbb{1}_{\left(\frac{k}{j},\frac{k}{j}+\frac{1}{j^2}\right)}(x) \begin{bmatrix}\cos(2\pi j^2 x)\\ \sin(2\pi j^2 x)\end{bmatrix},$$

we get (see [3] for details)

$$\nu_x = \delta_0 \quad \mathcal{L}^d\text{-a.e.}, \qquad \lambda_\nu = \mathcal{L}^1 \llcorner (0,1), \qquad \nu_x^\infty = \frac{1}{2\pi}\mathcal{H}^1 \llcorner \mathbb{S}^1 \quad \mathcal{L}^1\text{-a.e.}$$

Example 2.38 (Cantor–Vitali functions). Fix $0 < \delta < 1/2$ and let C_j be the jth step in the construction of the Cantor set associated with δ, i.e. $C_0 := [0,1]$ and we move from C_j to C_{j+1} by removing from each interval in C_j the centered open interval of length $\delta^j(1-2\delta)$. Then, the associated Cantor set is $C := \bigcap_j C_j$. The Cantor–Vitali function $f_C \in \mathrm{C}([0,1])$ can be defined as the uniform limit of the functions

$$f_j(x) := \frac{1}{(2\delta)^j} \int_0^x \mathbb{1}_{C_j} \,\mathrm{d}y.$$

It can be shown that $f_j \to f_C$ $\langle\boldsymbol{\cdot}\rangle$-strictly in $\mathrm{BV}(0,1)$ and that $f_C(x) = \alpha^{-1}\mathcal{H}^\gamma(C \cap [0,x])$, corresponding to $Df_C = \alpha^{-1}\mathcal{H}^\gamma \llcorner C$. Here, $\gamma = \ln 2/\ln(1/\delta)$ and $\alpha = \mathcal{H}^\gamma(C) = 2^{-\gamma}\omega_\gamma = 2^{-\gamma}\pi^{\gamma/2}/\Gamma(1+\gamma/2)$, where Γ denotes Euler's Γ-function. See [65], pp. 15–21, for the details. Consequently, by Proposition 2.26 the Young measure ν generated by (Df_j) is

$$\nu_x = \delta_0 \quad \mathcal{L}^d\text{-a.e.}, \qquad \lambda_\nu = \alpha^{-1}\,\mathcal{H}^\gamma \llcorner C, \qquad \nu_x^\infty = \delta_{+1} \quad \mathcal{H}^\gamma\text{-a.e.}$$

Chapter 3

Tangent measures and tangent Young measures

As important technical tools for the lower semicontinuity proofs in subsequent chapters, following [108] we present two localization principles for Young measures, one at regular and one at singular points. The Young measures obtained in the blow-up are called *tangent Young measures* and complement classical tangent measures (reviewed in the first section of this chapter) in blow-up arguments involving Young measures.

The localization principles are only exhibited for gradient Young measures (to show that the gradient property is preserved in the blow-up), but we remark that all the results also extend to the BD-Young measures introduced in Section 5.1, or even non-constrained Young measures, by obvious generalizations.

3.1 Tangent measures

Tangent measures are a powerful tool in Geometric Measure Theory for investigating the local structure of Radon measures. In contrast to the apparently prevalent notion of tangent measure in the Calculus of Variations, which restricts tangent measures to the unit ball (see for example Section 2.7 in [10]), we here use Preiss's original definition [105]. This has several advantages from a technical point of view. In particular, we can use the general theory for tangent measures, which will prove to be useful. General information on tangent measures can be found in Chapter 14 of [89], Section 2.7 in [10], and [105].

Let $T^{(x_0,r)}(x) := (x - x_0)/r$ for $x_0 \in \mathbb{R}^d$ and $r > 0$. For a vector-valued Radon measure $\mu \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$ and $x_0 \in \mathbb{R}^d$, a **tangent measure** to μ at x_0 is any (local) weak* limit in the space $\mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$ of the rescaled measures $c_n T_*^{(x_0,r_n)}\mu$ for some sequence $r_n \downarrow 0$ of radii and strictly positive rescaling constants $c_n > 0$. The set of all such tangent measures is denoted by $\mathrm{Tan}(\mu, x_0)$ and the sequence $c_n T_*^{(x_0,r_n)}\mu$ is called a **blow-up sequence**. From the definition it follows that $\mathrm{Tan}(\mu, x_0) = \{0\}$ for all $x_0 \notin \operatorname{supp} \mu$ and moreover $0 \in \mathrm{Tan}(\mu, x_0)$ for all $x_0 \in \mathbb{R}^d$. Preiss originally excluded the zero measure from $\mathrm{Tan}(\mu, x_0)$ explicitly, but for us it has some technical advantages to include it.

Is is a fundamental result of Preiss that the set $\mathrm{Tan}(\mu, x_0)$ contains non-zero measures at $|\mu|$-almost every $x_0 \in \operatorname{supp}\mu$ (or, equivalently, at $|\mu|$-almost every $x_0 \in \mathbb{R}^d$). This is proved in Theorem 2.5 of [105], but since this is the only result from Preiss's paper needed here, we reproduce its proof in our notation.

Lemma 3.1. *Let $\mu \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^N)$. At $|\mu|$-almost every $x_0 \in \mathbb{R}^d$, the set $\mathrm{Tan}(\mu, x_0)$ contains a non-zero measure.*

Proof. Using Proposition 3.2 below, we may assume that μ is a positive measure. Moreover, restricting if necessary to a sufficiently large closed ball containing x_0, we can even assume $\mu \in \mathbf{M}^+(K)$ for some compact set $K \subset \mathbb{R}^d$ with $x_0 \in K$.

Step 1. First, we note that for all relatively compact Borel sets $A \subset \mathbb{R}^d$ it holds that

$$\mu(A) = \frac{1}{\omega_d r^d} \int \mu(A \cap B(x,r)) \,\mathrm{d}x, \tag{3.1}$$

where ω_d denotes the volume of the unit ball in $\mathbb{R}^d$. This follows with the aid of Fubini's Theorem,

$$\begin{aligned} \int \mu(A \cap B(x,r)) \,\mathrm{d}x &= \int \int \mathbb{1}_A(y) \mathbb{1}_{B(x,r)}(y) \,\mathrm{d}\mu(y) \,\mathrm{d}x \\ &= \int \mathbb{1}_A(y) \int \mathbb{1}_{B(y,r)}(x) \,\mathrm{d}x \,\mathrm{d}\mu(y) = \omega_d r^d \mu(A). \end{aligned}$$

Step 2. We now show that for all $t > 1$ it holds that

$$\lim_{\beta \to \infty} \limsup_{r \downarrow 0} \mu\bigl(\bigl\{ x \in K \;:\; \mu(B(x,tr)) \geq \beta \mu(B(x,r)) \bigr\}\bigr) = 0. \tag{3.2}$$

For this, let $\varepsilon > 0$, $\beta > (2(t+1))^d \mu(K)/\varepsilon$ and fix any $r > 0$. Also define

$$E := \bigl\{ x \in K \;:\; \mu(B(x,tr)) \geq \beta \mu(B(x,r)) \bigr\}.$$

Whenever $B(x, r/2) \cap E \neq \emptyset$ for some $r > 0$, take $z \in B(x, r/2) \cap E$ to estimate

$$\beta \mu(B(x, r/2)) \leq \beta \mu(B(z,r)) \leq \mu(B(z,tr)) \leq \mu(B(x,(t+1)r)).$$

Hence we get from (3.1),

$$\begin{aligned} \mu(E) &= \frac{1}{\omega_d \cdot (r/2)^d} \int \mu(E \cap B(x, r/2)) \,\mathrm{d}x \\ &\leq \frac{(2(t+1))^d}{\beta} \cdot \frac{1}{\omega_d \cdot ((t+1)r)^d} \int \mu(B(x,(t+1)r)) \,\mathrm{d}x \\ &= \frac{(2(t+1))^d}{\beta} \mu(K) < \varepsilon. \end{aligned}$$

This clearly implies (3.2). In fact, it even implies this assertion with the limes superior replaced by the supremum over all $r > 0$. This, however, is due to the fact that we without loss of generality restricted the measure μ to the compact set K, and so a smallness assumption on r is already implicit.

Step 3. From (3.2) we see that for all $\varepsilon > 0$ and all $k = 2, 3, \ldots$ there exists constants $\beta_k > 0$ and $t_k > 0$ such that

$$\mu\big(\{\, x \in K \ \mathbf{:} \ \mu(B(x,kr)) \geq \beta_k \mu(B(x,r)) \,\}\big) \leq \frac{\varepsilon}{2^k} \qquad \text{whenever } r \in (0,t_k).$$

Then, for $r > 0$ set

$$A_r := \Big\{\, x \in K \ \mathbf{:} \ \text{there exists } k \in \{2,3,\ldots\} \text{ with } r \in (0,t_k) \text{ such that} \\ \mu(B(x,kr)) \geq \beta_k \mu(B(x,r)) \,\Big\}$$

and observe that $\mu(E_r) \leq \varepsilon$ by the previous estimate. Hence, also

$$A := \bigcup_{i=1}^{\infty} \bigcap_{j=i}^{\infty} A_{1/j}$$

satisfies $\mu(A) \leq \varepsilon$. Since $\varepsilon > 0$ was arbitrary, this implies $\mu(A) = 0$.

Let now $x \in K \setminus A$. Then, for all $i \in \mathbb{N}$ there exists $j \geq i$ such that $x \notin A_{1/j}$, i.e. for all $k \in \mathbb{N}$ with $1/j \leq t_k$,

$$\mu(B(x,k/j)) \leq \beta_k \mu(B(x,1/j)).$$

Therefore, for μ-almost every $x_0 \in \operatorname{supp} \mu$ (and hence μ-almost every $x_0 \in K$), there exists a sequence $r_n \downarrow 0$ with

$$\limsup_{n\to\infty} \frac{\mu(B(x_0,kr_n))}{\mu(B(x_0,r_n))} \leq \beta_k \qquad \text{for all } k \in \mathbb{N}.$$

This allows us to infer that the sequence $c_n T_*^{(x_0,r_n)} \mu$ with $c_n := \mu(B(x_0,r_n))^{-1}$ is weakly* compact in $\mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d)$ and every weak* limit of a subsequence is a non-zero tangent measure to μ at x_0. $\square$

One can show (see Remark 14.4 (i) of [89]) that for any non-zero $\tau \in \operatorname{Tan}(\mu, x_0)$ we may always choose the rescaling constants c_n in the blow-up sequence $c_n T_*^{(x_0,r_n)} \mu \overset{*}{\rightharpoonup} \tau$ to be

$$\tilde{c}_n := c\big[|\mu|(x_0 + r_n \overline{U})\big]^{-1} \tag{3.3}$$

for any bounded open set $U \subset \mathbb{R}^d$ containing the origin such that $|\tau|(U) > 0$ and some constant $c = c(U) > 0$. This involves passing to a (non-relabeled) subsequence if necessary. Indeed, we have

$$0 < |\tau|(U) \leq \liminf_{n\to\infty} c_n |\mu|(x_0 + r_n U) \leq \limsup_{n\to\infty} c_n |\mu|(x_0 + r_n \overline{U}) < \infty$$

by the boundedness of $c_n T_*^{(x_0,r_n)} |\mu|$. Hence, after selection of a subsequence, $c_n |\mu|(x_0 + r_n \overline{U}) \to c$, and so $\tilde{c}_n T_*^{(x_0,r_n)} \mu \overset{*}{\rightharpoonup} \tau$.

It turns out that at a given point $x_0 \in \operatorname{supp} \mu$, different blow-up sequences might behave very differently. Most starkly, this phenomenon can be observed for the O'Neil measure [102], which has *every* measure as tangent measure at almost every point (with respect to the measure itself). Therefore, we need to distinguish several classes of blow-up sequences $\gamma_n := c_n T_*^{(x_0,r_n)}\mu$ for a measure $\mu \in \mathbf{M}(\mathbb{R}^d;\mathbb{R}^N)$ at a point $x_0 \in \operatorname{supp} \mu$:

- **Regular blow-up:** x_0 is a regular point of μ (that is $\lim_{r\downarrow 0} r^{-d}|\mu|(B(x_0,r)) < \infty$) and $\gamma_j \overset{*}{\rightharpoonup} \alpha A_0 \mathcal{L}^d$, where $A_0 = \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^d}(x_0)$ and $\alpha \in \mathbb{R}$ is a constant.
- **Semi-regular blow-up:** x_0 is a singular point of μ (that is $\lim_{r\downarrow 0} r^{-d}|\mu|(B(x_0,r)) = \infty$), but nevertheless $\gamma_j \overset{*}{\rightharpoonup} \alpha A_0 \mathcal{L}^d$, where $A_0 = \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x_0)$ and $\alpha \in \mathbb{R}$
- **Fully singular blow-up:** x_0 is a singular point of μ and $\gamma_j \overset{*}{\rightharpoonup} \tau$, but $\tau \neq A\mathcal{L}^d$ for any $A \in \mathbb{R}^N$.

In some sense conversely to the O'Neil measure alluded to above, Preiss exhibited a positive and purely singular measure on a bounded interval (in particular a BV-derivative) such that all tangent measures are a fixed multiple of Lebesgue measure, see Example 5.9 (1) in [105]. In our terminology this means that all blow-ups at singular points are semi-regular.

However, despite the possibly diverse behavior of different tangent measures at the same point, we can still establish several important structure properties.

Proposition 3.2. *Let $\mu \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$. At $|\mu|$-almost every $x_0 \in \mathbb{R}^d$ and for all sequences $r_n \downarrow 0$, $c_n > 0$, it holds that*

$$\tau = \operatorname*{w^*\text{-}lim}_{n\to\infty} c_n T_*^{(x_0,r_n)}\mu \qquad \textit{if and only if} \qquad |\tau| = \operatorname*{w^*\text{-}lim}_{n\to\infty} c_n T_*^{(x_0,r_n)}|\mu|, \tag{3.4}$$

and in this case

$$\tau = \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x_0) \cdot |\tau|.$$

In particular,

$$\mathrm{Tan}(\mu, x_0) = \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x_0) \cdot \mathrm{Tan}(|\mu|, x_0). \tag{3.5}$$

Proof. This proof is adapted from Theorem 2.44 in [10].

Let $f := \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}$, so $\mu = f|\mu|$ and $|f| = 1$ $|\mu|$-almost everywhere, and take $x_0 \in \mathbb{R}^d$ to be any Lebesgue point of f with respect to $|\mu|$. Then, for every $\varphi \in \mathrm{C}_c(\mathbb{R}^d)$ observe

$$\begin{aligned}
&\int \varphi \,\mathrm{d}T_*^{(x_0,r_n)}|\mu| - \int \varphi\, f(x_0) \cdot \mathrm{d}T_*^{(x_0,r_n)}\mu \\
&\qquad = \int \varphi(y)\bigl[1 - f(x_0)\cdot f(x_0 + r_n y)\bigr] \,\mathrm{d}T_*^{(x_0,r_n)}|\mu|(y) \\
&\qquad = \int \varphi\Bigl(\frac{x - x_0}{r_n}\Bigr)\bigl[1 - f(x_0)\cdot f(x)\bigr] \,\mathrm{d}|\mu|(x).
\end{aligned}$$

Further, from (3.3) we can assume that $c_n = c[|\mu|(B(x_0, r_n R))]^{-1}$ for some $c \geq 0$ and a large ball $B(0,R) \supset\supset \operatorname{supp}\varphi$ ($R > 0$). The inequality $|1 - f(x_0) \cdot f(x)| \leq |f(x) - f(x_0)|$ and the Lebesgue point property of x_0 therefore imply

$$\begin{aligned} &c_n \left| \int \varphi \, \mathrm{d}T_*^{(x_0,r_n)}|\mu| - \int \varphi \, f(x_0) \cdot \mathrm{d}T_*^{(x_0,r_n)}\mu \right| \\ &\quad \leq c\|\varphi\|_\infty \fint_{B(x_0,r_n R)} |f(x) - f(x_0)| \, \mathrm{d}|\mu|(x) \\ &\quad \to 0 \qquad \text{as } n \to \infty. \end{aligned}$$

This shows that if $c_n T_*^{(x_0,r_n)}\mu$ converges weakly* to $\tau \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$, then $c_n T_*^{(x_0,r_n)}|\mu|$ converges weakly* to the positive measure $\sigma = f(x_0) \cdot \tau$. If $\tau = g|\tau|$, then we get from general measure theory results (see for example Proposition 1.62(b) in [10]) that

$$|\tau| \leq \sigma = f(x_0) \cdot \tau = (f(x_0) \cdot g)|\tau|,$$

which immediately yields $g = f(x_0)$ $|\tau|$-almost everywhere, i.e. $\sigma = |\tau|$.

On the other hand, if $c_n T_*^{(x_0,r_n)}|\mu|$ converges weakly* to a measure $\sigma \in \mathbf{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$, then we also get (selecting a subsequence if necessary) that $c_n T_*^{(x_0,r_n)}\mu \overset{*}{\rightharpoondown} \tau \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$. Immediately, the previous argument applies and again yields $\sigma = |\tau|$.

Finally, (3.5) follows immediately from (3.4). ☐

If $\mu \in \mathbf{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ is absolutely continuous with respect to a positive measure $\lambda \in \mathbf{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$, then $\mathrm{Tan}(\mu, x_0) = \mathrm{Tan}(\lambda, x_0)$ for μ-almost all $x_0 \in \mathbb{R}^d$. This fact is proved in Lemma 14.6 of [89] and is particularly powerful in conjunction with the following result, see Lemma 14.5 of [89]: For a Borel set $E \subset \mathbb{R}^d$, at all μ-density points $x_0 \in \operatorname{supp}\mu$ of E, i.e. all points $x_0 \in \operatorname{supp}\mu$ such that

$$\lim_{r \downarrow 0} \frac{\mu(B(x_0,r) \setminus E)}{\mu(B(x_0,r))} = 0,$$

it holds that

$$\mathrm{Tan}(\mu, x_0) = \mathrm{Tan}(\mu \llcorner E, x_0).$$

In particular, this relation holds for μ-almost every $x_0 \in E$.

As an application, we can first cut off the singular part of an arbitrary measure $\mu \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$, then use the first fact on the remaining (absolutely continuous) part, and also (3.5), to see

$$\mathrm{Tan}(\mu, x_0) = \left\{ \alpha \, \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^d}(x_0) \, \mathcal{L}^d \;\; \mathbf{:} \;\; \alpha \in \mathbb{R} \right\} \qquad \text{for } \mathcal{L}^d\text{-a.e. } x_0 \in \mathbb{R}^d. \tag{3.6}$$

The next fact, that tangent measures to tangent measures are again tangent measures, is very important for our theory.

Proposition 3.3. *Let $\mu \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$. For $|\mu|$-almost every $x_0 \in \mathbb{R}^d$ and every $\tau \in \mathrm{Tan}(\mu, x_0)$, it holds that $\mathrm{Tan}(\tau, y_0) \subset \mathrm{Tan}(\mu, x_0)$ for all $z \in \mathbb{R}^d$.*

The following proof is from Theorem 14.16 in [89]. Note that since in our definition $\mathrm{Tan}(\mu, x_0)$ contains the zero measure for *every* $x_0 \in \mathbb{R}^d$, in the statement above we can allow y_0 arbitrary instead of just from $\operatorname{supp}\tau$ as in *loc. cit.*

Proof. First, by Proposition 3.2 we can confine ourselves to the case of positive measures.

For this proof we recall that there exist pseudometrics $F_k\colon \mathbf{M}^+_{\mathrm{loc}}(\mathbb{R}^d) \times \mathbf{M}^+_{\mathrm{loc}}(\mathbb{R}^d) \to [0,\infty]$, $k \in \mathbb{N}$, such that

$$\mu_j \overset{*}{\rightharpoondown} \mu \quad \text{in } \mathbf{M}^+_{\mathrm{loc}}(\mathbb{R}^d) \qquad \text{if and only if} \qquad F_k(\mu_j, \mu) \to 0 \quad \text{for all } k \in \mathbb{N}.$$

The F_k satisfy all properties of a metric except that $F_k(\mu, \nu) = 0$ for only one $k \in \mathbb{N}$ does not imply $\mu = \nu$ (but if $F_k(\mu, \nu) = 0$ for *all* $k \in \mathbb{N}$, then again $\mu = \nu$ follows). Moreover, the convergence induced by the collection $\{F_k\}$ is separable in the sense that there exists a countable set $S \subset \mathbf{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$ such that for given $\mu \in \mathbf{M}^+_{\mathrm{loc}}(\mathbb{R}^d)$, $\varepsilon > 0$, and $k \in \mathbb{N}$, we can find $\nu \in S$ with $F_k(\mu, \nu) < 0$. The definition and properties of these pseudometrics can be found in Chapter 14 of [89].

Step 1. First we show that for μ-almost every $x_0 \in \mathbb{R}^d$, all $\tau \in \mathrm{Tan}(\mu, x_0)$ satisfy the following assertion:

$$T^{(z,\rho)}_* \tau \in \mathrm{Tan}(\mu, x_0) \qquad \text{for all } z \in \operatorname{supp}\tau \text{ and all } \rho > 0. \tag{3.7}$$

It is clear that for $\rho \neq 1$ we can write $T^{(z,\rho)}_* \tau = T^{(z,1)}_* T^{(0,\rho)}_* \tau$, and the assertion for $z = 0$ is trivial. Thus, from now on assume $\rho = 1$.

Define the exceptional sets $E_{k,m}$, where $k, m \in \mathbb{N}$, to contain all $x \in \mathbb{R}^d$ such that there exist $\tau_x \in \mathrm{Tan}(\mu, x) \setminus \{0\}$ and $z_x \in \operatorname{supp}\tau$ with

$$F_k\big(T^{(z_x,1)}_* \tau_x, cT^{(x,r)}_* \mu\big) > \frac{1}{k} \qquad \text{for all } c > 0 \text{ and all } r \in (0, 1/m).$$

We will show that $\mu(E_{k,m}) = 0$ for all $k, m \in \mathbb{N}$. Then, the claim at the beginning of this step holds for all $x_0 \in \mathbb{R}^d \setminus \bigcup_{k,m} E_{k,m}$.

To the contrary assume that for $k, m \in \mathbb{N}$ we have $\mu(E_{k,m}) > 0$. With the countable $\{F_k\}$-dense set S (see above), we know that the sets

$$\big\{\, x \in E_{k,m} \ : \ F_k(T^{(z_x,1)}_* \tau_x, \nu) < 1/(4k) \,\big\}, \qquad \nu \in S,$$

form a countable covering of $E_{k,m}$, hence one of them must have positive μ-measure. Therefore, we infer the existence of a set $E \subset E_{k,m}$ with $\mu(E) > 0$ and

$$F_k\big(T^{(z_x,1)}_* \tau_x, T^{(z_y,1)}_* \tau_y\big) < \frac{1}{2k} \qquad \text{for all } x, y \in E. \tag{3.8}$$

Let $w \in E$ be a μ-density point of E, that is,

$$\lim_{r \downarrow 0} \frac{\mu(\overline{B(w,r)} \setminus E)}{\mu(\overline{B(w,r)})} = 0;$$

such a point always exists by Lebesgue's Differentiation Theorem. Since $w \in E \subset E_{k,m}$, there exists $\tau_w \in \mathrm{Tan}(\mu, w) \setminus \{0\}$ and $z_w \in \operatorname{supp} \tau_w$ as above. Take sequences $r_n \downarrow 0$, $c_n > 0$ with $c_n T_*^{(w,r_n)} \mu \xrightarrow{*} \tau_w$. Then, we will show that

$$\lim_{n\to\infty} \frac{\operatorname{dist}(w + r_n z_w, E)}{r_n} = 0. \tag{3.9}$$

If this is false, we can find $\delta \in (0, |z_w|)$ with $\operatorname{dist}(w + r_n z_w, E) > \delta r_n$ for infinitely many n; without loss of generality we assume that this holds for all n. Because the closed ball $\overline{B(w + r_n z_w, \delta r_n)}$ lies completely outside E, we get

$$\begin{aligned}\frac{\tau_w(B(z_w,\delta))}{\tau_w(\overline{B(0,2|z_w|)})} &\le \liminf_{n\to\infty} \frac{c_n T_*^{(w,r_n)}\mu(\overline{B(z_w,\delta)})}{c_n T_*^{(w,r_n)}\mu(\overline{B(0,2|z_w|)})} = \liminf_{n\to\infty} \frac{\mu(\overline{B(w + r_n z_w, \delta r_n)})}{\mu(\overline{B(w, 2r_n|z_w|)})} \\ &\le \lim_{n\to\infty} \frac{\mu(\overline{B(w,2r_n|z_w|)} \setminus E)}{\mu(\overline{B(w,2r_n|z_w|)})} = 0,\end{aligned}$$

where the last assertion stems from the fact that w is a μ-density point of E. But this is a contradiction to $z_w \in \operatorname{supp} \tau_w$, whence (3.9) follows.

Thus, for a sequence $(w_n) \subset E$ with

$$|w_n - (w + r_n z_w)| \le \operatorname{dist}(w + r_n z_w, E) + \frac{r_n}{n},$$

we have

$$\lim_{n\to\infty} \frac{|w_n - (w + r_n z_w)|}{r_n} = 0.$$

Then, with $h_n := (w_n - w)/r_n \to z_w$,

$$c_n T_*^{(w_n,r_n)}\mu = T_*^{(h_n,1)}\big(c_n T_*^{(w,r_n)}\mu\big) \xrightarrow{*} T_*^{(z_w,1)}\tau_w \qquad \text{as } n \to \infty.$$

Finally, we can select n so large that $r_n < 1/m$ and

$$F_k\big(T_*^{(z_w,1)}\tau_w, c_n T_*^{(w_n,r_n)}\mu\big) < \frac{1}{2k}.$$

Then use (3.8) for the z_{w_n} associated to $w_n \in E$ in conjunction with the preceding estimate to conclude

$$\begin{aligned}&F_k\big(T_*^{(z_{w_n},1)}\tau_{w_n}, c_n T_*^{(w_n,r_n)}\mu\big) \\ &\quad\le F_k\big(T_*^{(z_{w_n},1)}\tau_{w_n}, T_*^{(z_w,1)}\tau_w\big) + F_k\big(T_*^{(z_w,1)}\tau_w, c_n T_*^{(w_n,r_n)}\mu\big) < \frac{1}{k}.\end{aligned}$$

But this is a contradiction to $w_n \in E_{k,m}$, hence the assertion at the beginning of this step is established.

Step 2. Take $x_0 \in \mathbb{R}^d$ and $\tau \in \mathrm{Tan}(\mu, x_0)$ such that the assertion (3.7) holds. If $y_0 \notin \operatorname{supp} \tau$, the statement of the proposition is trivial, so assume $y_0 \in \operatorname{supp} \tau$. The previous step shows that $\tilde{c}_n T_*^{(y_0,\rho_n)}\tau \in \mathrm{Tan}(\mu, x_0)$ for any sequences $\rho_n \downarrow 0$, $\tilde{c}_n > 0$. Since the set of tangent measures is weakly* closed (this can be seen by a diagonal argument using the $\{F_k\}$), we immediately get that $\mathrm{Tan}(\tau, y_0) \subset \mathrm{Tan}(\mu, x_0)$. This finishes the proof. $\square$

The preceding proposition will be used as a tool to iterate the blow-up construction in order to show the existence of at least one "good" tangent measure (or later tangent Young measure, see Section 5.4). A toy example of this strategy is the proof of the following simple dichotomy:

Proposition 3.4. *Let $\mu \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^N)$. At $|\mu|$-almost every $x_0 \in \mathbb{R}^d$, either all non-zero tangent measures are purely singular (with respect to $\mathcal{L}^d$), or $P_0\mathcal{L}^d \in \mathrm{Tan}(\mu, x_0)$ with $P_0 = \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x_0)$.*

Proof. If not all non-zero tangent measures are purely singular, there exists $\tau \in \mathrm{Tan}(\mu, x_0) \setminus \{0\}$ such that its absolutely continuous part τ^a (with respect to $\mathcal{L}^d$) is non-zero. By the arguments before (3.6), this implies that there exists $y_0 \in \operatorname{supp}\tau$ and $\sigma \in \mathrm{Tan}(\tau, y_0) \setminus \{0\}$ that is a constant multiple of $\mathcal{L}^d$. The preceding proposition entails that at $|\mu|$-almost every $x_0 \in \operatorname{supp}\mu$ it holds that $\mathrm{Tan}(\tau, y_0) \subset \mathrm{Tan}(\mu, x_0)$, whereby $\sigma \in \mathrm{Tan}(\mu, x_0)$. Finally, by Proposition 3.2, σ must be of the form $\alpha P_0\mathcal{L}^d$ with P_0 as in the statement of the proposition and $\alpha \in \mathbb{R} \setminus \{0\}$. Rescaling to get $\alpha = 1$, we conclude the proof. □

Finally, we remark that for all $x_0 \in \mathbb{R}^d$ such that $\mathrm{Tan}(\mu, x_0)$ contains a non-zero measure, and any convex set $C \subset \mathbb{R}^d$ with $0 \in C$, we can always find a tangent measure $\tau \in \mathrm{Tan}(\mu, x_0)$ with $|\tau|(C) = 1$ and $|\tau|(\partial C) = 0$. This simply follows by rescaling the blow-up sequence. A similar result was given in Lemma 5.1 of [87], also see the different proof in Lemma 3.1 of [107].

3.2 Localization at regular points

In order to carry out blow-up constructions involving Young measures, we will need localization principles for these objects, one at regular and one at singular points. These results should be considered complements to the theory of tangent measures in the previous section and thus the Young measures obtained in the blow-up limit are called "tangent" Young measures.

We start with regular points.

Proposition 3.5 (Localization at regular points). *Let $\nu \in \mathbf{GY}(\Omega;\mathbb{R}^{m\times d})$ be a gradient Young measure. Then, for $\mathcal{L}^d$-almost every $x_0 \in \Omega$ there exists a* ***regular tangent Young measure*** *$\sigma \in \mathbf{GY}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^{m\times d})$ satisfying*

$$[\sigma] \in \mathrm{Tan}([\nu], x_0), \qquad \sigma_y = \nu_{x_0} \quad a.e., \tag{3.10}$$

$$\lambda_\sigma = \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0)\,\mathcal{L}^d \in \mathrm{Tan}(\lambda_\nu, x_0), \qquad \sigma_y^\infty = \nu_{x_0}^\infty \quad a.e. \tag{3.11}$$

In particular, for all bounded open sets $U \subset \mathbb{R}^d$ with $\mathcal{L}^d(\partial U) = 0$ and all $h \in \mathrm{C}(\mathbb{R}^{m\times d})$ such that the recession function h^∞ exists in the sense of (2.6), it holds that

$$\big\langle\!\big\langle \mathbb{1}_U \otimes h, \sigma \big\rangle\!\big\rangle = \left[\big\langle h, \nu_{x_0}\big\rangle + \big\langle h^\infty, \nu_{x_0}^\infty\big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0)\right]|U|. \tag{3.12}$$

Proof. Take a set $\{\varphi_k \otimes h_k\} \subset \mathbf{E}_c(\mathbb{R}^d;\mathbb{R}^{m\times d})$ determining the (local) Young measure convergence analogous to the collection from Lemma 2.16 and let $x_0 \in \mathbb{R}^d$ be as follows:

(i) There exists a sequence $r_n \downarrow 0$ such that (with $T^{(x_0,r)}(x) := (x-x_0)/r$)

$$\gamma_n := r_n^{-d} T_*^{(x_0,r_n)}[\nu] \quad \overset{*}{\rightharpoonup} \quad \frac{\mathrm{d}[\nu]}{\mathrm{d}\mathcal{L}^d}(x_0)\,\mathcal{L}^d \quad \in \mathrm{Tan}([\nu],x_0).$$

(ii) It holds that

$$\lim_{r\downarrow 0} \frac{\lambda_\nu^s(B(x_0,r))}{r^d} = 0 \qquad \text{and} \qquad \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0)\,\mathcal{L}^d \in \mathrm{Tan}(\lambda_\nu,x_0).$$

(iii) The point x_0 is an $\mathcal{L}^d$-Lebesgue point for the functions

$$x \mapsto \big\langle h_k, \nu_x \big\rangle + \big\langle h_k^\infty, \nu_x^\infty \big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x), \qquad k \in \mathbb{N}.$$

By results recalled in Sections 2.1.2, 3.1 and standard results in measure theory, the above three conditions are satisfied simultaneously at $\mathcal{L}^d$-almost every $x_0 \in \mathbb{R}^d$.

Take a BV-norm bounded generating sequence $(u_j) \subset \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ for ν, i.e. $Du_j \overset{\mathbf{Y}}{\to} \nu$ (see Proposition 2.30) and denote by $\tilde{u}_j \in \mathrm{BV}(\mathbb{R}^d;\mathbb{R}^m)$ the extension of u_j by zero onto all of $\mathbb{R}^d$. For each $n \in \mathbb{N}$ set

$$v_j^{(n)}(y) := \frac{\tilde{u}_j(x_0 + r_n y)}{r_n}, \qquad y \in \mathbb{R}^d.$$

Thus, also employing (2.1),

$$\begin{aligned} Dv_j^{(n)} &= r_n^{-d} T_*^{(x_0,r_n)} D\tilde{u}_j \\ &= \nabla u_j(x_0 + r_n \bullet)\,\mathcal{L}^d + r_n^{-1}\big(u_j(x_0+r_n\bullet)|_{\partial\Omega_n} \otimes n_{\Omega_n}\big)\,\mathcal{H}^{d-1} \llcorner \partial\Omega_n, \end{aligned}$$

where $\Omega_n := r_n^{-1}(\Omega - x_0)$, $u_j(x_0+r_n\bullet)|_{\partial\Omega_n}$ is the (inner) trace of the function $y \mapsto u_j(x_0+r_n y)$ onto $\partial\Omega_n$ and $n_{\Omega_n} \colon \partial\Omega_n \to \mathbb{S}^{d-1}$ is the unit inner normal to $\partial\Omega_n$.

We can use the previous formula together with a Poincaré inequality in BV and the boundedness of the BV-trace operator, see Section 2.6.1, to get

$$\|v_j^{(n)}\|_{\mathrm{BV}(\mathbb{R}^d;\mathbb{R}^m)} \le C(n)|Dv_j^{(n)}|(\mathbb{R}^d) = C(n)|D\tilde{u}_j|(\mathbb{R}^d) \le C(n)\|u_j\|_{\mathrm{BV}(\Omega;\mathbb{R}^m)}, \tag{3.13}$$

where $C(n)$ absorbs all n-dependent constants (including r_n^{-d}). For fixed n, this last expression is j-uniformly bounded. Hence, we may select a subsequence of the js (not explicitly labeled and depending on n) such that the sequence $(Dv_j^{(n)})_j$ generates a Young measure $\sigma^{(n)} \in \mathbf{GY}(\mathbb{R}^d)$.

For every $\varphi \otimes h \in \mathbf{E}_c(\mathbb{R}^d;\mathbb{R}^{m\times d})$ let $n \in \mathbb{N}$ be so large that $\operatorname{supp}\varphi \subset\subset \Omega_n$ (then the boundary measure in $Dv_j^{(n)}$ can be neglected), and calculate

$$\begin{aligned}\big\langle\!\big\langle \varphi \otimes h, \sigma^{(n)} \big\rangle\!\big\rangle &= \lim_{j\to\infty} \int \varphi(y) h\big(\nabla v_j^{(n)}(y)\big)\,\mathrm{d}y \\ &= \lim_{j\to\infty} \int \varphi(y) h\big(\nabla u_j(x_0 + r_n y)\big)\,\mathrm{d}y \\ &= \lim_{j\to\infty} \frac{1}{r_n^d} \int \varphi\Big(\frac{x - x_0}{r_n}\Big) h\big(\nabla u_j(x)\big)\,\mathrm{d}x \\ &= \frac{1}{r_n^d} \Big\langle\!\Big\langle \varphi\Big(\frac{\bullet - x_0}{r_n}\Big) \otimes h, \nu \Big\rangle\!\Big\rangle.\end{aligned}$$

First, we examine the regular part of the last expression:

$$\begin{aligned}&\frac{1}{r_n^d} \int \varphi\Big(\frac{x - x_0}{r_n}\Big) \Big[\big\langle h, \nu_x \big\rangle + \big\langle h^\infty, \nu_x^\infty \big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x)\Big]\,\mathrm{d}x \\ &\qquad = \int \varphi(y) \Big[\big\langle h, \nu_{x_0 + r_n y} \big\rangle + \big\langle h^\infty, \nu_{x_0 + r_n y}^\infty \big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0 + r_n y)\Big]\,\mathrm{d}y,\end{aligned}$$

which as $n \to \infty$ ($r_n \downarrow 0$) converges to

$$\int \varphi(y) \Big[\big\langle h, \nu_{x_0} \big\rangle + \big\langle h^\infty, \nu_{x_0}^\infty \big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0)\Big]\,\mathrm{d}y.$$

The latter convergence first holds for the collection of $\varphi_k \otimes h_k$ by the corresponding Lebesgue point properties of x_0, and then also for all $\varphi \otimes h \in \mathbf{E}_c(\mathbb{R}^d;\mathbb{R}^{m\times d})$ by density.

For the singular part, let $N > 0$ be so large that $\operatorname{supp}\varphi \subset B(0,N)$ and observe by virtue of assumption (ii) on x_0 that as $n \to \infty$,

$$\bigg| \frac{1}{r_n^d} \int \varphi\Big(\frac{x - x_0}{r_n}\Big) \big\langle h^\infty, \nu_x^\infty \big\rangle\,\mathrm{d}\lambda_\nu^s(x) \bigg| \le M \|\varphi\|_\infty \cdot \frac{\lambda_\nu^s(B(x_0, N r_n))}{r_n^d} \quad \to \quad 0,$$

where $M := \sup\{\, |h^\infty(A)| \;:\; A \in \partial\mathbb{B}^{m\times d} \,\}$.

In particular, we have proved so far that

$$\sup_{n\in\mathbb{N}} \big| \big\langle\!\big\langle \varphi \otimes |\bullet|, \sigma^{(n)} \big\rangle\!\big\rangle \big| < \infty \qquad \text{for all } \varphi \in \mathrm{C}_c(\mathbb{R}^d).$$

Thus, by the Young measure compactness result in Corollary 2.9, selecting a further subsequence if necessary, we may assume that $\sigma^{(n)} \overset{*}{\to} \sigma$ for a Young measure $\sigma \in \mathbf{Y}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^{m\times d})$. From a diagonal argument we get that in fact $\sigma \in \mathbf{GY}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^{m\times d})$. We also have $[\sigma] \in \operatorname{Tan}([\nu], x_0)$ because $[\sigma^{(n)}] = \gamma_n$ (from assumption (i) on x_0) plus a jump part that moves out to infinity in the limit. This proves the first assertion in (3.10).

Our previous considerations yield

$$\big\langle\!\big\langle \varphi \otimes h, \sigma \big\rangle\!\big\rangle = \int \varphi(y) \Big[\big\langle h, \nu_{x_0} \big\rangle + \big\langle h^\infty, \nu_{x_0}^\infty \big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0)\Big]\,\mathrm{d}y$$

for all $\varphi \otimes h \in \mathbf{E}_c(\mathbb{R}^d;\mathbb{R}^{m\times d})$. Varying first φ and then h, we see that $\sigma_y = \nu_{x_0}$ and $\sigma_y^\infty = \nu_{x_0}^\infty$ holds for $\mathcal{L}^d$-almost every $y \in \mathbb{R}^d$, i.e. the second assertions in (3.10) and (3.11), respectively.

The first assertion of (3.11) follows since the previous formula also implies $\lambda_\sigma = \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0)\,\mathcal{L}^d$ and the latter measure lies in $\mathrm{Tan}(\lambda_\nu, x_0)$ by (ii). Finally, as an immediate consequence of (3.10) and (3.11) in conjunction with Proposition 2.17, we get (3.12). This concludes the proof. □

3.3 Localization at singular points

We now turn to singular points, i.e. points in the support of the singular part of the concentration measure λ_ν of a Young measure ν.

Proposition 3.6 (Localization at singular points). *Let $\nu \in \mathbf{GY}(\Omega;\mathbb{R}^{m\times d})$ be a gradient Young measure. Then, for λ_ν^s-almost every $x_0 \in \Omega$, there exists a **singular tangent Young measure** $\sigma \in \mathbf{GY}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^{m\times d})$ satisfying*

$$[\sigma] \in \mathrm{Tan}([\nu], x_0), \qquad \sigma_y = \delta_0 \quad a.e., \tag{3.14}$$

$$\lambda_\sigma \in \mathrm{Tan}(\lambda_\nu^s, x_0) \setminus \{0\}, \qquad \sigma_y^\infty = \nu_{x_0}^\infty \quad \lambda_\sigma\text{-}a.e. \tag{3.15}$$

In particular, for all bounded open sets $U \subset \mathbb{R}^d$ with $(\mathcal{L}^d + \lambda_\sigma)(\partial U) = 0$ and all positively 1-homogeneous $g \in \mathrm{C}(\mathbb{R}^{m\times d})$ it holds that

$$\big\langle\!\big\langle \mathbb{1}_U \otimes g, \sigma \big\rangle\!\big\rangle = \big\langle g, \nu_{x_0}^\infty \big\rangle\, \lambda_\sigma(U). \tag{3.16}$$

Proof. Take a dense and countable set $\{g_k\} \subset \mathrm{C}(\partial\mathbb{B}^{m\times d})$ and consider all g_k to be extended to $\mathbb{R}^{m\times d}$ by positive 1-homogeneity. Then, let $x_0 \in \operatorname{supp}\lambda_\nu^s$ be such that:

(i) 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

$$c_n T_*^{(x_0,r_n)} \lambda_\nu^s \overset{*}{\rightharpoonup} \lambda_0. \tag{3.17}$$

(ii) It holds that

$$\lim_{r\downarrow 0} \frac{1}{\lambda_\nu^s(B(x_0,r))} \int_{B(x_0,r)} 1 + \big\langle |\boldsymbol{\cdot}|, \nu_x \big\rangle + \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x)\,\mathrm{d}x = 0. \tag{3.18}$$

(iii) The point x_0 is a λ_ν^s-Lebesgue point for the functions

$$x \mapsto \big\langle \mathrm{id}, \nu_x^\infty \big\rangle \qquad \text{and} \qquad x \mapsto \big\langle g_k, \nu_x^\infty \big\rangle, \quad k \in \mathbb{N}.$$

By the usual measure-theoretic results and Preiss's existence theorem for non-zero tangent measures, see Lemma 3.1, this can be achieved at λ_ν^s-almost every $x_0 \in \Omega$.

The constants c_n in (3.17) can always be chosen as

$$c_n = c\big[\lambda_\nu^s(\overline{B(x_0, Rr_n)})\big]^{-1} \qquad (< \infty)$$

for some fixed $R > 0$, $c > 0$, such that $\lambda_0(B(0,R)) > 0$, see (3.3). Also notice that we may increase R such that $\lambda_\nu^s(\partial B(x_0, Rr_n)) = 0$ for all n. In conjunction with (3.17) this further yields for each $n \in \mathbb{N}$ the existence of a constant $\beta_N > 0$ satisfying

$$\limsup_{n\to\infty} c \cdot \frac{\lambda_\nu^s(B(x_0, Nr_n))}{\lambda_\nu^s(B(x_0, Rr_n))} \le \beta_N.$$

Combining this with (3.18), we get

$$\begin{aligned}
&\limsup_{n\to\infty} c_n \big\langle\!\big\langle \mathbb{1}_{B(x_0,Nr_n)} \otimes |\cdot|, \nu \big\rangle\!\big\rangle \\
&\qquad = \limsup_{n\to\infty} \Bigg[\frac{c}{\lambda_\nu^s(B(x_0, Rr_n))} \int_{B(x_0,Nr_n)} \big\langle |\cdot|, \nu_x \big\rangle + \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x) \,\mathrm{d}x \\
&\qquad\qquad\qquad + \frac{c}{\lambda_\nu^s(B(x_0, Rr_n))}\, \lambda_\nu^s(B(x_0, Nr_n)) \Bigg] \\
&\qquad \le 0 + \beta_N.
\end{aligned}$$

Hence, for all $N \in \mathbb{N}$,

$$\limsup_{n\to\infty} c_n \lambda_\nu^s(B(x_0, Nr_n)) \le \limsup_{n\to\infty} c_n \big\langle\!\big\langle \mathbb{1}_{B(x_0,Nr_n)} \otimes |\cdot|, \nu \big\rangle\!\big\rangle \le \beta_N \tag{3.19}$$

Furthermore,

$$\limsup_{n\to\infty} \big(c_n T_*^{(x_0,r_n)} |[\nu]|\big)(B(0,N)) \le \beta_N, \qquad \text{for all } N \in \mathbb{N},$$

and so, taking a (non-relabeled) subsequence of the r_n, we may assume

$$c_n T_*^{(x_0,r_n)}[\nu] \overset{*}{\rightharpoonup} \tau \quad \in \mathrm{Tan}([\nu], x_0). \tag{3.20}$$

Notice that τ might be the zero measure ($\tau \neq 0$ could only be ensured for $[\nu]$-almost every $x_0 \in \operatorname{supp}[\nu]$, but not necessarily for λ_ν^s-almost every $x_0 \in \Omega$).

For a norm-bounded generating sequence $(u_j) \subset \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ of ν, that is $Du_j \overset{\mathbf{Y}}{\to} \nu$, we denote by $\tilde{u} \in \mathrm{BV}(\mathbb{R}^d;\mathbb{R}^m)$ the extension by zero, and set

$$v_j^{(n)}(y) := r_n^{d-1} c_n \tilde{u}_j(x_0 + r_n y), \qquad y \in \mathbb{R}^d.$$

We can then compute

$$\begin{aligned}
Dv_j^{(n)} &= c_n T_*^{(x_0,r_n)} D\tilde{u}_j \\
&= r_n^d c_n \nabla u_j(x_0 + r_n \cdot)\, \mathcal{L}^d + r_n^{d-1} c_n \big(u_j(x_0 + r_n \cdot)|_{\partial\Omega_n} \otimes n_{\Omega_n}\big)\, \mathcal{H}^{d-1} \llcorner \Omega_n,
\end{aligned}$$

where as before $\Omega_n := r_n^{-1}(\partial\Omega - x_0)$. In complete analogy to (3.13) we may also derive

$$\|v_j^{(n)}\|_{\mathrm{BV}(\mathbb{R}^d;\mathbb{R}^m)} \le C(n) \|u_j\|_{\mathrm{BV}(\Omega;\mathbb{R}^m)}.$$

The latter estimate implies that up to an n-dependent subsequence of js, $(Dv_j^{(n)})_j$ generates a Young measure $\sigma^{(n)} \in \mathbf{GY}(\mathbb{R}^d)$.

Let $g \in \mathrm{C}(\mathbb{R}^{m\times d})$ be positively 1-homogeneous and let $\varphi \in \mathrm{C}_c(\mathbb{R}^d)$. Then we have for all n so large that $\operatorname{supp}\varphi \subset\subset \Omega_n$ (and hence we may neglect the boundary jump part of $Dv_j^{(n)}$),

$$\begin{aligned}
\big\langle\!\big\langle \varphi \otimes g, \sigma^{(n)} \big\rangle\!\big\rangle &= \lim_{j\to\infty} \int \varphi(y) g\big(\nabla v_j^{(n)}(y)\big) \;\mathrm{d}y \\
&= \lim_{j\to\infty} r_n^d c_n \int \varphi(y) g\big(\nabla u_j(x_0 + r_n y)\big) \;\mathrm{d}y \\
&= \lim_{j\to\infty} c_n \int \varphi\Big(\frac{x - x_0}{r_n}\Big) g\big(\nabla u_j(x)\big) \;\mathrm{d}y \\
&= c_n \Big\langle\!\Big\langle \varphi\Big(\frac{\bullet - x_0}{r_n}\Big) \otimes g, \nu \Big\rangle\!\Big\rangle.
\end{aligned} \tag{3.21}$$

For the regular part of the last expression, set $M := \sup\{\, |g(A)| \;\mathbf{:}\; A \in \partial\mathbb{B}^{m\times d} \,\}$ and choose $N \in \mathbb{N}$ so large that $\operatorname{supp}\varphi \subset B(0, N)$. Possibly increasing n as to ensure

$$c_n \lambda_\nu^s(B(x_0, N r_n)) \leq \beta_N + 1,$$

see (3.19), we have

$$\begin{aligned}
&\bigg| c_n \int \varphi\Big(\frac{x - x_0}{r_n}\Big) \Big[\big\langle g, \nu_x \big\rangle + \big\langle g, \nu_x^\infty \big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x) \Big] \;\mathrm{d}x \bigg| \\
&\qquad \leq c_n M \|\varphi\|_\infty \int_{B(x_0, N r_n)} \big\langle |\bullet|, \nu_x \big\rangle + \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x) \;\mathrm{d}x \\
&\qquad \leq \frac{M \|\varphi\|_\infty (\beta_N + 1)}{\lambda_\nu^s(B(x_0, N r_n))} \int_{B(x_0, N r_n)} \big\langle |\bullet|, \nu_x \big\rangle + \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x) \;\mathrm{d}x \\
&\qquad \to 0 \qquad \text{as } n \to \infty,
\end{aligned} \tag{3.22}$$

the convergence following by virtue of (3.18). Hence, we get from (3.21),

$$\limsup_{n\to\infty} \big\langle\!\big\langle \varphi \otimes g, \sigma^{(n)} \big\rangle\!\big\rangle = \limsup_{n\to\infty} c_n \int \varphi\Big(\frac{x - x_0}{r_n}\Big) \big\langle g, \nu_x^\infty \big\rangle \;\mathrm{d}\lambda_\nu^s(x). \tag{3.23}$$

Taking $g = |\bullet|$ in the previous equality,

$$\begin{aligned}
\limsup_{n\to\infty} \big\langle\!\big\langle \varphi \otimes |\bullet|, \sigma^{(n)} \big\rangle\!\big\rangle &= \limsup_{n\to\infty} c_n \int \varphi\Big(\frac{x - x_0}{r_n}\Big) \;\mathrm{d}\lambda_\nu^s(x) \\
&= \limsup_{n\to\infty} \int \varphi \;\mathrm{d}\big(c_n T_*^{(x_0, r_n)} \lambda_\nu^s\big) = \int \varphi \;\mathrm{d}\lambda_0,
\end{aligned}$$

where the convergence follows from (3.17). In particular, $\langle\!\langle \varphi \otimes |\bullet|, \sigma^{(n)} \rangle\!\rangle$ is asymptotically as $n \to \infty$ bounded by $\|\varphi\|_\infty \lambda_0(\operatorname{supp}\varphi)$, and hence by the Young measure compactness there exists a subsequence of the r_ns (not relabeled) with

$$\sigma^{(n)} \overset{*}{\rightharpoonup} \sigma \qquad \text{in } \mathbf{Y}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^{m\times d}).$$

Again by a diagonal argument, we see $\sigma \in \mathbf{GY}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^{m\times d})$. From (3.23) we also get

$$\big\langle\!\big\langle \varphi \otimes g, \sigma \big\rangle\!\big\rangle = \lim_{n\to\infty} c_n \int \varphi\Big(\frac{x - x_0}{r_n}\Big) \big\langle g, \nu_x^\infty \big\rangle \;\mathrm{d}\lambda_\nu^s(x). \tag{3.24}$$

We now turn to the verification of (3.14), (3.15), and (3.16). The barycenters of the $\sigma^{(n)}$ satisfy

$$[\sigma^{(n)}] = c_n T_*^{(x_0,r_n)}[\nu] + \mu_n,$$

where $\mu_n \in \mathbf{M}(\partial\Omega_n;\mathbb{R}^{m\times d})$ are boundary measures satisfying $\mu_n \overset{*}{\rightharpoondown} 0$. Hence, by (3.20), $[\sigma^{(n)}] \overset{*}{\rightharpoondown} \tau$ as $n \to \infty$ and so $[\sigma] = \tau \in \mathrm{Tan}([\nu], x_0)$, which is the first assertion of (3.14).

For the second assertion of (3.14), we take cut-off functions $\varphi \in \mathrm{C}_c(\mathbb{R}^d;[0,1])$, $\chi \in \mathrm{C}_c(\mathbb{R}^{m\times d};[0,1])$ and calculate similarly to (3.21),

$$\begin{aligned}\left\langle\!\left\langle \varphi \otimes |\cdot|\chi(\cdot), \sigma^{(n)} \right\rangle\!\right\rangle &= \lim_{j\to\infty} c_n \int \varphi\Big(\frac{x-x_0}{r_n}\Big)|\nabla u_j(x)|\chi\big(r_n^d c_n \nabla u_j(x)\big)\,\mathrm{d}x \\ &= c_n \left\langle\!\!\left\langle \varphi\Big(\frac{\cdot - x_0}{r_n}\Big) \otimes |\cdot|\chi(r_n^d c_n \cdot), \nu \right\rangle\!\!\right\rangle.\end{aligned}$$

Then use a reasoning analogous to (3.22) to see that the regular part of the previous expression converges to zero as $n \to \infty$. On the other hand, because χ has compact support in $\mathbb{R}^{m\times d}$, the singular part is identically zero. So we have shown

$$\left\langle\!\left\langle \varphi \otimes |\cdot|\chi(\cdot), \sigma \right\rangle\!\right\rangle = 0$$

for all φ, χ as above. Hence, $\sigma_y = \delta_0$ for $\mathcal{L}^d$-almost every $y \in \mathbb{R}^d$.

To see the first assertion from (3.15), plug $g := |\cdot|$ into (3.24) and use $\sigma_y = \delta_0$ a.e. to derive for any $\varphi \in \mathrm{C}_c(\mathbb{R}^d)$,

$$\begin{aligned}\int \varphi \,\mathrm{d}\lambda_\sigma = \left\langle\!\left\langle \varphi \otimes |\cdot|, \sigma \right\rangle\!\right\rangle &= \lim_{n\to\infty} c_n \int \varphi\Big(\frac{x-x_0}{r_n}\Big)\,\mathrm{d}\lambda_\nu^s(x) \\ &= \lim_{n\to\infty} \int \varphi \,\mathrm{d}\big(c_n T_*^{(x_0,r_n)}\lambda_\nu^s\big) = \int \varphi \,\mathrm{d}\lambda_0,\end{aligned}$$

the last equality by (3.17). Hence, $\lambda_\sigma = \lambda_0 \in \mathrm{Tan}(\lambda_\nu^s, x_0)$.

We postpone the proof of the second assertion of (3.15) for a moment and instead turn to the verification of (3.16) first. Let $U \subset \mathbb{R}^d$ be a bounded open set with $(\mathcal{L}^d+\lambda_\sigma)(\partial U) = 0$. If $\lambda_\sigma(U) = 0$, then (3.16) holds trivially, so assume $\lambda_\sigma(U) > 0$. Use $\varphi = \mathbb{1}_U$ in (3.24), which is allowed by virtue of Proposition 2.17, to get

$$\int_U \big\langle g, \sigma_y^\infty \big\rangle \,\mathrm{d}\lambda_\sigma(y) = \left\langle\!\left\langle \mathbb{1}_U \otimes g, \sigma \right\rangle\!\right\rangle = \lim_{n\to\infty} c_n \int_{x_0+r_nU} \big\langle g, \nu_x^\infty \big\rangle \,\mathrm{d}\lambda_\nu^s(x).$$

Because $\lambda_\sigma \in \mathrm{Tan}(\lambda_\nu^s, x_0)$ and $\lambda_\sigma(U) > 0$, by (3.3) in Section 3.1 we infer that $c_n = \tilde{c}(U)[\lambda_\nu^s(x_0 + r_nU)]^{-1}$ for some constant $\tilde{c}(U) > 0$. With this, the right hand side is

$$\begin{aligned}\lim_{n\to\infty} c_n \int_{x_0+r_nU} \big\langle g, \nu_x^\infty \big\rangle \,\mathrm{d}\lambda_\nu^s(x) &= \lim_{n\to\infty} \tilde{c}(U) \fint_{x_0+r_nU} \big\langle g, \nu_x^\infty \big\rangle \,\mathrm{d}\lambda_\nu^s(x) \\ &= \tilde{c}(U)\,\big\langle g, \nu_{x_0}^\infty \big\rangle\end{aligned}$$

by the Lebesgue point properties of x_0 (first ascertain this for the collection $\{g_k\}$ and then for the general case). Hence we have shown

$$\left\langle\!\left\langle \mathbb{1}_U \otimes g, \sigma \right\rangle\!\right\rangle = \int_U \big\langle g, \sigma_y^\infty \big\rangle \,\mathrm{d}\lambda_\sigma(y) = \tilde{c}(U)\,\big\langle g, \nu_{x_0}^\infty \big\rangle$$

and testing this with $g = |\bullet|$, we get $\tilde{c}(U) = \lambda_\sigma(U)$. Thus we have proved (3.16). But clearly, varying U and g, this also implies $\sigma_y^\infty = \nu_{x_0}^\infty$ for λ_σ-almost every $y \in \mathbb{R}^d$, which is the second assertion of (3.15). $\square$

Chapter 4

Lower semicontinuity in BV

This chapter is devoted to the new proof from [107] of (a slightly generalized version of) the classical lower semicontinuity theorem for integral functionals on the space $\mathrm{BV}(\Omega;\mathbb{R}^m)$, first established by Ambrosio & Dal Maso [8] and Fonseca & Müller [57]:

Theorem 4.1. *Let $\Omega \subset \mathbb{R}^d$ be a bounded open Lipschitz domain and let $f \in \mathbf{R}(\Omega;\mathbb{R}^{m\times d})$ be a quasiconvex integrand, that is, $f\colon \overline{\Omega} \times \mathbb{R}^{m\times d} \to \mathbb{R}$ satisfies the following assumptions:*

(i) *f is a Carathéodory function,*

(ii) *$|f(x,A)| \le M(1+|A|)$ for some $M > 0$ and all $x \in \overline{\Omega}$, $A \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$,*

(iii) *the (strong) recession function f^∞ exists in the sense of* (2.6) *and is (jointly) continuous on $\overline{\Omega} \times \mathbb{R}^{m\times d}$,*

(iv) *$f(x,\,\bullet\,)$ is quasiconvex for all $x \in \overline{\Omega}$.*

Then, the functional

$$\begin{aligned}\mathcal{F}(u) := \int_\Omega f\bigl(x,\nabla u(x)\bigr)\,\mathrm{d}x + \int_\Omega f^\infty\left(x,\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\right)\mathrm{d}|D^s u|(x)\\ + \int_{\partial\Omega} f^\infty\bigl(x,u|_{\partial\Omega}(x)\otimes n_\Omega(x)\bigr)\,\mathrm{d}\mathcal{H}^{d-1}(x), \qquad u \in \mathrm{BV}(\Omega;\mathbb{R}^m),\end{aligned}$$

is lower semicontinuous with respect to weak convergence in the space* $\mathrm{BV}(\Omega;\mathbb{R}^m)$.

An outline of the proof was already given in the introduction.

4.1 Integral functionals and quasiconvexity

We first record the following proposition, which tells us that we are indeed considering the correct BV-extension of our functional.

Proposition 4.2. *Let $f \in \mathbf{R}(\Omega;\mathbb{R}^{m\times d})$ (not necessarily quasiconvex). Then, $\mathcal{F}$ from Theorem 4.1 is the $\langle\bullet\rangle$-strictly continuous extension of the functional*

$$\int_\Omega f\big(x,\nabla u(x)\big)\,\mathrm{d}x + \int_{\partial\Omega} f^\infty\big(x,u|_{\partial\Omega}(x)\otimes n_\Omega(x)\big)\,\mathrm{d}\mathcal{H}^{d-1}(x), \qquad u \in \mathrm{W}^{1,1}(\Omega;\mathbb{R}^m),$$

onto the space $\mathrm{BV}(\Omega;\mathbb{R}^m)$.

Proof. By Lemma 2.24, the space $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^m)$ is $\langle\bullet\rangle$-strictly dense in $\mathrm{BV}(\Omega;\mathbb{R}^m)$, whence we may deduce the result from Proposition 2.26 (a corollary to Reshetnyak's Continuity Theorem 2.3) in conjunction with Proposition 2.17 (i) on the extended representation of Young measure limits (for integrands $f \in \mathbf{R}(\Omega;\mathbb{R}^{m\times d})$). □

As the decisive convexity property of the integrand, Morrey [93, 94] introduced the following notion: A locally bounded Borel function $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ is called **quasiconvex** if

$$h(A) \le \fint_\omega h\big(A+\nabla\psi(z)\big)\,\mathrm{d}z \qquad \text{for all } A \in \mathbb{R}^{m\times d} \text{ and all } \psi \in \mathrm{W}_0^{1,\infty}(\omega;\mathbb{R}^m),$$

where $\omega \subset \mathbb{R}^d$ is an arbitrary bounded open Lipschitz domain. This definition does not depend on the particular choice of the open Lipschitz domain ω (by a covering argument) and it can be shown that quasiconvex functions are **rank-one convex**, meaning that they are convex along rank-one lines. Moreover, if h has linear growth at infinity, the requirement that $\psi \in \mathrm{W}_0^{1,\infty}$ may equivalently be replaced by $\psi \in \mathrm{W}_0^{1,1}(\omega;\mathbb{R}^m)$, cf. [20]. See [34, 94] for more on quasiconvexity.

The **quasiconvex envelope** Qh of a continuous function $h\colon \mathbb{R}^{m\times d} \to \mathbb{R}$ is

$$Qh = \sup\big\{\, g \;:\; g \text{ quasiconvex and } g \le h \,\big\},$$

which is either identically $-\infty$ or real-valued and quasiconvex. It is well-known (cf. [34,75]) that for continuous h the quasiconvex envelope can be written via Dacorogna's formula as

$$Qh(A) := \inf\left\{ \fint_\omega h\big(A+\nabla\psi(z)\big)\,\mathrm{d}z \;:\; \psi \in \mathrm{W}_0^{1,\infty}(\omega;\mathbb{R}^m) \right\} \tag{4.1}$$

and this formula is again independent of ω. Note that the equality also holds if one of the two expressions is $-\infty$. If h has linear growth at infinity, we may again replace the space $\mathrm{W}_0^{1,\infty}(\omega;\mathbb{R}^m)$ in the above formula with $\mathrm{W}_0^{1,1}(\omega;\mathbb{R}^m)$.

The next lemma introduces a class of non-trivial quasiconvex functions with p-growth, $p \in [1,\infty)$ (including the linear growth case), based on a construction of Šverák [113] and some ellipticity arguments similar to those in Lemma 2.7 of [96].

Lemma 4.3. *Let $M \in \mathbb{R}^{2\times 2}$ with $\operatorname{rank} M = 2$ and let $p \in [1,\infty)$. Define*

$$h(A) := \operatorname{dist}(A,\{-M,M\})^p, \qquad A \in \mathbb{R}^{2\times 2}.$$

Then, $h(0) > 0$ and the quasiconvex envelope Qh is not convex (at zero). Moreover, h has a (strong) p-recession function h^∞, defined via $h^\infty(A) := \lim_{A'\to A,t\to\infty} t^{-p}h(tA')$ for $A \in \mathbb{R}^{2\times 2}$, and it holds that $(Qh)^\infty = h^\infty = |\bullet|^p$.

Proof. We first show $h(0) > 0$. Assume to the contrary that $Qh(0) = 0$. Then by (4.1) there exists a sequence $(\psi_j) \subset \mathrm{W}_0^{1,\infty}(\mathbb{B}^2;\mathbb{R}^2)$ with

$$\fint_{\mathbb{B}^2} h(\nabla\psi_j)\,\mathrm{d}z \to 0. \tag{4.2}$$

Let $P\colon \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$ be the projection onto the orthogonal complement of $\operatorname{span}\{M\}$. Some linear algebra shows $|P(A)|^p \le h(A)$ for all $A \in \mathbb{R}^{2\times 2}$. Therefore,

$$P(\nabla\psi_j) \to 0 \quad \text{in } \mathrm{L}^p(\mathbb{B}^2;\mathbb{R}^{2\times 2}). \tag{4.3}$$

With the definition

$$\hat{g}(\xi) := \int_{\mathbb{B}^d} g(x)\mathrm{e}^{2\pi\mathrm{i}\xi\cdot x}\,\mathrm{d}x, \qquad \xi \in \mathbb{R}^d,$$

for the Fourier transform $\hat{g}$ of the function $g \in \mathrm{L}^1(\mathbb{B}^d;\mathbb{R}^N)$, we get $\widehat{\nabla\psi_j}(\xi) = (2\pi\mathrm{i})\,\hat{\psi}_j(\xi)\otimes\xi$. The fact that $a\otimes\xi \notin \operatorname{span}\{M\}$ for all $a,\xi \in \mathbb{R}^2\setminus\{0\}$ implies $P(a\otimes\xi) \neq 0$. We may write

$$\widehat{\nabla\psi_j}(\xi) = M(\xi)P(\hat{\psi}_j(\xi)\otimes\xi), \qquad \xi \in \mathbb{R}^d,$$

where $M(\xi)\colon \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$ is linear and depends smoothly and positively 0-homogeneously on ξ. We omit the details of this construction here, see Section 5.4.1 for an analogous argument. Applying the Mihlin Multiplier Theorem (see for example Theorem 5.2.7 in [68] and Theorem 6.1.6 in [24]) first for $p > 1$ and combine with (4.3) to get

$$\nabla\psi_j \to 0 \quad \text{in } \mathrm{L}^p(\mathbb{B}^2;\mathbb{R}^{2\times 2}),$$

contradicting (4.2).

If $p = 1$, we get from the Mihlin Multiplier Theorem that only

$$\{\, x \in \mathbb{B}^2 \;:\; |\nabla\psi_j|(x) \ge \alpha \,\} \to 0 \qquad \text{for all } \alpha > 0. \tag{4.4}$$

Then,

$$\begin{aligned}\int_{\{|\nabla\psi_j|\ge\alpha\}} |\nabla\psi_j|\,\mathrm{d}z &\le \int_{\{|\nabla\psi_j|\ge\alpha\}} h(\nabla\psi_j)\,\mathrm{d}z + |M|\cdot\{\, x\in\mathbb{B}^2 \;:\; |\nabla\psi_j|(x)\ge\alpha \,\}\\ &\to 0.\end{aligned}$$

Since from (4.4) we may, after selecting a non-relabeled subsequence if necessary, assume that also $\nabla\psi_j \to 0$ pointwise almost everywhere, we can conclude that

$$\int_{\mathbb{B}^2} |\nabla\psi_j|\,\mathrm{d}z \le \int_{\{|\nabla\psi_j|\ge\alpha\}} |\nabla\psi_j|\,\mathrm{d}z + \int_{\{|\nabla\psi_j|\le\alpha\}} |\nabla\psi_j|\,\mathrm{d}z \to 0,$$

which implies that $\nabla\psi_j \to 0$ in $\mathrm{L}^1(\mathbb{R}^2;\mathbb{R}^2)$, contradicting (4.2). Hence we have shown $Qh(0) > 0$ for all $p \in [1,\infty)$.

Finally, if Qh was convex at zero, we would have

$$Qh(0) \le \frac{1}{2}\big(Qh(M) + Qh(-M)\big) \le \frac{1}{2}\big(h(M) + h(-M)\big) = 0,$$

a contradiction. For the addition it suffices to notice that there exists a constant $C > 0$ such that

$$|A|^p - C \le h(A) \le |A|^p + C \qquad \text{for all } A \in \mathbb{R}^{m\times d}.$$

This concludes the proof. □

We refer to [129] for another construction of quasiconvex functions with linear growth and to [95] for an example of a non-convex quasiconvex function that is even positively 1-homogeneous. Also, the non-constructive result in Theorem 8.1 of [81] shows that "many" quasiconvex functions with linear growth must exist.

As it turns out, not every quasiconvex function with linear growth at infinity has a recession function h^∞ in the sense of (2.6) (see Theorem 2 in [95] for a counterexample), so we have to use the generalized recession function $h^\#$ as defined in (2.7) instead. However, we remark that by the rank-one convexity one may replace the limes superior in the definition of $h^\#$ by a proper limit for all matrices of rank one. Notice also that under the assumption of linear growth it follows from Fatou's Lemma that the generalized recession function $h^\#$ is quasiconvex whenever h is, see for example [10], pp. 303–304; the same applies to h^∞ if it exists. It is shown in Morrey's book [94] that quasiconvex functions with linear growth are Lipschitz continuous (see Lemma 2.2 of [19] for an explicit Lipschitz constant), hence we may use the simpler definition (2.8) for the recession functions.

4.2 Rigidity

In this section we establish that functions $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$ with the property that $Du = P|Du|$, where $P \in \mathbb{R}^{m\times d}$ is a fixed matrix, have a very special structure. The origins of this observation can be traced back to Hadamard's jump condition, Proposition 2 in [18], and Lemma 1.4 of [38]; also see the proof of Theorem 3.95 in [10]. Other rigidity results may be found in [77, 78, 96] and the references cited therein.

Lemma 4.4 (Rigidity of BV-functions). *Let $C \subset \mathbb{R}^d$ be open and convex (not necessarily bounded) and let $u \in \mathrm{BV}_{\mathrm{loc}}(C;\mathbb{R}^m)$ such that for a fixed matrix $P \in \mathbb{R}^{m\times d}$ with $|P| = 1$ it holds that*

$$Du = P|Du| \qquad \textit{or, equivalently,} \qquad \frac{\mathrm{d}Du}{\mathrm{d}|Du|} = P \quad |Du|\text{-a.e.}$$

Then:

(i) If $\operatorname{rank} P \ge 2$, then $u(x) = u_0 + \alpha P x$ a.e., where $\alpha \in \mathbb{R}$, $u_0 \in \mathbb{R}^m$.

(ii) If $P = a \otimes \xi$ ($a \in \mathbb{R}^m$, $\xi \in \mathbb{S}^{d-1}$), then there exist $\psi \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$, $u_0 \in \mathbb{R}^m$ such that $u(x) = u_0 + \psi(x \cdot \xi)a$ a.e..

Proof. First assume $u \in (\mathrm{W}^{1,1}_{\mathrm{loc}} \cap \mathrm{C}^\infty)(C;\mathbb{R}^m)$. The idea of the proof is that the curl of ∇u vanishes, i.e.

$$\partial_i(\nabla u)^k_j = \partial_j(\nabla u)^k_i \qquad \text{for all } i,j = 1,\dots,d \text{ and } k = 1,\dots,m.$$

For our special $\nabla u = Pg$, where $g \in \mathrm{C}^\infty(C)$ is a smooth function, this gives the conditions

$$P^k_j \partial_i g = P^k_i \partial_j g \qquad \text{for all } i,j = 1,\dots,d \text{ and } k = 1,\dots,m. \tag{4.5}$$

Under the assumptions of (i), we claim that $\nabla g \equiv 0$. If otherwise $\xi(x) := \nabla g(x) \neq 0$ for an $x \in C$, then with $a_k(x) := P^k_j/\xi_j(x)$ ($k = 1,\dots,m$) for any j such that $\xi_j(x) \neq 0$ (the quantity $a_k(x)$ is well-defined by the relation (4.5)), we have $P^k_j = a_k(x)\xi_j(x)$, which immediately implies $P = a(x) \otimes \xi(x)$. This, however, is impossible if $\operatorname{rank} P \geq 2$. Hence, $\nabla g \equiv 0$ and u is an affine function, which must be of the form exhibited in assertion (i).

For part (ii), that is $P = a \otimes \xi$, we additionally assume $a \neq 0$. Observe that in this case $\nabla g(x) = \theta(x)\xi^T$ for some function $\theta \in \mathrm{C}^\infty(C)$. Indeed, (4.5) entails

$$\xi \nabla g(x) = \nabla g(x)^T \xi^T \in \mathbb{R}^{d\times d}, \qquad x \in C,$$

which gives the projection relation

$$\nabla g(x) = (\xi \cdot \nabla g(x)^T)\xi^T =: \theta(x)\xi^T.$$

Next, since level sets of a function are always orthogonal to the function's gradient, we infer that g is constant on all hyperplanes orthogonal to ξ intersected with C. As C is assumed convex, we may therefore write

$$g(x) = \tilde{g}(x \cdot \xi), \qquad x \in C$$

with $\tilde{g} \in \mathrm{C}^\infty(\mathbb{R})$. Hence, $\nabla u(x) = Pg(x) = (a \otimes \xi)\tilde{g}(x \cdot \xi)$, and so $u(x) = u_0 + \psi(x \cdot \xi)a$ for some $\psi \in \mathrm{C}^\infty(\mathbb{R})$ with $\psi' = \tilde{g}$ and $u_0 \in \mathbb{R}^m$.

For general u as in the statement of the proposition, we employ a mollification argument as follows: Consider a convex subdomain $C' \subset\subset C$ and mollify the original u by a smooth kernel with support inside $B(0,d)$, where $d > 0$ is the distance from C' to $\mathbb{R}^d \setminus C$. Then, in C' this yields a smooth function $\tilde{u} \in (\mathrm{W}^{1,1}_{\mathrm{loc}} \cap \mathrm{C}^\infty)(C';\mathbb{R}^m)$, which still satisfies $D\tilde{u} = P_0|D\tilde{u}|$, and we can apply the above reasoning to that function. Since C' was arbitrary, we conclude the proof. □

Remark 4.5. Statement (ii) can also be proved in a slightly different fashion: Let $P = a \otimes \xi$ and pick any $b \perp \xi$. Then,

$$\partial_t u(x + tb) = \nabla u(x)b = [a\xi^T b]g(x) = 0.$$

This implies that u is constant in direction b. As $b \perp \xi$ was arbitrary and Ω is assumed convex, $u(x)$ can only depend on $x \cdot \xi$.

Remark 4.6. The elementary argument of the last remark can also be carried out directly on the level of BV-functions using the theory of one-dimensional sections (see for example Section 3.11 of [10]). Let $P = a \otimes \xi$ and $b \perp \xi$. Then, the theory of sections implies that (the scalar product here is to be taken row-wise)

$$Du \cdot b = \mathcal{L}^{d-1} \llcorner \Omega_b \otimes Du_y^b = \int_{\Omega_b} Du_y^b \,\mathrm{d}y,$$

where Ω_b is the orthogonal projection of Ω onto the hyperplane orthogonal to b and $u_y^b(t) := u(y + tb)$ for any $y \in \Omega_b$ and $t \in \mathbb{R}$ such that $y + tb \in \Omega$. Also,

$$|Du \cdot b| = \mathcal{L}^{d-1} \llcorner \Omega_b \otimes |Du_y^b|.$$

For $Du = P|Du|$ we have $|Du \cdot b| = |a\xi^T b||Du| = 0$ and hence $|Du_y^b| = 0$ for almost every $y \in \Omega_b$. But this implies that u is constant in direction b. As $b \perp \xi$ was arbitrary, $u(x)$ can only depend on $x \cdot \xi$ and we have shown the claim.

Remark 4.7 (Differential inclusions). Restating the preceding lemma, we have proved rigidity for the **differential inclusion**

$$Du \in \operatorname{span}\{P\}, \qquad u \in \mathrm{BV}(C; \mathbb{R}^m), \qquad P \in \mathbb{R}^{m \times d} \text{ fixed}, \tag{4.6}$$

which, if we additionally assume $|P| = 1$, is to be interpreted as

$$\frac{\mathrm{d}Du}{\mathrm{d}|Du|} = P \qquad |Du|\text{-a.e.}$$

Notice that for $u \in \mathrm{W}^{1,1}(C; \mathbb{R}^m)$, this simply means

$$\nabla u(x) \in \operatorname{span}\{P\}, \qquad \text{for a.e. } x \in C.$$

Rigidity here refers to the fact that (4.6) has either only affine solutions if $\operatorname{rank} P \geq 2$, or one-directional solutions ("plane waves") in direction ξ if $P = a \otimes \xi$. For the terminology also cf. Definition 1.1 in [77].

The following corollary is not needed in the sequel, but is included for completeness.

Corollary 4.8 (Rigidity for approximate solutions). *Let $\Omega \subset \mathbb{R}^d$ be a bounded open connected Lipschitz domain. If $P \in \mathbb{R}^{m \times d}$ with $\operatorname{rank} P \geq 2$ and $(u_j) \subset \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m)$ is a sequence such that*

$$\begin{aligned} u_j &\overset{*}{\rightharpoonup} u \quad \text{in } \mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^m) \qquad \text{and} \\ \operatorname{dist}\big(\nabla u_j, \operatorname{span}\{P\}\big) &\to 0 \quad \text{in measure,} \end{aligned}$$

then even

$$\nabla u_j \to \text{const} \quad \text{in measure.}$$

Proof. This follows from Lemma 2.7 (ii) in [96] and Lemma 4.4. □

The following lemma applies to all types of blow-ups and is essentially a local version of the Rigidity Lemma 4.4.

Lemma 4.9 (Local structure of BV-derivatives). *Let $u \in \mathrm{BV}(\Omega;\mathbb{R}^m)$. Then, 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 $\operatorname{rank} P_0 \geq 2$, then $v(x) = v_0 + \alpha P_0 x$ a.e., where $\alpha \in \mathbb{R}$, $v_0 \in \mathbb{R}^m$.

(ii) If $P_0 = a \otimes \xi$ ($a \in \mathbb{R}^m$, $\xi \in \mathbb{S}^{d-1}$), then there exist $\psi \in \mathrm{BV}(\mathbb{R})$, $v_0 \in \mathbb{R}^m$ such that $v(x) = v_0 + \psi(x \cdot \xi)a$ a.e.

Proof. By an argument similar to the one in the proof of localization principles, Propositions 3.5, 3.6, we get that every $\tau \in \mathrm{Tan}(Du, x_0)$ is a BV-derivative of a function $v \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^m)$. Moreover, from (3.5) we get that $Dv = P_0|Dv|$. Hence, we are in the situation of Lemma 4.4 and the conclusion follows from this. □

Remark 4.10 (Comparison to Alberti's Rank One Theorem). The preceding result can be seen as a weaker version of Alberti's Rank One Theorem 2.25. Indeed, from the Local Structure Lemma 4.9 we get that every tangent measure $\tau \in \mathrm{Tan}(Du, x_0)$ at almost every point $x_0 \in \Omega$ is the derivative of a BV-function $v \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^m)$, which either has the form $v(x) = v_0 + \alpha P_0 x$ or $v(x) = v_0 + \psi(x \cdot \xi)a$ with $a \in \mathbb{R}^m$, $\xi \in \mathbb{S}^{d-1}$, $\psi \in \mathrm{BV}(\mathbb{R})$, $v_0 \in \mathbb{R}^m$, and $\alpha \in \mathbb{R}$. In the second case also $P_0 = a \otimes \xi$ and only at these points we can assert that the conclusion of Alberti's Theorem holds. But Preiss's Example 5.9 (1) from [105] shows that the first case may even occur almost everywhere, so the result is potentially much weaker than Alberti's. Nevertheless, our lemma still asserts that locally at singular points, $\tau = Dv$ is always one-directional, i.e. translation-invariant in all but at most one direction (which usually is proved as a corollary to Alberti's Theorem) and this will suffice later on. On a related note, Preiss's example alluded to above is also the reason why Alberti's Theorem cannot be proved by a blow-up argument, see Section 3 of [38] for further explanation.

4.3 Jensen-type inequalities

This section establishes Jensen-type inequalities for gradient Young measures (defined in Section 2.6). We proceed separately for the regular and the singular part of the Young measure and employ in particular the localization principles of Chapter 3 and the Rigidity Lemma 4.4.

The proof of the Jensen-type inequality at a regular point is rather straightforward:

Proposition 4.11. *Let $\nu \in \mathbf{GY}(\Omega;\mathbb{R}^{m\times d})$ be a gradient Young measure. Then, for $\mathcal{L}^d$-almost every $x_0 \in \Omega$ it holds that*

$$h\biggl(\bigl\langle \mathrm{id}, \nu_{x_0}\bigr\rangle + \bigl\langle \mathrm{id}, \nu_{x_0}^\infty\bigr\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0)\biggr) \leq \bigl\langle h, \nu_{x_0}\bigr\rangle + \bigl\langle h^\#, \nu_{x_0}^\infty\bigr\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0)$$

for all quasiconvex $h \in \mathrm{C}(\mathbb{R}^{m\times d})$ with linear growth at infinity.

Proof. Let $\sigma \in \mathbf{GY}(\mathbb{B}^d;\mathbb{R}^{m\times d})$ be (the restriction of) a regular tangent Young measure to ν at a suitable $x_0 \in \Omega$ as in Proposition 3.5. In particular, $[\sigma] = A_0\,\mathcal{L}^d \llcorner \mathbb{B}^d$, where

$$A_0 = \bigl\langle \mathrm{id}, \nu_{x_0}\bigr\rangle + \bigl\langle \mathrm{id}, \nu_{x_0}^\infty\bigr\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0) \quad \in \mathbb{R}^{m\times d}.$$

Use Proposition 2.30 (ii) to get a sequence $(v_n) \subset \mathrm{W}^{1,1}(\mathbb{B}^d;\mathbb{R}^m)$ with $Dv_n \xrightarrow{\mathbf{Y}} \sigma$ and $v_n(x) = A_0 x$ on $\mathbb{S}^{d-1}$. Hence, for all quasiconvex $h \in \mathrm{C}(\mathbb{R}^{m\times d})$ with $\mathbb{1} \otimes h \in \mathbf{E}(\mathbb{B}^d;\mathbb{R}^{m\times d})$ we have

$$h(A_0) \leq \fint_{\mathbb{B}^d} h(\nabla v_n)\,\mathrm{d}x \qquad \text{for all } n \in \mathbb{N}.$$

Now, use Lemma 2.2 to get a sequence $(\mathbb{1} \otimes h_k) \subset \mathbf{E}(\mathbb{B}^d;\mathbb{R}^{m\times d})$ such that $h_k \downarrow h$, $h_k^\infty \downarrow h^\#$ pointwise, and all h_k have uniformly bounded linear growth constants. Then, by (3.12), for all $k \in \mathbb{N}$,

$$\begin{aligned} h(A_0) &\leq \limsup_{n\to\infty} \fint_{\mathbb{B}^d} h(\nabla v_n)\,\mathrm{d}x \leq \lim_{n\to\infty} \fint_{\mathbb{B}^d} h_k(\nabla v_n)\,\mathrm{d}x \\ &= \frac{1}{\omega_d}\bigl\langle\!\bigl\langle \mathbb{1} \otimes h_k, \sigma \bigr\rangle\!\bigr\rangle = \bigl\langle h_k, \nu_{x_0}\bigr\rangle + \bigl\langle h_k^\infty, \nu_{x_0}^\infty\bigr\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0). \end{aligned}$$

Letting $k \to \infty$ together with the Monotone Convergence Theorem proves the claim of the proposition. □

Establishing a Jensen-type inequality for the singular points is more involved, but the basic principle of blowing up around a point x_0 and then applying quasiconvexity remains the same. Additionally, however, we need to include an averaging procedure since tangent Young measures at singular points might not have an affine function as underlying deformation. For this, we need the information provided by the Rigidity Lemma 4.4.

Proposition 4.12. *Let $\nu \in \mathbf{GY}(\Omega;\mathbb{R}^{m\times d})$ be a gradient Young measure. Then, for λ_ν^s-almost every $x_0 \in \Omega$ it holds that*

$$g\bigl(\bigl\langle \mathrm{id}, \nu_{x_0}^\infty\bigr\rangle\bigr) \leq \bigl\langle g, \nu_{x_0}^\infty\bigr\rangle$$

for all quasiconvex and positively 1-homogeneous functions $g \in \mathrm{C}(\mathbb{R}^{m\times d})$.

Notice that we did not say anything about the validity of a singular Jensen-type inequality at boundary points $x_0 \in \partial\Omega$. This is also not needed in the sequel.

Proof. Let $x_0 \in \Omega$ be as in Proposition 3.6 and define (which is possible λ_ν^s-almost everywhere)

$$A_0 := \langle \mathrm{id}, \nu_{x_0}^\infty \rangle.$$

We distinguish two cases:

Case 1: $\operatorname{rank} A_0 \geq 2$ *or* $A_0 = 0$ *(semi-regular blow-up).*
Let $\sigma \in \mathbf{GY}(Q;\mathbb{R}^{m\times d})$ be (the restriction of) a singular tangent Young measure to ν at x_0, whose existence is ascertained by Proposition 3.6, where $Q = (-1/2, 1/2)^d$ is the unit cube. Moreover, we can assume without loss of generality that $\lambda_\sigma(Q) = 1$ and $\lambda_\sigma(\partial Q) = 0$, possibly taking a concentric larger cube $Q' \supset\supset Q$ and then adjusting the blow-up sequence to scale back to the unit cube. In this semi-regular case, we can proceed analogously to the regular blow-up in Proposition 4.11: From Proposition 2.30 (ii) take $(v_n) \subset \mathrm{W}^{1,1}(Q;\mathbb{R}^m)$ with $\nabla v_n \overset{\mathbf{Y}}{\to} \sigma$, $v_n \overset{*}{\rightharpoonup} v \in \mathrm{BV}(Q;\mathbb{R}^m)$, and $v_n|_{\partial Q} = v|_{\partial Q}$. Also, $Dv = A_0\lambda_\sigma$ and so, invoking the Rigidity Lemma 4.4 (i),

$$v(x) = A_0 x, \quad x \in Q, \qquad \text{and} \qquad Dv = A_0\, \mathcal{L}^d \llcorner Q,$$

where, without loss of generality, we assumed that the constant part of v is zero. The quasiconvexity of g immediately yields

$$g(A_0) \leq \fint_Q g(\nabla v_n)\,\mathrm{d}x \qquad \text{for all } n \in \mathbb{N}.$$

Then,

$$g(A_0) \leq \limsup_{n\to\infty} \fint_Q g(\nabla v_n)\,\mathrm{d}x = \langle\!\langle \mathbb{1}_{\overline{Q}} \otimes g, \sigma \rangle\!\rangle = \langle g, \nu_{x_0}^\infty \rangle,$$

where the last equality follows from (3.16).

Case 2: $A_0 = a \otimes \xi$ *for* $a \in \mathbb{R}^m \setminus \{0\}$, $\xi \in \mathbb{S}^{d-1}$ *(fully singular blow-up).*
To simplify notation we assume that $\xi = \mathrm{e}_1 = (1, 0, \ldots, 0)^T$; otherwise the unit cube $Q = (-1/2, 1/2)^d$ in the following proof has to be replaced by a rotated unit cube with one face orthogonal to ξ.

Like in Case 1, take $\sigma \in \mathbf{GY}(Q;\mathbb{R}^{m\times d})$ to be a singular tangent Young measure to ν at x_0 as in Proposition 3.6 with $\lambda_\sigma(\partial Q) = 0$. Also let $(v_n) \subset \mathrm{W}^{1,1}(Q;\mathbb{R}^m)$ with $\nabla v_n \overset{\mathbf{Y}}{\to} \sigma$, $v_n \overset{*}{\rightharpoonup} v \in \mathrm{BV}(Q;\mathbb{R}^m)$, and $v_n|_{\partial Q} = v|_{\partial Q}$. Using case (ii) of the Rigidity Lemma 4.4 and adding a constant function if necessary, v can be written in the form

$$v(x) = \psi(x_1)a \qquad \text{for some } \psi \in \mathrm{BV}(-1/2, 1/2).$$

Observe that we may also require $\lambda_\sigma(Q) = 1$, whereby

$$A_0 = A_0\lambda_\sigma(Q) = [\sigma](Q) = Dv(Q) = D\psi((-1/2, 1/2))A_0,$$

and hence

$$D\psi((-1/2, 1/2)) = \psi(1/2 - 0) - \psi(-1/2 + 0) = 1.$$

Set

$$\tilde{u}_n(x) := v_n\left(x - \left\lfloor x + \frac{1}{2} \right\rfloor\right) + a\left\lfloor x_1 + \frac{1}{2} \right\rfloor, \qquad x \in \mathbb{R}^d,$$

where $\lfloor s \rfloor$ is the largest integer smaller than or equal to $s \in \mathbb{R}$ and $\lfloor x \rfloor := (\lfloor x_1 \rfloor, \ldots, \lfloor x_d \rfloor)$ for $x \in \mathbb{R}^d$. Then define $(u_n) \subset \mathrm{BV}(Q;\mathbb{R}^m)$ by

$$u_n = \frac{\tilde{u}_n(nx)}{n}, \qquad x \in Q,$$

and observe that

$$\nabla u_n(x) = \sum_{z \in \{0,\ldots,n-1\}^d} \nabla v_n(nx - z) \mathbb{1}_{Q(z/n,1/n)}(x), \qquad x \in Q.$$

Furthermore, Du_n has zero singular part because the gluing discontinuities over the hyperplanes $\{x_1 = \ell/n + 1/(2n)\}$, are compensated by the jumps of the staircase term in the definition of $\tilde{u}_n$.

It is easy to see that $u_n \to A_0 x$ in $\mathrm{L}^1(Q;\mathbb{R}^m)$ by a change of variables and since ψ is bounded. Therefore, using Proposition 2.30 (ii) again, we find another sequence $(w_n) \subset \mathrm{W}^{1,1}(Q;\mathbb{R}^m)$ with $w_n(x) = A_0 x$ for all $x \in \partial Q$, and such that the sequences (∇u_n) and (∇w_n) generate the same (unnamed) Young measure, in particular

$$\lim_{n\to\infty} \int_Q g(\nabla w_n) \,\mathrm{d}x = \lim_{n\to\infty} \int_Q g(\nabla u_n) \,\mathrm{d}x$$

for all g as in the statement of the proposition. Consequently, we get by quasiconvexity

$$g(A_0) \le \fint_Q g(\nabla w_n) \,\mathrm{d}x \qquad \text{for all } n \in \mathbb{N}.$$

Then use (3.16) to deduce

$$\begin{aligned}
g(A_0) &\le \limsup_{n\to\infty} \fint_Q g(\nabla w_n) \,\mathrm{d}x = \lim_{n\to\infty} \fint_Q g(\nabla u_n) \,\mathrm{d}x \\
&= \lim_{n\to\infty} \sum_{z \in \{0,\ldots,n-1\}^d} \int_{Q(z/n,1/n)} g\big(\nabla v_n(nx - z)\big) \,\mathrm{d}x \\
&= \lim_{n\to\infty} \sum_{z \in \{0,\ldots,n-1\}^d} \frac{1}{n^d} \int_Q g(\nabla v_n) \,\mathrm{d}y \\
&= \lim_{n\to\infty} \int_Q g(\nabla v_n) \,\mathrm{d}y = \big\langle\!\big\langle \mathbb{1}_{\overline{Q}} \otimes g, \sigma \big\rangle\!\big\rangle = \big\langle g, \nu_{x_0}^\infty \big\rangle.
\end{aligned}$$

This concludes the proof. □

Summing up, the following theorem exhibits necessary conditions for a Young measure to be a gradient Young measure.

Theorem 4.13 (Jensen-type inequalities). *Let $\nu \in \mathbf{GY}(\Omega; \mathbb{R}^{m\times d})$ be a gradient Young measure. Then, for all quasiconvex $h \in \mathrm{C}(\mathbb{R}^{m\times d})$ with linear growth at infinity, it holds that*

$$h\biggl(\bigl\langle \mathrm{id}, \nu_x \bigr\rangle + \bigl\langle \mathrm{id}, \nu_x^\infty \bigr\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x)\biggr) \le \bigl\langle h, \nu_x \bigr\rangle + \bigl\langle h^\#, \nu_x^\infty \bigr\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x)$$

for $\mathcal{L}^d$-almost every $x \in \Omega$, and

$$h^\#\bigl(\bigl\langle \mathrm{id}, \nu_x^\infty \bigr\rangle\bigr) \le \bigl\langle h^\#, \nu_x^\infty \bigr\rangle$$

for λ_ν^s-almost every $x \in \Omega$.

The proof is contained in Propositions 4.11 and 4.12 once we notice that if h is quasiconvex, then its generalized recession function $h^\#$ is quasiconvex as well (by Fatou's lemma), and hence continuous (see Section 4.1).

4.4 Lower semicontinuity

We can now prove the main lower semicontinuity result of this chapter.

Proof of Theorem 4.1. Let $u_j \overset{*}{\rightharpoonup} u$ in $\mathrm{BV}(\Omega; \mathbb{R}^m)$. Consider all u_j, u to be extended to $\mathbb{R}^d$ by zero. Assume also that $Du_j \overset{\mathbf{Y}}{\to} \nu \in \mathbf{GY}(\mathbb{R}^d; \mathbb{R}^{m\times d})$, for which it follows that

$$[\nu] = Du + (u|_{\partial\Omega} \otimes n_\Omega)\, \mathcal{H}^{d-1} \llcorner \partial\Omega.$$

This entails taking a subsequence if necessary, but since we will show an inequality for all such subsequences, it also holds for the original sequence. Observe that if λ_ν^* is the singular part of λ_ν with respect to $|D^s u| + |u|\, \mathcal{H}^{d-1} \llcorner \partial\Omega$, i.e. λ_ν^* is concentrated in an $(|D^s u| + |u|\, \mathcal{H}^{d-1} \llcorner \partial\Omega)$-negligible set, then

$$\bigl\langle \mathrm{id}, \nu_x \bigr\rangle + \bigl\langle \mathrm{id}, \nu_x^\infty \bigr\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x) = \frac{\mathrm{d}[\nu]}{\mathrm{d}\mathcal{L}^d}(x) = \begin{cases} \nabla u(x) & \text{for } \mathcal{L}^d\text{-a.e. } x \in \Omega, \\ 0 & \text{for } \mathcal{L}^d\text{-a.e. } x \in \mathbb{R}^d \setminus \Omega, \end{cases}$$

$$\frac{\bigl\langle \mathrm{id}, \nu_x^\infty \bigr\rangle}{|\bigl\langle \mathrm{id}, \nu_x^\infty \bigr\rangle|} = \frac{\mathrm{d}[\nu]^s}{\mathrm{d}|[\nu]^s|}(x) = \begin{cases} \dfrac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x) & \text{for } |D^s u|\text{-a.e. } x \in \Omega, \\ \dfrac{u|_{\partial\Omega}(x)}{|u|_{\partial\Omega}(x)|} \otimes n_\Omega(x) & \text{for } |u|\, \mathcal{H}^{d-1}\text{-a.e. } x \in \partial\Omega, \end{cases}$$

$$\bigl\langle \mathrm{id}, \nu_x^\infty \bigr\rangle = 0 \qquad \text{for } \lambda_\nu^*\text{-a.e. } x \in \mathbb{R}^d,$$

$$\bigl|\bigl\langle \mathrm{id}, \nu_x^\infty \bigr\rangle\bigr|\, \lambda_\nu^s = |D^s u| + |u|\, \mathcal{H}^{d-1} \llcorner \partial\Omega,$$

$$\bigl\langle \mathrm{id}, \nu_x \bigr\rangle = 0 \qquad \text{for all } x \in \mathbb{R}^d \setminus \overline{\Omega},$$

$$\lambda_\nu \llcorner (\mathbb{R}^d \setminus \overline{\Omega}) = 0.$$

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 function, f^∞ is jointly continuous and $f(x, 0) = 0$ for all $x \in \mathbb{R}^d \setminus \overline{\Omega}$.

Then, from the Jensen-inequalities in Theorem 4.13 together with the extended representation result for Young measures, Proposition 2.17 in Section 2.3.1, we get

$$\begin{aligned}
\liminf_{j\to\infty} \mathcal{F}(u_j) &= \int \big\langle f(x,\bullet),\nu_x \big\rangle + \big\langle f^\infty(x,\bullet),\nu_x^\infty \big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x)\,\mathrm{d}x \\
&\quad + \int \big\langle f^\infty(x,\bullet),\nu_x^\infty \big\rangle \,\mathrm{d}\lambda_\nu^s(x) \\
&\geq \int f\Big(x, \langle \mathrm{id},\nu_x\rangle + \langle \mathrm{id},\nu_x^\infty\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x)\Big)\,\mathrm{d}x \\
&\quad + \int f^\infty\big(x, \langle \mathrm{id},\nu_x^\infty\rangle\big)\,\mathrm{d}\lambda_\nu^s(x) \\
&= \mathcal{F}(u).
\end{aligned}$$

This proves the claim. □

Remark 4.14 (Recession functions). In comparison to previously known results, we have to assume that the "strong" recession function f^∞ exists instead of merely using the generalized recession function $f^\#$. This is in fact an unavoidable phenomenon of our proof strategy without Alberti's Rank One Theorem: It is well-known (see for instance Theorem 2.5 (iii) in [3]) that the natural recession function for lower semicontinuity is the *lower* generalized recession function

$$f_\#(x,A) := \liminf_{\substack{x'\to x\\ A'\to A\\ t\to\infty}} \frac{f(tA')}{t}, \qquad x\in\overline{\Omega},\ A\in\mathbb{R}^{m\times d}.$$

Unfortunately, we cannot easily determine whether this function is quasiconvex, so the singular Jensen-type inequality from Prosposition 4.12 is not applicable. The usual proof that $f^\#$ (and hence f^∞) is quasiconvex whenever f is, proceeds by virtue of Fatou's lemma, and this method fails for $f_\#$. One can show, however, that if $f_\#$ is known to be quasiconvex, then the lower semicontinuity theorem also holds with $f_\#$ in place of f^∞. Indeed, take a sequence $(f_k) \subset \mathbf{E}(\Omega;\mathbb{R}^{m\times d})$ with $f_k \uparrow f$, $f_k^\infty \uparrow f_\#$, and define $\mathcal{F}_k$ like $\mathcal{F}$, but with f replaced by f_k. Also, let $\mathcal{F}_\#$ be the functional with f^∞ replaced by $f_\#$. Then,

$$\begin{aligned}
\liminf_{j\to\infty} \mathcal{F}_\#(u_j) &\geq \lim_{k\to\infty}\lim_{j\to\infty} \mathcal{F}_k(u_j) = \lim_{k\to\infty} \big\langle\!\big\langle f_k,\nu \big\rangle\!\big\rangle \\
&= \int \big\langle f(x,\bullet),\nu_x\big\rangle\,\mathrm{d}x + \int \big\langle f_\#(x,\bullet),\nu_x^\infty\big\rangle\,\mathrm{d}\lambda_\nu^s(x) \\
&\geq \mathcal{F}_\#(u)
\end{aligned}$$

by the Monotone Convergence Theorem and the Jensen-type inequalities. Hence, $\mathcal{F}_\#$ is weakly* lower semicontinuous. By Alberti's Rank One Theorem we know that this is the same functional since $f_\#(x,A) = f^\#(x,A)$ if $\operatorname{rank} A \leq 1$ (by the rank-one convexity of f). This should be contrasted with the fact that $f_\#$ and $f^\#$ may differ outside the rank-one cone, see [95].

We also immediately get the following corollary on the functional without the boundary term:

Corollary 4.15. *For every quasiconvex $f \in \mathbf{E}(\Omega; \mathbb{R}^{m\times d})$, the functional*

$$\mathcal{F}(u) := \int_\Omega f\bigl(x, \nabla u(x)\bigr) \, \mathrm{d}x + \int_\Omega f^\infty\biggl(x, \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x)\biggr) \, \mathrm{d}|D^s u|(x), \qquad u \in \mathrm{BV}(\Omega; \mathbb{R}^m),$$

is sequentially lower semicontinuous with respect to all weakly-converging sequences $u_j \overset{*}{\rightharpoonup} u$ in $\mathrm{BV}(\Omega; \mathbb{R}^m)$ if $f \geq 0$ or $u_j|_{\partial\Omega} = u|_{\partial\Omega}$ for all $j \in \mathbb{N}$.*

This follows from Theorem 4.1 since in the above cases the boundary term can be neglected. Note that for signed integrands, the above corollary might be *false*, as can be seen from easy counterexamples.

Remark 4.16. We note that it is also possible to show lower semicontinuity of integral functionals in $\mathrm{BV}(\Omega; \mathbb{R}^m)$ (or relaxation theorems) without the use of Young measures. For example, for the functional

$$\mathcal{F}(u) := \int_\Omega f(\nabla u) \, \mathrm{d}x + \int_\Omega f^\infty\biggl(\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}\biggr) \, \mathrm{d}|D^s u|, \qquad u \in \mathrm{BV}(\Omega; \mathbb{R}^m),$$

where $f \in \mathbf{E}(\Omega; \mathbb{R}^{m\times d})$ takes only non-negative values, does not depend on x, and is quasiconvex, one can follow the proof of lower semicontinuity in [8] (reproduced in Section 5.5 of [10]) almost completely line-by-line, but replacing the use of Alberti's Rank One Theorem with the Rigidity Lemma 4.4 (or the Local Structure Lemma 4.9). Indeed, in the estimate of the singular part from below, Alberti's Theorem is only used to show that the blow-up limit is one-directional and we can reach that conclusion also by our Rigidity Lemma. The other occurrence of Alberti's Rank One Theorem in that proof concerns the fact that we can use the generalized recession function $f^\#$ instead of requiring the existence of the (strong) recession function f^∞. This, in fact, cannot be avoided without Alberti's Theorem, cf. Remark 4.14. However, while this alternative proof circumvents the framework of Young measures, it uses other technical results instead (like the De Giorgi–Letta Theorem).

Chapter 5

Lower semicontinuity in BD

The aim of this chapter is to prove the following general lower semicontinuity theorem from [108]:

Theorem 5.1. *Let $\Omega \subset \mathbb{R}^d$ be a bounded open Lipschitz domain and let $f \in \mathbf{R}(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}})$ be a symmetric-quasiconvex integrand, that is, $f \colon \overline{\Omega} \times \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}$ satisfies the following assumptions:*

(i) f is a Carathéodory function,

(ii) $|f(x, A)| \le M(1 + |A|)$ for some $M > 0$ and all $x \in \overline{\Omega}$, $A \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$,

(iii) the (strong) recession function f^∞ exists in the sense of (2.6) *and is (jointly) continuous on $\overline{\Omega} \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$,*

(iv) $f(x, \cdot)$ is symmetric-quasiconvex for all $x \in \overline{\Omega}$ (see Section 5.3 below).

Then, the functional

$$\begin{aligned}\mathcal{F}(u) := &\int_\Omega f\bigl(x, \mathcal{E}u(x)\bigr) \,\mathrm{d}x + \int_\Omega f^\infty\biggl(x, \frac{\mathrm{d}E^s u}{\mathrm{d}|E^s u|}(x)\biggr) \,\mathrm{d}|E^s u|(x) \\ &+ \int_{\partial\Omega} f^\infty\bigl(x, u|_{\partial\Omega}(x) \odot n_\Omega(x)\bigr) \,\mathrm{d}\mathcal{H}^{d-1}(x), \qquad u \in \mathrm{BD}(\Omega),\end{aligned}$$

is lower semicontinuous with respect to weak-convergence in the space* $\mathrm{BD}(\Omega)$

As in the previous chapter, we use a Young measure approach based on the localization principles from Chapter 3 in order to prove suitable Jensen-type inequalities, which then immediately yield lower semicontinuity.

5.1 Functions of bounded deformation and BD-Young measures

This section collects some well-known properties of functions of bounded deformation (BD) and introduces the corresponding class of Young measures. More information on BD and applications to plasticity theory can be found in [7, 61, 122] and also in [79, 80, 86, 88, 111, 112, 123].

Let Ω be an open domain with Lipschitz boundary. For a function $u \in \mathrm{W}^{1,1}_{\mathrm{loc}}(\Omega;\mathbb{R}^d)$ define the **symmetrized gradient** (or **deformation tensor**) via

$$\mathcal{E}u := \frac{1}{2}\big(\nabla u + \nabla u^T\big), \qquad \mathcal{E}u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}}). \tag{5.1}$$

The space $\mathrm{BD}(\Omega)$ of **functions of bounded deformation** is the space of functions $u \in \mathrm{L}^1(\Omega;\mathbb{R}^d)$ such that the distributional **symmetrized derivative** (defined by duality with $\mathcal{E}u$)

$$Eu = \frac{1}{2}\big(Du + Du^T\big)$$

is (representable as) a finite Radon measure, $Eu \in \mathbf{M}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$. The space $\mathrm{BD}(\Omega)$ is a Banach space under the norm

$$\|u\|_{\mathrm{BD}(\Omega)} := \|u\|_{\mathrm{L}^1(\Omega;\mathbb{R}^d)} + |Eu|(\Omega).$$

We split Eu according to the Lebesgue–Radon–Nikodým decomposition

$$Eu = \mathcal{E}u\,\mathcal{L}^d + E^s u,$$

where $\mathcal{E}u = \frac{\mathrm{d}Eu}{\mathrm{d}\mathcal{L}^d} \in \mathrm{L}^1(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ denotes the Radon–Nikodým derivative of Eu with respect to Lebesgue measure and $E^s u$ is singular. We call $\mathcal{E}u$ the **approximate symmetrized gradient** (the reason for the word "approximate" can be found in Section 4 of [7]).

The subspace $\mathrm{LD}(\Omega)$ of $\mathrm{BD}(\Omega)$ consists of all BD-functions such that Eu is absolutely continuous with respect to Lebesgue measure (i.e. $E^s u = 0$). The space $\mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$ is the space of functions $u \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)$ such that the restriction of u to every relatively compact open subset $U \subset \mathbb{R}^d$ lies in $\mathrm{BD}(U)$. Since there is no Korn inequality in L^1, see [32, 103], it can be shown that $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^d)$ is a proper subspace of $\mathrm{LD}(\Omega)$ and also that the space $\mathrm{BV}(\Omega;\mathbb{R}^d)$ is a proper subspace of $\mathrm{BD}(\Omega)$.

A **rigid deformation** is a skew-symmetric affine map $u\colon \mathbb{R}^d \to \mathbb{R}^d$, i.e. u is of the form

$$u(x) = u_0 + Rx, \qquad \text{where } u_0 \in \mathbb{R}^d,\ R \in \mathbb{R}^{d\times d}_{\mathrm{skew}}.$$

The following lemma is well-known and will be used many times in the sequel, usually without mentioning. We reproduce its proof here because the central formula (5.2) will be of use later.

Lemma 5.2. *The kernel of the linear operator $\mathcal{E}\colon \mathrm{C}^1(\mathbb{R}^d;\mathbb{R}^d) \to \mathrm{C}(\mathbb{R}^d;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ given in* (5.1) *is the space of rigid deformations.*

Proof. It is obvious that $\mathcal{E}u$ vanishes for a rigid deformation u.

For the other direction, let $u \in \mathrm{C}^1(\mathbb{R}^d;\mathbb{R}^d)$ with $\mathcal{E}u \equiv 0$ and define

$$\mathcal{W}u := \frac{1}{2}\big(\nabla u - \nabla u^T\big).$$

Then, for all $i, j, k = 1, \ldots, d$, we have in the sense of distributions (recall that we denote by A^i_j the element of A in the ith row and jth column),

$$\begin{aligned} \partial_k \mathcal{W}u^i_j &= \frac{1}{2}\big(\partial_{kj}u^i - \partial_{ki}u^j\big) = \frac{1}{2}\big(\partial_{jk}u^i + \partial_{ji}u^k\big) - \frac{1}{2}\big(\partial_{ij}u^k + \partial_{ik}u^j\big) \\ &= \partial_j \mathcal{E}u^i_k - \partial_i \mathcal{E}u^j_k \equiv 0. \end{aligned} \tag{5.2}$$

As $\nabla u = \mathcal{E}u + \mathcal{W}u$, this entails that ∇u is a constant, hence u is affine and it is clear that it in fact must be a rigid deformation. □

It is an easy consequence of the previous lemma that $u \in \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$ with $Eu = A\mathcal{L}^d$, where $A \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ is a fixed symmetric matrix, is an affine function. More precisely, $u(x) = u_0 + (A + R)x$ for some $u_0 \in \mathbb{R}^d$ and $R \in \mathbb{R}^{d\times d}_{\mathrm{skew}}$.

As notions of convergence in $\mathrm{BD}(\Omega)$ we have the **strong (norm) convergence** $u_j \to u$ and the **weak* convergence** $u_j \overset{*}{\rightharpoonup} u$ in $\mathrm{BD}(\Omega)$, meaning $u_j \to u$ strongly in $\mathrm{L}^1(\Omega;\mathbb{R}^m)$ and $Eu_j \overset{*}{\rightharpoonup} Eu$ in $\mathbf{M}(\Omega;\mathbb{R}^m)$. If $\sup_j \|Eu_j\|_{\mathrm{BD}(\Omega)} < \infty$, then there exists a weakly* converging subsequence. Moreover, we say that (u_j) converges **strictly** or **$\langle\boldsymbol{\cdot}\rangle$-strictly (area-strictly)** to u if $u_j \overset{*}{\rightharpoonup} u$ in $\mathrm{BD}(\Omega)$ and additionally $|Eu_j|(\Omega) \to |Eu|(\Omega)$ or $\langle Eu_j\rangle(\Omega) \to \langle Eu\rangle(\Omega)$, respectively.

In $\mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$ we let **weak* convergence** mean $u_j \to u$ in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)$ (i.e. in L^1 on all compact subsets of $\mathbb{R}^d$) and $Eu_j \overset{*}{\rightharpoonup} Eu$ in $\mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^{d\times d})$. If $(u_j) \subset \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$ and $\sup_j \|u_j\|_{\mathrm{BD}(U)} < \infty$ for all relatively compact open $U \subset \mathbb{R}^d$, then there exists a weakly* converging subsequence.

Since Ω has a Lipschitz boundary, the **trace** $u|_{\partial\Omega}$ of u onto $\partial\Omega$ is well-defined in the sense that there exists a bounded linear operator $u \mapsto u|_{\partial\Omega}$ mapping $\mathrm{BD}(\Omega)$ (surjectively) onto $\mathrm{L}^1(\partial\Omega, \mathcal{H}^{d-1};\mathbb{R}^d)$ that coincides with the natural trace for all $u \in \mathrm{BD}(\Omega) \cap \mathrm{C}(\overline{\Omega};\mathbb{R}^d)$, see Theorem II.2.1 of [123].

If $u \in \mathrm{BD}(\Omega)$ with $u|_{\partial\Omega} = 0$, then we have the Poincaré inequality

$$\|u\|_{\mathrm{BD}(\Omega)} \leq C|Eu|(\Omega),$$

where $C = C(\Omega)$ only depends on the domain Ω, see Proposition II.2.4 in [122]. Moreover, it is shown for example in [123] (or see Remark II.2.5 of [122]) that for each $u \in \mathrm{BD}(\Omega)$ there exists a rigid deformation r such that

$$\|u + r\|_{\mathrm{L}^{d/(d-1)}(\Omega;\mathbb{R}^d)} \leq C|Eu|(\Omega),$$

where again $C = C(\Omega)$.

Since $\mathrm{LD}(\Omega)$ is $\langle\boldsymbol{\cdot}\rangle$-strictly dense in $\mathrm{BD}(\Omega)$ (by a mollification argument), the Reshetnyak Continuity Theorem 2.3 immediately implies the following result, which is analogous to Proposition 4.2 for the BV-case.

Proposition 5.3. *Let $f \in \mathbf{R}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ (not necessarily symmetric-quasiconvex). Then, $\mathcal{F}$ from Theorem 5.1 is the $\langle\,\boldsymbol{\cdot}\,\rangle$-strictly continuous extension of the functional*

$$\int_\Omega f\bigl(x,\mathcal{E}u(x)\bigr)\,\mathrm{d}x + \int_{\partial\Omega} f^\infty\bigl(x,u|_{\partial\Omega}(x)\odot n_\Omega(x)\bigr)\,\mathrm{d}\mathcal{H}^{d-1}(x), \qquad u \in \mathrm{LD}(\Omega),$$

onto the space $\mathrm{BD}(\Omega)$.

A Young measure $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ is called a **BD-Young measure**, in symbols $\nu \in \mathbf{BDY}(\Omega)$, if it can be generated by a sequence of elementary Young measures corresponding to symmetrized derivatives. That is, for all $\nu \in \mathbf{BDY}(\Omega)$, there exists a (norm-bounded) sequence $(u_j) \subset \mathrm{BD}(\Omega)$ with $Eu_j \overset{\mathbf{Y}}{\to} \nu$. It is easy to see that for a BD-Young measure $\nu \in \mathbf{BDY}(\Omega)$, there exists $u \in \mathrm{BD}(\Omega)$ satisfying $Eu = [\nu]\llcorner\Omega$, any such u is called an **underlying deformation** of ν. Similarly, define $\mathbf{BDY}_{\mathrm{loc}}(\mathbb{R}^d)$ by replacing $\mathbf{Y}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ and $\mathrm{BD}(\Omega)$ by their respective local counterparts. When working with $\mathbf{BDY}(\Omega)$ or $\mathbf{BDY}_{\mathrm{loc}}(\mathbb{R}^d)$, the appropriate spaces of integrands are $\mathbf{E}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ and $\mathbf{E}_c(\mathbb{R}^d;\mathbb{R}^{d\times d}_{\mathrm{sym}})$, respectively, since it is clear that both ν_x and ν_x^∞ only take values in $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ whenever $\nu \in \mathrm{BD}(\Omega)$ or $\nu \in \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$.

The whole theory of Young measures as developed in Chapter 2 is applicable here, including the theory for gradient Young measures (such as the localization principles from Chapter 3) with the necessary adjustments.

5.2 Symmetric tensor products

Recall the definition $a \odot b := (a\otimes b + b\otimes a)/2$ of the **symmetric tensor product**. We record the following lemma about these special matrices in $\mathbb{R}^{2\times 2}$:

Lemma 5.4. *Let $M \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}$ be a non-zero symmetric matrix.*

(i) If $\operatorname{rank} M = 1$, then $M = \pm a\odot a = \pm a\otimes a$ for a vector $a \in \mathbb{R}^2$.

(ii) If $\operatorname{rank} M = 2$, then $M = a\odot b$ for vectors $a,b \in \mathbb{R}^2$ if and only if the two (non-zero, real) eigenvalues of M have opposite signs.

Proof. Ad (i). Every rank-one matrix M can be written as a tensor product $M = c\otimes d$ for some vectors $c,d \in \mathbb{R}^2\setminus\{0\}$. By the symmetry, we get $c_1d_2 = c_2d_1$, which implies that the vectors c and d are multiples of each other. We therefore find $a \in \mathbb{R}^2$ with $M = \pm a\otimes a$.

Ad (ii). Assume first that $M = a\odot b$ for some vectors $a,b \in \mathbb{R}^2$ and take an orthogonal matrix $Q \in \mathbb{R}^{2\times 2}$ such that QMQ^T is diagonal. Moreover,

$$QMQ^T = \frac{1}{2}Q\bigl(a\otimes b + b\otimes a\bigr)Q^T = \frac{1}{2}\bigl(Qa\otimes Qb + Qb\otimes Qa\bigr) = Qa\odot Qb,$$

whence we may always assume without loss of generality that M is already diagonal,

$$a\odot b = M = \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix},$$

where $\lambda_1, \lambda_2 \neq 0$ are the two eigenvalues of M. Writing this out componentwise, we get

$$a_1 b_1 = \lambda_1, \qquad a_2 b_2 = \lambda_2, \qquad a_1 b_2 + a_2 b_1 = 0.$$

As $\lambda_1, \lambda_2 \neq 0$, also $a_1, a_2, b_1, b_2 \neq 0$, and hence

$$0 = a_1 b_2 + a_2 b_1 = \frac{a_1}{a_2}\lambda_2 + \frac{a_2}{a_1}\lambda_1.$$

Thus, λ_1 and λ_2 must have opposite signs.

For the other direction, by transforming as before we may assume again that M is diagonal, $M = \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix}$ and that λ_1 and λ_2 do not have the same sign. Then, with $\gamma := \sqrt{-\lambda_1/\lambda_2}$, we define

$$a := \begin{pmatrix} \gamma \\ 1 \end{pmatrix}, \qquad b := \begin{pmatrix} \lambda_1 \gamma^{-1} \\ \lambda_2 \end{pmatrix}.$$

For $\lambda_1 > 0$, $\lambda_2 < 0$ say (the other case is analogous),

$$\lambda_1 \gamma^{-1} + \lambda_2 \gamma = \lambda_1 \sqrt{\frac{|\lambda_2|}{\lambda_1}} - |\lambda_2| \sqrt{\frac{\lambda_1}{|\lambda_2|}} = 0,$$

and therefore

$$a \odot b = \frac{1}{2}\begin{pmatrix} \lambda_1 & \lambda_2\gamma \\ \lambda_1\gamma^{-1} & \lambda_2 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \lambda_1 & \lambda_1\gamma^{-1} \\ \lambda_2\gamma & \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} = M.$$

This proves the claim. □

5.3 Symmetric quasiconvexity

A locally bounded Borel function $h\colon \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}$ is called **symmetric-quasiconvex** if

$$h(A) \leq \fint_\omega h\big(A + \mathcal{E}\psi(z)\big) \,\mathrm{d}z \qquad \text{for all } A \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \text{ and all } \psi \in \mathrm{W}^{1,\infty}_0(\omega;\mathbb{R}^d),$$

where $\omega \subset \mathbb{R}^d$ is an arbitrary bounded open Lipschitz domain (by standard covering arguments it suffices to check this for one particular choice of ω only). It is not difficult to see that $h\colon \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}$ is symmetric-quasiconvex if and only if $h \circ P\colon \mathbb{R}^{d\times d} \to \mathbb{R}$ is quasiconvex in the usual (gradient) sense, where $P(A) := (A + A^T)/2$ is the projection onto the symmetric part of $A \in \mathbb{R}^{d\times d}$.

Notice that if h has linear growth at infinity, we may replace the space $\mathrm{W}^{1,\infty}_0(\omega;\mathbb{R}^d)$ by $\mathrm{LD}_0(\omega)$ (LD-functions with zero boundary values in the sense of trace) in the above definition, see Remark 3.2 in [22].

Using one-directional oscillations one can prove that if the function $h\colon \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}$ is symmetric-quasiconvex, then it holds that

$$h(\theta A_1 + (1-\theta)A_2) \leq \theta h(A_1) + (1-\theta) h(A_2) \tag{5.3}$$

whenever $A_1, A_2 \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ with $A_2 - A_1 = a \odot b$ for some $a, b \in \mathbb{R}^d$ and $\theta \in [0,1]$, cf. Proposition 3.4 in [58] for a more general statement in the framework of $\mathcal{A}$-quasiconvexity.

If we consider $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ to be identified with $\mathbb{R}^{d(d+1)/2}$ and $h\colon \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}$ with $\tilde{h}\colon \mathbb{R}^{d(d+1)/2} \to \mathbb{R}$, then the convexity in (5.3) implies that $\tilde{h}$ is separately convex and so, by a well-known result, h is locally Lipschitz continuous, see for example Lemma 2.2 in [19]. If additionally h has linear growth at infinity, then the formula from *loc. cit.* even implies that h is globally Lipschitz. In particular, we may use the simplified definitions for the recession function in (2.8).

Notice that from Fatou's Lemma we get that the recession function $f^{\#}$, and hence also f^{∞} if it exists, is symmetric-quasiconvex whenever f is, this is completely analogous to the situation for ordinary quasiconvexity. Hence, $f^{\#}$ and f^{∞} are also continuous on $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ in this situation.

We mention that symmetric-quasiconvex functions with linear growth at infinity exist. In fact, we have analogously to Lemma 4.3 for quasiconvexity:

Lemma 5.5. *Let $M \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}$ be a matrix that cannot be written in the form $a \odot b$ for $a, b \in \mathbb{R}^2$ (see Lemma 5.4) and let $p \in [1, \infty)$. Define*

$$h(A) := \operatorname{dist}(A, \{-M, M\})^p, \qquad A \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}.$$

Then, $h(0) > 0$ and the symmetric-quasiconvex envelope SQh (defined analogously to the quasiconvex envelope) is not convex (at zero). Moreover, h has a (strong) p-recession function h^{∞}, and it holds that $(SQh)^{\infty} = h^{\infty} = |\cdot|^p$.

The proof of this is similar to the one for Lemma 5.5. A direct argument for $M = I_2$ can be found in [46].

5.4 Construction of good singular blow-ups

This section combines the localization principles from Chapter 3 with rigidity arguments to show that among the possibly many singular tangent Young measures of a BD-Young measure $\nu \in \mathbf{BDY}(\Omega)$, there are always "good" ones at λ_ν^s-almost every point $x_0 \in \Omega$. More concretely, we will construct blow-ups that are either affine or that are sums of one-directional functions, see Figure 5.1. A key point in this construction is an iteration of the blow-up construction via Proposition 3.3; this technique could already be observed in the proof of Proposition 3.4. Some concrete differential inclusions involving the symmetrized gradient $\mathcal{E}u$ in two dimensions are treated more elaborately in Section 5.7 for the purpose of illustration.

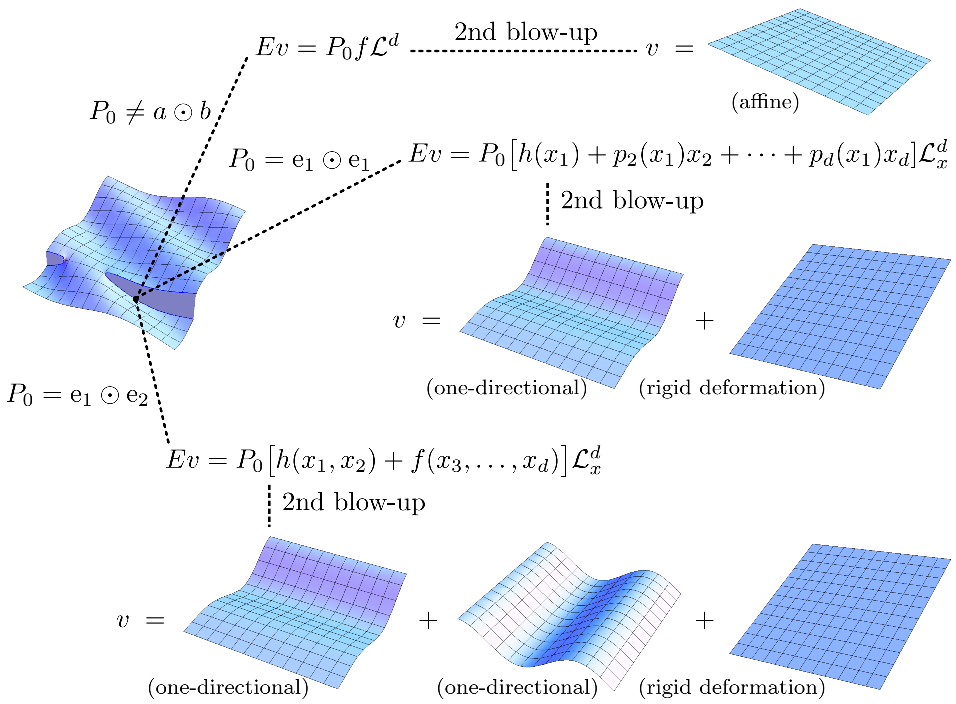

Figure 5.1: Constructing good singular blow-ups.

Theorem 5.6 (Good singular blow-ups). *Let $\nu \in \mathbf{BDY}(\Omega)$ be a BD-Young measure. For λ_ν^s-almost every $x_0 \in \Omega$, there exists a singular tangent Young measure $\sigma \in \mathbf{BDY}_{\mathrm{loc}}(\mathbb{R}^d)$ as in Proposition 3.6 such that additionally for any $v \in \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$ with $Ev = [\sigma]$:*

(i) If $\langle \mathrm{id}, \nu_{x_0}^\infty \rangle \notin \{\, a \odot b \;:\; a, b \in \mathbb{R}^d \setminus \{0\}\,\}$ (this includes the case $\langle \mathrm{id}, \nu_{x_0}^\infty \rangle = 0$), then v is equal to an affine function almost everywhere.

(ii) If $\langle \mathrm{id}, \nu_{x_0}^\infty \rangle = a \odot b$ ($a, b \in \mathbb{R}^d \setminus \{0\}$) with $a \neq b$, then there exist functions $h_1, h_2 \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$, $v_0 \in \mathbb{R}^d$, and a skew-symmetric matrix $R \in \mathbb{R}^{d\times d}_{\mathrm{skew}}$ such that

$$v(x) = v_0 + h_1(x \cdot a)b + h_2(x \cdot b)a + Rx, \qquad x \in \mathbb{R}^d \text{ a.e.}$$

(iii) If $\langle \mathrm{id}, \nu_{x_0}^\infty \rangle = a \odot a$ ($a \in \mathbb{R}^d \setminus \{0\}$), then there exists a function $h \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$, $v_0 \in \mathbb{R}^d$ and a skew-symmetric matrix $R \in \mathbb{R}^{d\times d}_{\mathrm{skew}}$ such that

$$v(x) = v_0 + h(x \cdot a)a + Rx, \qquad x \in \mathbb{R}^d \text{ a.e.}$$

Remark 5.7. In contrast to the situation for the space BV, where *all* blow-ups could be shown to have a good structure, in BD we may only ascertain that there *exists* at least one good blow-up. Moreover, no analogue of Alberti's Rank One Theorem 2.25 is known.

Example 5.8. Let $\Omega = (-1,1)^2 \subset \mathbb{R}^2$ and let

$$u := \begin{pmatrix} \mathbb{1}_{\{x_2>0\}} \\ \mathbb{1}_{\{x_1>0\}} \end{pmatrix}.$$

Then, $u \in \mathbf{BDY}(\Omega)$ and

$$Eu = (\mathrm{e}_1 \odot \mathrm{e}_2)\big[\mathcal{H}^1 \llcorner \{x_1 = 0\} + \mathcal{H}^1 \llcorner \{x_2 = 0\}\big].$$

Hence, for the elementary BD-Young measure ε_{Eu} at the origin, case (ii) of the preceding theorem is applicable; notice that indeed we need both h_1 and h_2 for the result to be true.

With the notation of the theorem we set

$$P_0 := \begin{cases} \frac{\langle \mathrm{id}, \nu_{x_0}^\infty \rangle}{|\langle \mathrm{id}, \nu_{x_0}^\infty \rangle|} & \text{if } \langle \mathrm{id}, \nu_{x_0}^\infty \rangle \neq 0, \\ 0 & \text{if } \langle \mathrm{id}, \nu_{x_0}^\infty \rangle = 0. \end{cases}$$

The proof will be accomplished in the following three sections, its main scheme is shown in Figure 5.1.

5.4.1 The case $P_0 \neq a \odot b$

The proof technique for this case consists of a rigidity/ellipticity argument using Fourier multipliers and projections (inspired by the proof of Lemma 2.7 in [96]) together with an iterated blow-up.

Proof of Theorem 5.6 (i). Take a singular tangent Young measure $\nu \in \mathbf{BDY}_{\mathrm{loc}}(\mathbb{R}^d)$ at a point $x_0 \in \Omega$ as in Proposition 3.6 and let $v \in \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$ with $Ev = [\sigma]$. This v then satisfies (by the properties of singular tangent Young measures, see (3.16))

$$Ev = P_0|Ev|.$$

If $P_0 = 0$ (i.e. $\langle \mathrm{id}, \nu_{x_0}^\infty \rangle = 0$), then we immediately have that v is affine. Hence from now on we assume $P_0 \neq 0$.

Step 1. Suppose first that v is smooth. Let $A\colon \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d}$ be the orthogonal projection onto $(\mathrm{span}\{P_0\})^\perp$. Then,

$$A(\mathcal{E}v) \equiv 0. \tag{5.4}$$

For every smooth cut-off function $\varphi \in \mathrm{C}_c^\infty(\mathbb{R}^d; [0,1])$, the function $w := \varphi v$ satisfies (here exceptionally considering $\nabla\varphi$ as a column vector)

$$\mathcal{E}w = \varphi\mathcal{E}v + v \odot \nabla\varphi.$$

Combining this with (5.4), we get

$$A(\mathcal{E}w) = A(v \odot \nabla\varphi) =: f, \tag{5.5}$$

where by means of an embedding result in BD [123], $f \in \mathrm{L}^{d/(d-1)}(\mathbb{R}^d;\mathbb{R}^{d\times d})$ (L^∞ if $d = 1$).

If for the Fourier transform $\hat{g}$ of a function $g \in \mathrm{L}^1(\mathbb{R}^d;\mathbb{R}^N)$ we use the definition

$$\hat{g}(\xi) := \int g(x)\mathrm{e}^{-2\pi\mathrm{i}x\cdot\xi}\,\mathrm{d}x, \qquad \xi \in \mathbb{R}^d,$$

then it can be checked easily that

$$\widehat{\mathcal{E}w}(\xi) = (2\pi\mathrm{i})\,\hat{w}(\xi)\odot\xi.$$

Hence, applying the Fourier transform to both sides of (5.5) and considering A to be identified with its complexification (that is, $A(M+\mathrm{i}N) = A(M)+\mathrm{i}A(N)$ for $M,N \in \mathbb{R}^{d\times d}$), we arrive at

$$(2\pi\mathrm{i})\,A(\hat{w}(\xi)\odot\xi) = \hat{f}(\xi) \qquad \text{for all } \xi \in \mathbb{R}^d. \tag{5.6}$$

Step 2. We will now use some linear algebra to rewrite (5.6) as a Fourier multiplier equation and then apply a version of the Mihlin multiplier theorem.

Notice first that (the complexification of) the projection $A\colon \mathbb{C}^{d\times d} \to \mathbb{C}^{d\times d}$ has kernel $\mathrm{span}\{P_0\}$ (here and in the following all spans are understood in $\mathbb{C}^{d\times d}$) and hence descends to the quotient

$$[A]\colon \mathbb{C}^{d\times d}/\,\mathrm{span}\{P_0\} \to \mathrm{ran}\,A,$$

and $[A]$ is an invertible linear map. By assumption, $P_0 \neq a\odot b$ for any $a,b \in \mathbb{R}^d$. Then, for $\xi \in \mathbb{R}^d \setminus \{0\}$, let

$$\bigl\{P_0, \mathrm{e}_1\odot\xi,\ldots,\mathrm{e}_d\odot\xi, G_{d+1}(\xi),\ldots,G_{d^2-1}(\xi)\bigr\} \subset \mathbb{R}^{d\times d}$$

be a basis of $\mathbb{C}^{d\times d}$ with the property that the matrices $G_{d+1}(\xi),\ldots,G_{d^2-1}(\xi)$ depend smoothly on ξ and are positively 1-homogeneous in ξ. For all $\xi \in \mathbb{R}^d\setminus\{0\}$, denote by $B(\xi)\colon \mathbb{C}^{d\times d} \to \mathbb{C}^{d\times d}$ the (non-orthogonal) projection with

$$\begin{aligned}
&\ker B(\xi) = \mathrm{span}\{P_0\},\\
&\mathrm{ran}\,B(\xi) = \mathrm{span}\bigl\{\mathrm{e}_1\odot\xi,\ldots,\mathrm{e}_d\odot\xi, G_{d+1}(\xi),\ldots,G_{d^2-1}(\xi)\bigr\}.
\end{aligned}$$

If we interpret $\mathrm{e}_1\odot\xi,\ldots,\mathrm{e}_d\odot\xi, G_{d+1}(\xi),\ldots,G_{d^2-1}(\xi)$ as vectors in $\mathbb{R}^{d^2}$, collect them into the columns of the matrix $X(\xi) \in \mathbb{R}^{d^2\times(d^2-1)}$, and if we further let $Y \in \mathbb{R}^{d^2\times(d^2-1)}$ be a matrix whose columns comprise an orthonormal basis of $(\mathrm{span}\{P_0\})^\perp$, then $B(\xi)$ can be written explicitly as (it is elementary to see that $Y^TX(\xi)$ is invertible)

$$B(\xi) = X(\xi)(Y^TX(\xi))^{-1}Y^T.$$

This implies that $B(\xi)$ is positively 0-homogeneous and using Cramer's Rule we also see that $B(\xi)$ depends smoothly on $\xi \in \mathbb{R}^d\setminus\{0\}$ (if $\det(Y^TX(\xi))$ was not bounded away from zero for $\xi \in \mathbb{S}^{d-1}$, then by compactness there would exist $\xi_0 \in \mathbb{S}^{d-1}$ with $\det(Y^T X(\xi_0)) = 0$, a contradiction). Of course, also $B(\xi)$ descends to a quotient

$$[B(\xi)] \colon \mathbb{C}^{d\times d}/\operatorname{span}\{P_0\} \to \operatorname{ran} B(\xi),$$

which is now invertible. It is not difficult to see that $\xi \mapsto [B(\xi)]$ is still positively 0-homogeneous and smooth in $\xi \neq 0$ (for example by utilizing the basis given above).

Since $\hat{w}(\xi) \odot \xi \in \operatorname{ran} B(\xi)$, we notice that $[B(\xi)]^{-1}(\hat{w}(\xi) \odot \xi) = [\hat{w}(\xi) \odot \xi]$, the equivalence class of $\hat{w}(\xi) \odot \xi$ in $\mathbb{C}^{d\times d}/\operatorname{span}\{P_0\}$. This allows us to rewrite (5.6) in the form

$$(2\pi \mathrm{i})\, [A][B(\xi)]^{-1}(\hat{w}(\xi) \odot \xi) = \hat{f}(\xi),$$

or equivalently as

$$(2\pi \mathrm{i})\, \hat{w}(\xi) \odot \xi = [B(\xi)][A]^{-1}\hat{f}(\xi).$$

The function $M \colon \mathbb{R}^d \setminus \{0\} \to \mathbb{R}^{d^2 \times d^2}$ given by $\xi \mapsto [B(\xi)][A]^{-1}$ is smooth and positively 0-homogeneous, and we have the multiplier equation

$$\widehat{\mathcal{E}w}(\xi) = (2\pi \mathrm{i})\, \hat{w}(\xi) \odot \xi = M(\xi)\hat{f}(\xi).$$

A matrix-version of the Mihlin Multiplier Theorem, see Theorem 6.1.6 in [24] and also Theorem 5.2.7 in [68], now yields

$$\|\mathcal{E}w\|_{\mathrm{L}^{d/(d-1)}} \le C\|f\|_{\mathrm{L}^{d/(d-1)}} \le C\|v\|_{\mathrm{L}^{d/(d-1)}(K;\mathbb{R}^d)}, \tag{5.7}$$

where $K := \operatorname{supp} \varphi$ and $C = C(K, \|A\|, \|\nabla\varphi\|_\infty)$ is a constant.

Step 3. If $v \in \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$ is not smooth, we take a family of mollifiers $(\rho_\delta)_{\delta>0}$ and define by convolution $v_\delta := \rho_\delta \star v \in \mathrm{C}^\infty(\mathbb{R}^d;\mathbb{R}^d)$. Correspondingly, with a fixed cut-off function φ as above we set $w_\delta := \varphi v_\delta$. This mollification preserves the property $Ev_\delta = P_0|Ev_\delta|$ and so (5.7) gives for $\delta < 1$,

$$\|\mathcal{E}w_\delta\|_{\mathrm{L}^{d/(d-1)}} \le C\|v_\delta\|_{\mathrm{L}^{d/(d-1)}(K;\mathbb{R}^d)} \le C\|v\|_{\mathrm{L}^{d/(d-1)}(K_1;\mathbb{R}^d)},$$

where again $K := \operatorname{supp} \varphi$ and $K_1 := K + \mathbb{B}^d$.

Since $\mathcal{E}w_\delta\, \mathcal{L}^d \overset{*}{\rightharpoondown} Ew$ as $\delta \downarrow 0$, the previous δ-uniform estimate implies that Ew is absolutely continuous with respect to Lebesgue measure, i.e. $Ew = \mathcal{E}w\, \mathcal{L}^d$ with $\mathcal{E}w \in \mathrm{L}^{d/(d-1)}(K;\mathbb{R}^{d\times d}_{\mathrm{sym}})$. Finally, varying φ, we get that also Ev is absolutely continuous with respect to Lebesgue measure and $\mathcal{E}v \in \mathrm{L}^{d/(d-1)}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^{d\times d}_{\mathrm{sym}})$.

Step 4. We have shown so far that $[\sigma] = Ev$ is absolutely continuous with respect to Lebesgue measure. Now apply Proposition 3.5 and Preiss's existence result for non-zero tangent measures to σ in order to infer the existence of a *regular* tangent Young measure κ to σ at $\mathcal{L}^d$-almost every point $y_0 \in \operatorname{supp}[\sigma]$ with $[\kappa] \neq 0$. It is not difficult to see that κ is still a singular tangent measure to ν at x_0 in the sense of Proposition 3.6. Indeed, one may observe first that (3.14), (3.15) with κ in place of σ still hold by the conclusion of Proposition 3.5

and (3.14), (3.15) for σ together with the fact that tangent measures to tangent measures are tangent measures, see Proposition 3.3 (we need to select $x_0 \in \Omega$ according to that lemma, which is still possible λ_ν^s-almost everywhere). Finally, we see that (3.16) also holds with κ in place of σ because this assertion always follows from (3.14), (3.15).

On the other hand, by the absolute continuity of Ev with respect to $\mathcal{L}^d$ and standard results on tangent measures, we may in fact choose y_0 such that $[\kappa] \in \mathrm{Tan}(Ev, y_0)$ is a constant multiple of Lebesgue measure, see Section 3.1. Thus, any $\tilde{v} \in \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$ with $E\tilde{v} = [\kappa]$ is affine. This shows the claim of Theorem 5.6 (i) with κ in place of σ and $\tilde{v}$ in place of v. □

5.4.2 The case $P_0 = a \odot b$

This case is more involved, yet essentially elementary. We first examine the situation in two dimensions and then, via a dimension reduction lemma, extend the result to an arbitrary number of dimensions.

Lemma 5.9 (2D rigidity). *A function $u \in \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^2)$ satisfies*

$$Eu = \frac{a \odot b}{|a \odot b|}|Eu| \qquad \text{for fixed } a, b \in \mathbb{R}^2 \setminus \{0\} \text{ with } a \neq b, \tag{5.8}$$

if and only if u has the form

$$u(x) = h_1(x \cdot a)b + h_2(x \cdot b)a, \qquad x \in \mathbb{R}^2 \text{ a.e.}, \tag{5.9}$$

where $h_1, h_2 \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$.

Notice that we are only imposing a condition on the symmetric derivative, which only determines a function up to a rigid deformation. In the above case, however, since a and b are linearly independent, we may absorb this rigid deformation into h_1 and h_2.

Proof. By the chain rule in BV, it is easy to see that all u of the form (5.9) satisfy (5.8).

For the other direction, without loss of generality we suppose that $a = \mathrm{e}_1$, $b = \mathrm{e}_2$ (see Step 2 in the proof of Theorem 5.6 (ii) below for an explicit reduction; in fact, this lemma will only be used in the case $a = \mathrm{e}_1$, $b = \mathrm{e}_2$ anyway).

We will use a slicing result, Proposition 3.2 in [7], which essentially follows from Fubini's Theorem: If for $\xi \in \mathbb{R}^2 \setminus \{0\}$ we define

$$H_\xi := \{\, x \in \mathbb{R}^2 \;:\; x \cdot \xi = 0 \,\},$$
$$u_y^\xi(t) := \xi^T u(y + t\xi), \qquad \text{where } t \in \mathbb{R},\ y \in H_\xi,$$

then the result in *loc. cit.* states

$$\left|\xi^T Eu \xi\right| = \int_{H_\xi} \left|Du_y^\xi\right| \,\mathrm{d}\mathcal{H}^1(y) \qquad \text{as measures.} \tag{5.10}$$

By assumption, $Eu = \sqrt{2}(\mathrm{e}_1 \odot \mathrm{e}_2)|Eu|$ with $|Eu| \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^2)$, so if we apply (5.10) for $\xi = \mathrm{e}_1$, we get

$$0 = \frac{\sqrt{2}}{2}\left|\mathrm{e}_1^T\left(\mathrm{e}_1\mathrm{e}_2^T + \mathrm{e}_2\mathrm{e}_1^T\right)\mathrm{e}_1\right| |Eu| = \int_{H_\xi} \left|\partial_t u^1(y + t\mathrm{e}_1)\right| \,\mathrm{d}\mathcal{H}^1(y),$$

where we wrote $u = (u^1, u^2)^T$. This yields $\partial_1 u^1 \equiv 0$ distributionally, whence $u^1(x) = h_2(x_2)$ for some $h_2 \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$. Analogously, we find that $u^2(x) = h_1(x_1)$ with $h_1 \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$. Thus, we may decompose

$$u(x) = \begin{pmatrix} 0 \\ h_1(x_1) \end{pmatrix} + \begin{pmatrix} h_2(x_2) \\ 0 \end{pmatrix} = h_1(x \cdot \mathrm{e}_1)\mathrm{e}_2 + h_2(x \cdot \mathrm{e}_2)\mathrm{e}_1$$

and it only remains to show that $h_1, h_2 \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$. For this, fix $\eta \in \mathrm{C}^1_c(\mathbb{R}; [-1, 1])$ with $\int \eta \,\mathrm{d}t = 1$ and calculate for all $\varphi \in \mathrm{C}^1_c(\mathbb{R}; [-1, 1])$ by Fubini's Theorem,

$$\begin{aligned} 2\int \varphi \otimes \eta \,\mathrm{d}(Eu)^1_2 &= -\int u^2(\varphi' \otimes \eta) \,\mathrm{d}x - \int u^1(\varphi \otimes \eta') \,\mathrm{d}x \\ &= -\int h_1\varphi' \,\mathrm{d}x_1 \cdot \int \eta \,\mathrm{d}x_2 - \int u^1(\varphi \otimes \eta') \,\mathrm{d}x. \end{aligned}$$

So, with $K := \operatorname{supp} \varphi \times \operatorname{supp} \eta$,

$$\left|\int h_1\varphi' \,\mathrm{d}x\right| \leq 2|Eu|(K) + \|u^1\|_{\mathrm{L}^1(K)} \cdot \|\eta'\|_\infty < \infty$$

for all $\varphi \in \mathrm{C}^1_c(\mathbb{R})$ with $\|\varphi\|_\infty \leq 1$, hence $h_1 \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$. Likewise, $h_2 \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$, and we have shown the lemma. □

Next we need to extend the preceding rigidity lemma to an arbitrary number of dimensions. This is the purpose of the following lemma, which we only formulate for the case $P = \mathrm{e}_1 \odot \mathrm{e}_2$ to avoid notational clutter (we will only need this special case later).

Lemma 5.10 (Dimension reduction). *Let $u \in \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$ be such that*

$$Eu = \sqrt{2}(\mathrm{e}_1 \odot \mathrm{e}_2)|Eu|.$$

Then, there exist a Radon measure $\mu \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^2)$ and a linear function $f\colon \mathbb{R}^{d-2} \to \mathbb{R}$ such that

$$Eu = \sqrt{2}(\mathrm{e}_1 \odot \mathrm{e}_2)\left[\mu \otimes \mathcal{L}^{d-2} + f(x_3, \ldots, x_d)\mathcal{L}^d_x\right].$$

Proof. We suppose $d \geq 3$. In all of the following, let

$$P_0 := \sqrt{2}(\mathrm{e}_1 \odot \mathrm{e}_2).$$

Step 1. We first assume that u is smooth. In this case, there exists $g \in \mathrm{C}^\infty(\mathbb{R}^d)$ with

$$\mathcal{E}u = P_0 g \qquad \text{and} \qquad E^s u = 0.$$

Clearly,

$$\mathcal{E}u(x)^j_k = \frac{\sqrt{2}}{2}\begin{cases} g(x) & \text{if } (j,k) = (1,2) \text{ or } (j,k) = (2,1), \\ 0 & \text{otherwise.} \end{cases}$$

Fix $i \geq 3$. With

$$\mathcal{W}u := \frac{1}{2}\big(\nabla u - \nabla u^T\big)$$

we have from (5.2),

$$\partial_k \mathcal{W}u^i_j = \partial_j \mathcal{E}u^i_k - \partial_i \mathcal{E}u^j_k, \qquad \text{for } j,k = 1,\dots,d.$$

Since $i \geq 3$, only the second term is possibly non-zero, so

$$\nabla \mathcal{W}u(x)^i_j = -\partial_i \mathcal{E}u(x)^j = -\frac{\sqrt{2}}{2}\begin{cases} (0, \partial_i g(x), 0, \dots, 0) & \text{if } j = 1, \\ (\partial_i g(x), 0, 0, \dots, 0) & \text{if } j = 2, \\ (0, \dots, 0) & \text{if } j \geq 3. \end{cases}$$

It is elementary to see that if a function $h \in \mathrm{C}^\infty(\mathbb{R}^d)$ satisfies $\partial_k h \equiv 0$ for all $k = 2, \dots, d$, then, with a slight abuse of notation, $h(x) = h(x_1)$ and also $\partial_1 h(x) = \partial_1 h(x_1)$. In our situation this gives that $\partial_i g$ can be written both as a function of x_1 only and as a function of x_2 only. But this is only possible if $\partial_i g$ is constant, say $\partial_i g \equiv a_i \in \mathbb{R}$ for $i = 3, \dots, d$.

If we set

$$f(x) := a_3 x_3 + \cdots + a_d x_d,$$

we have that the function $h(x) := g(x) - f(x)$ only depends on the first two components x_1, x_2 of x, and thus

$$Eu = P_0\Big[h(x_1,x_2)\mathcal{L}^d_x + f(x_3,\dots,x_d)\mathcal{L}^d_x\Big].$$

Step 2. Now assume that only $u \in \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$. We will reduce this case to the previous one by a smoothing argument. Set $u_\delta := \rho_\delta \star u \in \mathrm{C}^\infty(\mathbb{R}^d;\mathbb{R}^d)$, where $(\rho_\delta)_{\delta>0}$ is a family of mollifying kernels. It can be seen that $Eu_\delta = \sqrt{2}(\mathrm{e}_1 \odot \mathrm{e}_2)|Eu_\delta|$ still holds, so we may apply the first step to get a smooth function $h_\delta \in \mathrm{C}^\infty(\mathbb{R}^2)$ and a linear function $f_\delta \colon \mathbb{R}^{d-2} \to \mathbb{R}$ such that

$$Eu_\delta = P_0\Big[h_\delta(x_1,x_2)\mathcal{L}^d_x + f_\delta(x_3,\dots,x_d)\mathcal{L}^d_x\Big].$$

We will show that also the limit has an analogous form: With the cube $Q^k(R) := (-R,R)^k$ $(R > 0)$, take $\varphi \in \mathrm{C}_c(Q^2(R);[-1,1])$ and define the measures

$$\mu_\delta := h_\delta(x_1,x_2)\mathcal{L}^2_{(x_1,x_2)} \quad \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^2).$$

We have from Fubini's Theorem,

$$\int \varphi \otimes \mathbb{1}_{Q^{d-2}(R)} \,\mathrm{d}Eu_\delta = P_0\left[(2R)^{d-2}\int \varphi \,\mathrm{d}\mu_\delta + \int \varphi \,\mathrm{d}x \cdot \int_{Q^{d-2}(R)} f_\delta \,\mathrm{d}x\right]$$

The second term on the right hand side is identically zero since f_δ is linear and $Q^{d-2}(R)$ is symmetric, so, with a constant $C = C(R)$,

$$\limsup_{\delta \downarrow 0} \int \varphi \, \mathrm{d}\mu_\delta \leq C \limsup_{\delta \downarrow 0} |Eu_\delta|(Q^d(R)) < \infty.$$

Therefore, selecting a subsequence of δs, we may assume that $\mu_\delta \overset{*}{\rightharpoondown} \mu \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^2)$, which entails $\mu_\delta \otimes \mathcal{L}^{d-2} \overset{*}{\rightharpoondown} \mu \otimes \mathcal{L}^{d-2}$. Moreover, if $f_\delta \mathcal{L}^d \overset{*}{\rightharpoondown} \gamma \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^d)$, then γ must be of the form $f\mathcal{L}^d$ with $f = f(x_3, \ldots, x_d)$ linear since the space of measures of this form is finite-dimensional and hence weakly* closed. Consequently, we see that there exists a Radon measure $\mu \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^2)$ and a linear map $f \colon \mathbb{R}^{d-2} \to \mathbb{R}$ such that

$$Eu = P_0 \Big[\mu \otimes \mathcal{L}^{d-2} + f(x_3, \ldots, x_d) \mathcal{L}^d_x \Big].$$

This proves the claim. □

We can now finish the proof of case (ii) of our theorem:

Proof of Theorem 5.6 (ii). Like in the proof of part (i) of the theorem, take a singular tangent Young measure $\nu \in \mathbf{BDY}_{\mathrm{loc}}(\mathbb{R}^d)$ at a point $x_0 \in \Omega$ as in Proposition 3.6 and let $v \in \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$ with $Ev = [\sigma]$. As before, it holds from the properties of tangent Young measures that

$$Ev = P_0 |Ev|.$$

Step 1. We first show the result in the case $a = \mathrm{e}_1$, $b = \mathrm{e}_2$, i.e. $P_0 = \sqrt{2}(\mathrm{e}_1 \odot \mathrm{e}_2)$. Under this assumption we may apply the dimensional reduction result from Lemma 5.10 to get a Radon measure $\mu \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R}^2)$ and a linear function $f \colon \mathbb{R}^{d-2} \to \mathbb{R}$, for which

$$Ev = P_0 \Big[\mu \otimes \mathcal{L}^{d-2} + f(x_3, \ldots, x_d) \mathcal{L}^d_x \Big].$$

If f is non-zero, $[\sigma] = Ev$ cannot be purely singular and so there exists an $\mathcal{L}^d$-negligible set $N \subset \mathbb{R}^d$ such that $[\sigma] \llcorner (\mathbb{R}^d \setminus N) = g\mathcal{L}^d$ for a non-zero $g \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^{d \times d}_{\mathrm{sym}})$. Hence, by virtue of Proposition 3.5 and Preiss's existence result for non-zero tangent measures, there is $y_0 \in \mathbb{R}^d$ and a regular tangent Young measure $\kappa \in \mathbf{BDY}_{\mathrm{loc}}(\mathbb{R}^d)$ to σ at y_0 with $[\kappa]$ a non-zero constant multiple of Lebesgue measure, namely $[\kappa] = \alpha P_0 \mathcal{L}^d$ for some $\alpha \neq 0$. Therefore, any $\tilde{v} \in \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$ with $[\kappa] = E\tilde{v}$ is affine and in particular of the form exhibited in case (ii) of the theorem (with h_1, h_2 linear). As in Step 4 of the proof of part (i) of the present theorem, we can show that κ is a singular tangent measure to ν at x_0 as well (in the sense that it satisfies the conclusion of Proposition 3.6). Thus, in the case that f is not identically zero, we have already shown part (ii) of the present theorem with $\tilde{v}$ and κ in place of v and σ, respectively.

Next we treat the remaining case where $f \equiv 0$ and Ev might be purely singular, that is

$$Ev = P_0 \, \mu \otimes \mathcal{L}^{d-2}. \tag{5.11}$$

In this situation we have that there exists a function $h \in \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^2)$ and $v_0 \in \mathbb{R}^d$ as well as a skew-symmetric matrix $R \in \mathbb{R}^{d\times d}_{\mathrm{skew}}$ such that

$$v(x) = v_0 + \begin{pmatrix} h^1(x_1,x_2) \\ h^2(x_1,x_2) \\ 0 \\ \vdots \\ 0 \end{pmatrix} + Rx.$$

This can roughly be seen as follows: By a mollification argument, we may assume that v is smooth. Then, (5.11) means that $\mathcal{E}v(x) = P_0 g(x_1,x_2)$ for some $g \in \mathrm{C}^\infty(\mathbb{R}^2)$, $x \in \mathbb{R}^d$. Hence, the function

$$h(x_1,x_2) := \begin{pmatrix} v^1(x_1,x_2,0,\dots,0) \\ v^2(x_1,x_2,0,\dots,0) \end{pmatrix},$$

has symmetrized gradient $\mathcal{E}h(x_1,x_2) = \tilde{P}_0 g(x_1,x_2)$, where $\tilde{P}_0$ is the leading principal (2×2)-submatrix of P_0. Considering h to be extended to a function on $\mathbb{R}^d$ (constant in $x_3,\dots,x_d$) and with d components ($h^3,\dots,h^d = 0$), we have that $\mathcal{E}h = \mathcal{E}v$ and so, v equals h modulo a rigid deformation.

But for h we can invoke Lemma 5.9 to deduce that

$$h(x_1,x_2) = h_1(x_1)\mathrm{e}_2 + h_2(x_2)\mathrm{e}_1.$$

where $h_1, h_2 \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$. Thus, we arrive at

$$v(x) = v_0 + h_1(x_1)\mathrm{e}_2 + h_2(x_2)\mathrm{e}_1 + Rx.$$

This proves the claim for $a = \mathrm{e}_1$, $b = \mathrm{e}_2$.

Step 2. For general $a, b \in \mathbb{R}^d$ with $a \neq b$ take an invertible matrix $G \in \mathbb{R}^{d\times d}$ with $Ga = \mathrm{e}_1$, $Gb = \mathrm{e}_2$. Then $G(a \odot b)G^T = \mathrm{e}_1 \odot \mathrm{e}_2$ and hence, replacing $v(x)$ by

$$\tilde{v}(x) := Gv(G^T x),$$

we have $E\tilde{v} = \sqrt{2}(\mathrm{e}_1 \odot \mathrm{e}_2)|E\tilde{v}|$. By the previous step, there exist $\tilde{v}_0 \in \mathbb{R}^d$ and a skew-symmetric matrix $\tilde{R} \in \mathbb{R}^{d\times d}_{\mathrm{skew}}$ such that

$$\tilde{v}(x) = \tilde{v}_0 + h_1(x_1)\mathrm{e}_2 + h_2(x_2)\mathrm{e}_1 + \tilde{R}x.$$

We can now transform back to the original $v(x) = G^{-1}\tilde{v}(G^{-T}x)$. In this process, we get

$$\begin{aligned} &G^{-1}h_1(G^{-T}x\cdot \mathrm{e}_1)\mathrm{e}_2 = h_1(x\cdot a)b, \qquad \text{and} \\ &G^{-1}h_2(G^{-T}x\cdot \mathrm{e}_2)\mathrm{e}_1 = h_2(x\cdot b)a. \end{aligned}$$

Also setting $v_0 := G^{-1}\tilde{v}_0$ and $R := G^{-1}RG^{-T}$, which is still skew-symmetric, we have proved the claimed splitting in the general situation as well. $\square$

5.4.3 The case $P_0 = a \odot a$

For this degenerate case we can essentially use the same techniques as in the previous sections, but there are some differences.

Proof of Theorem 5.6 (iii). Once more we take a singular tangent Young measure $\sigma \in \mathbf{BDY}_{\mathrm{loc}}(\mathbb{R}^d)$ at a point $x_0 \in \Omega$ from Proposition 3.6 and $v \in \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$ with

$$Ev = [\sigma] = (a \odot a)|Ev|.$$

Step 1. In case that v is smooth and $a = \mathrm{e}_1$, i.e. there exists $g \in \mathrm{C}^\infty(\mathbb{R}^d)$ such that

$$\mathcal{E}v = (\mathrm{e}_1 \odot \mathrm{e}_1)g \qquad \text{and} \qquad E^s v = 0,$$

we may proceed analogously to Step 1 in the proof of Lemma 5.10, to get for $i = 2, \dots, d$,

$$\nabla \mathcal{W}u(x)_1^i = -\partial_i Eu(x)^1 = (-\partial_i g(x), 0, \dots, 0),$$

where as before $\mathcal{W}u$ is the skew-symmetric part of ∇u. This gives that $\mathcal{W}u_1^i$ and hence also $\partial_i g$ only depend on the first component x_1 of x, $\partial_i g(x) = p_i(x_1)$ say. Define

$$h(x) := g(x) - p_2(x_1)x_2 - \cdots - p_d(x_1)x_d$$

and observe that $\partial_i h \equiv 0$ for $i = 2, \dots, d$. Hence we may write $h(x) = h(x_1)$ and have now decomposed g as

$$g(x) = h(x_1) + p_2(x_1)x_2 + \cdots + p_d(x_1)x_d. \tag{5.12}$$

Step 2. For v only from $\mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$, but still $a = \mathrm{e}_1$, we use a smoothing argument very similar to Step 2 in the proof to Lemma 5.10 together with the first step, to see that

$$\begin{aligned} Ev = (\mathrm{e}_1 \odot \mathrm{e}_1)\Big[\mu \otimes \mathcal{L}^{d-1} &+ \gamma_2 \otimes (x_2 \mathcal{L}_{x_2}) \otimes \mathcal{L}^{d-2} \\ &+ \gamma_3 \otimes \mathcal{L}^d \otimes (x_3 \mathcal{L}_{x_3}) \otimes \mathcal{L}^{d-3} \\ &+ \cdots \\ &+ \gamma_d \otimes \mathcal{L}^{d-2} \otimes (x_d \mathcal{L}_{x_d})\Big], \end{aligned} \tag{5.13}$$

where $\mu, \gamma_2, \dots, \gamma_d \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R})$ are signed measures. In fact, mollify v to get $v_\delta \in \mathrm{C}^\infty(\mathbb{R}^d; \mathbb{R}^d)$ and apply Step 1 to the v_δ to see that $Ev_\delta = (\mathrm{e}_1 \odot \mathrm{e}_1)g_\delta \mathcal{L}^d$ with g_δ of the form exhibited in (5.12). Then use test functions of the form

$$\varphi(x_1)\mathbb{1}_{Q^d(R)}(x), \qquad \varphi(x_1)x_2\mathbb{1}_{Q^d(R)}(x), \qquad \dots, \qquad \varphi(x_1)x_d\mathbb{1}_{Q^d(R)}(x)$$

for $\varphi \in \mathrm{C}_c((-R, R); [-1, 1])$, $R > 0$, in a similar argument as before to see that all parts of the measures $(\mathrm{e}_1 \odot \mathrm{e}_1)g_\delta \mathcal{L}^d$ converge separately. Thus, $Ev = \text{w*-}\lim_{\delta \downarrow 0} Ev_\delta$ has the form (5.13).

Let $y_0 \in \mathbb{R}^d$ be such that there exists another (non-zero) singular tangent Young measure $\kappa \in \mathbf{BDY}_{\mathrm{loc}}(\mathbb{R}^d)$ to σ at y_0 (in the sense of Proposition 3.6). Since then $[\kappa] \in \mathrm{Tan}(Ev, y_0)$ and all parts of Ev are smooth in the variables $x_2, \dots, x_d$ by (5.13), every tangent measure will be constant in these variables (one can see this for example by testing the blow-up sequence with tensor products of $\mathrm{C}_c(\mathbb{R})$-functions). Hence, $[\kappa]$ can be written in the form

$$[\kappa] = \tilde{\mu} \otimes \mathcal{L}^{d-1}$$

for some $\tilde{\mu} \in \mathbf{M}_{\mathrm{loc}}(\mathbb{R})$. As before we have that κ is also a singular tangent Young measure to ν at the point x_0.

Step 3. We may now argue similarly to the proof of part (ii) of the theorem in the previous section to get that there exists $h \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$ as well as $\tilde{v}_0 \in \mathbb{R}^d$ and a skew-symmetric matrix $\tilde{R} \in \mathbb{R}^{d\times d}_{\mathrm{skew}}$ with

$$\tilde{v}(x) = \tilde{v}_0 + h(x_1)\mathrm{e}_1 + \tilde{R}x.$$

This shows the claim of case (iii) of the theorem for $a = \mathrm{e}_1$.

Step 4. For general a, we use a transformation like in Step 2 of the proof of case (ii) in the previous section. $\square$

5.5 Jensen-type inequalities

In this section we establish the following necessary conditions for BD-Young measures, which will later yield general lower semicontinuity and relaxation results as corollaries.

Theorem 5.11 (Jensen-type inequalities). *Let $\nu \in \mathbf{BDY}(\Omega)$ be a BD-Young measure. Then, for all symmetric-quasiconvex $h \in \mathrm{C}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ with linear growth at infinity it holds that*

$$h\biggl(\langle \mathrm{id}, \nu_{x_0}\rangle + \langle \mathrm{id}, \nu_{x_0}^\infty\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0)\biggr) \le \langle h, \nu_{x_0}\rangle + \langle h^\#, \nu_{x_0}^\infty\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0)$$

for $\mathcal{L}^d$-almost every $x_0 \in \Omega$, and

$$h^\#\bigl(\langle \mathrm{id}, \nu_{x_0}^\infty\rangle\bigr) \le \langle h^\#, \nu_{x_0}^\infty\rangle$$

for λ_ν^s-almost every $x_0 \in \Omega$.

The proof is contained in Lemmas 5.12 and 5.13 below (notice that if h is symmetric-quasiconvex, then so is its generalized recession function $h^\#$).

5.5.1 Jensen-type inequality at regular points

The proof at regular points is straightforward.

Lemma 5.12. *Let $\nu \in \mathbf{BDY}(\Omega)$ be a BD-Young measure. Then, for $\mathcal{L}^d$-a.e. $x_0 \in \Omega$ it holds that*

$$h\bigg(\langle \mathrm{id}, \nu_{x_0}\rangle + \langle \mathrm{id}, \nu_{x_0}^\infty\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0)\bigg) \le \langle h, \nu_{x_0}\rangle + \langle h^\#, \nu_{x_0}^\infty\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0)$$

for all symmetric-quasiconvex $h \in \mathrm{C}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$ with linear growth at infinity.

Proof. Use Proposition 3.5 to get a regular tangent Young measure $\sigma \in \mathbf{BDY}(\mathbb{B}^d)$ to ν at a suitable $x_0 \in \Omega$ (this is possible for $\mathcal{L}^d$-almost every $x_0 \in \Omega$). With

$$A_0 := \langle \mathrm{id}, \nu_{x_0}\rangle + \langle \mathrm{id}, \nu_{x_0}^\infty\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0),$$

it holds that $[\sigma] = A_0\, \mathcal{L}^d \llcorner \mathbb{B}^d$. From Proposition 2.30 (ii) take a sequence $(v_n) \subset (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(\mathbb{B}^d;\mathbb{R}^d)$ with $Ev_n \xrightarrow{\mathbf{Y}} \sigma$ in $\mathbf{Y}(\mathbb{B}^d;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ and $v_n(x) = A_0 x$ on $\mathbb{S}^{d-1}$. Since the function h is quasiconvex,

$$h(A_0) \le \fint_{\mathbb{B}^d} h(\mathcal{E}v_n)\,\mathrm{d}x \qquad \text{for all } n \in \mathbb{N}.$$

By virtue of Lemma 2.2, we get a sequence $(\mathbb{1} \otimes h_k) \subset \mathbf{E}(\mathbb{B}^d;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ with $h_k \downarrow h$, $h_k^\infty \downarrow h^\#$ pointwise and $\sup_k \|\mathbb{1} \otimes h_k\|_{\mathbf{E}(\mathbb{B}^d;\mathbb{R}^{d\times d}_{\mathrm{sym}})} < \infty$. Thus, for all $k \in \mathbb{N}$,

$$\begin{aligned} h(A_0) &\le \limsup_{n\to\infty} \fint_{\mathbb{B}^d} h(\mathcal{E}v_n)\,\mathrm{d}x \le \lim_{n\to\infty} \fint_{\mathbb{B}^d} h_k(\mathcal{E}v_n)\,\mathrm{d}x \\ &= \frac{1}{\omega_d}\langle\!\langle \mathbb{1} \otimes h_k, \sigma\rangle\!\rangle = \langle h_k, \nu_{x_0}\rangle + \langle h_k^\infty, \nu_{x_0}^\infty\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x_0), \end{aligned}$$

where the last equality follows from (3.12). Now let $k \to \infty$ and invoke the Monotone Convergence Theorem to conclude. □

5.5.2 Jensen-type inequality at singular points

We now show a Jensen-type inequality at singular points, utilizing the good blow-ups from Theorem 5.6. At points where this (good) blow-up is affine, the proof is a straightforward application of the symmetric-quasiconvexity. At (almost all) other points, we can decompose the blow-up into one or two one-directional functions and an affine part (cf. Figure 5.1). This special structure allows us to average the functions into an affine function, which then allows the application of quasiconvexity, see Figure 5.3 for an illustration of this averaging procedure.

Lemma 5.13. *Let $\nu \in \mathbf{BDY}(\Omega)$ be a BD-Young measure. Then, for λ_ν^s-almost every $x_0 \in \Omega$ it holds that*

$$g\big(\langle \mathrm{id}, \nu_{x_0}^\infty\rangle\big) \le \langle g, \nu_{x_0}^\infty\rangle$$

for all symmetric-quasiconvex and positively 1-homogeneous $g \in \mathrm{C}(\mathbb{R}^{d\times d}_{\mathrm{sym}})$.

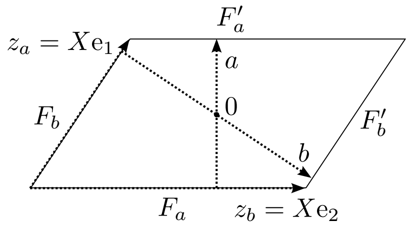

Figure 5.2: Parallelotope notation.

Proof. Theorem 5.6 on the existence of good blow-ups yields the existence of a singular tangent Young measure $\sigma \in \mathbf{BDY}_{\mathrm{loc}}(\mathbb{R}^d)$ to ν (in the sense of Proposition 3.6) at λ_ν^s-almost every $x_0 \in \Omega$. Let $[\sigma] = Ev$ for some $v \in \mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^d)$ and define

$$A_0 := \bigl\langle \mathrm{id}, \nu_{x_0}^\infty \bigr\rangle.$$

Observe that by (3.16), $Ev = [\sigma] = A_0\lambda_\sigma$. Moreover, depending on the value of A_0, one of the cases (i), (ii), (iii) in Theorem 5.6 holds.

Case 1: $A_0 \notin \{\, a \odot b \ \textbf{:}\ a, b \in \mathbb{R}^d \setminus \{0\} \,\}$ *(possibly* $A_0 = 0$*).*
By Theorem 5.6 (i), v is affine, and multiplying v by a constant, we may assume without loss of generality that $Ev = A_0\,\mathcal{L}^d \llcorner \mathbb{B}^d$. Adding a rigid deformation if necessary, we may also suppose $v(x) = A_0 x$. Now restrict σ to the unit ball $\mathbb{B}^d$ and by virtue of Proposition 2.30 (ii) take a sequence $(v_n) \subset (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(\mathbb{B}^d; \mathbb{R}^d)$ with $Ev_n \xrightarrow{\mathbf{Y}} \sigma$ in $\mathbf{Y}(\mathbb{B}^d; \mathbb{R}^{d\times d}_{\mathrm{sym}})$ and $v_n(x) = A_0 x$ on $\mathbb{S}^{d-1}$. Since g is quasiconvex,

$$g(A_0) \leq \fint_{\mathbb{B}^d} g(\mathcal{E}v_n) \,\mathrm{d}x \qquad \text{for all } n \in \mathbb{N}.$$

Finally, we may use (3.16) to get

$$g(A_0) \leq \limsup_{n\to\infty} \fint_{\mathbb{B}^d} g(\mathcal{E}v_n) \,\mathrm{d}x = \frac{1}{\omega_d} \bigl\langle\!\bigl\langle \mathbb{1} \otimes g, \sigma \bigr\rangle\!\bigr\rangle = \bigl\langle g, \nu_{x_0}^\infty \bigr\rangle.$$

This proves the claim in this case.

Case 2: $A_0 = q(a \odot b)$*, where* $a, b \in \mathbb{S}^{d-1}, q \in \mathbb{R} \setminus \{0\}$ *and* $a \neq b$*.*
Let P be an open unit parallelotope with mid-point at the origin and with two face normals a, b. The other face normals are orthogonal to a and b, yet otherwise arbitrary, i.e. if $\xi_3, \ldots, \xi_d \in \mathbb{S}^{d-1}$ extend a, b to a basis of $\mathbb{R}^d$ and satisfy $\xi_3, \ldots, \xi_d \perp \operatorname{span}\{a, b\}$, then

$$P = \bigl\{\, x \in \mathbb{R}^d \ \textbf{:}\ |x \cdot a|, |x \cdot b|, |x \cdot \xi_3|, \ldots, |x \cdot \xi_d| \leq \tfrac{1}{2} \,\bigr\}.$$

We also set $P(x_0, r) := x_0 + rP$, where $x_0 \in \mathbb{R}^d$, $r > 0$. Put all the principal vectors of P (i.e. the vectors lying in the edges) as columns into the matrix $X \in \mathbb{R}^{d\times d}$. See Figure 5.2 for notation.

By Theorem 5.6 (ii), there exist functions $h_1, h_2 \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$, a vector $v_0 \in \mathbb{R}^d$, and a skew-symmetric matrix $R \in \mathbb{R}^{d\times d}_{\mathrm{skew}}$ such that

$$v(x) = v_0 + h_1(x \cdot a)b + h_2(x \cdot b)a + Rx. \tag{5.14}$$

Without loss of generality we may assume that $v_0 = 0$ and $R = 0$. Moreover, we may additionally suppose that

$$\lambda_\sigma(P) > 0 \qquad \text{and} \qquad \lambda_\sigma(\partial P) = 0. \tag{5.15}$$

This can be achieved by taking a larger parallelotope $P' = tP \supset P$ ($t > 1$) with $\lambda_\sigma(P') > 0$, $\lambda_\sigma(\partial P') = 0$ if necessary, and then modifying the blow-up radii $r_n \downarrow 0$ to $r_n' := tr_n$.

Let $F_a, F_a' \subset \partial P$ be the two faces of P with normal a and such that F_a lies in the affine hyperplane $H_a - a/2$, where $H_a := \{ x \in \mathbb{R}^d \;:\; x \cdot a = 0 \}$. Likewise define F_b, F_b' and also $F_3, F_3', \ldots, F_d, F_d'$ for the remaining parallel face pairs. Then, the special form (5.14) of v and the observation that the vectors $z_a = X\mathrm{e}_1$, $z_b = X\mathrm{e}_2$ (say) with $F_a' = F_a + z_a$, $F_b' = F_b + z_b$ satisfy

$$z_a \perp b \qquad \text{and} \qquad z_b \perp a,$$

together yield

$$v|_{F_a'} - v|_{F_a}(\,\cdot\, - z_a) \equiv q_1 b, \qquad v|_{F_b'} - v|_{F_b}(\,\cdot\, - z_b) \equiv q_2 a$$

where $q_1 = h_1(1/2 - 0) - h_1(-1/2 + 0) = Dh_1((-1/2, 1/2))$ and $q_2 = Dh_2((-1/2, 1/2))$. Also,

$$v|_{F_k'} - v|_{F_k}(\,\cdot\, - X\mathrm{e}_k) \equiv 0 \qquad \text{for } k = 3, \ldots, d.$$

By the chain rule in BV,

$$Ev(P) = (q_1 + q_2) a \odot b,$$

but on the other hand from the properties of σ, see (3.14), we have

$$Ev(P) = [\sigma](P) = \langle \mathrm{id}, \nu_{x_0}^\infty \rangle \lambda_\sigma(P) = A_0 \lambda_\sigma(P) = q(a \odot b)\lambda_\sigma(P),$$

and so in particular

$$q \cdot \lambda_\sigma(P) = q_1 + q_2.$$

By virtue of the Boundary Adjustment Proposition 2.30 (ii), we take a sequence $(v_n) \subset (\mathrm{W}^{1,1} \cap \mathrm{C}^\infty)(P; \mathbb{R}^d)$ with $v_n|_{\partial P} = v|_{\partial P}$ such that $Ev_n \overset{\mathbf{Y}}{\to} \sigma$ in $\mathbf{Y}(P; \mathbb{R}^{d\times d}_{\mathrm{sym}})$. Extend v_n to all of $\mathbb{R}^d$ by periodicity (with respect to the periodicity cell P) and define

$$w_n(x) := v_n(x) + q_1 \left\lfloor x \cdot a + \frac{1}{2} \right\rfloor b + q_2 \left\lfloor x \cdot b + \frac{1}{2} \right\rfloor a, \qquad x \in P.$$

Clearly, $(w_n) \subset \mathrm{BD}(P)$ and one checks that the Ew_n in fact do not charge the gluing surfaces. Indeed, the size of the jump incurred over the boundary of each copy of P from the gluing of the v_n is exactly compensated by the staircase function. For example, over each (F_a, F_a')-interface, the first term in the definition of w_n incurs a jump of magnitude

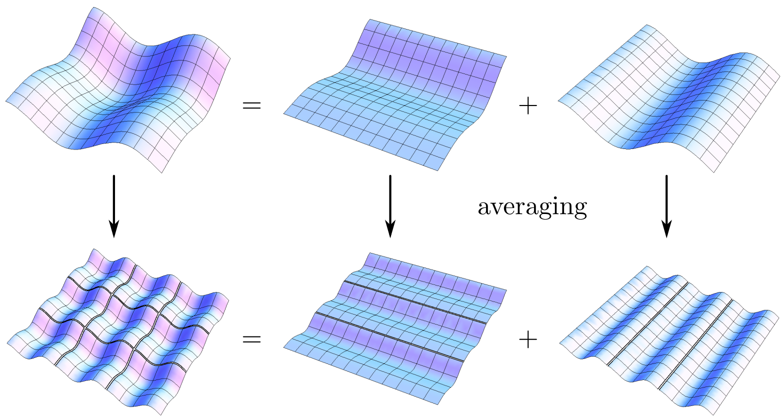

Figure 5.3: Staircase construction for the singular Jensen-type inequality.

$-q_1 b$, but at the same time the staircase term gives a jump of size $q_1 b$ over the same gluing interface, whence in w_n no jump remains. Thus, also extending w_n to $\mathbb{R}^d$ by P-periodicity, $(w_n) \subset \mathrm{LD}_{\mathrm{loc}}(\mathbb{R}^d)$. Now set

$$u_n(x) := \frac{w_n(nx)}{n} \qquad x \in P,$$

which lies in $\mathrm{LD}(P)$ and satisfies

$$\mathcal{E}u_n(x) = \sum_{z \in \{0,\ldots,n-1\}^d} \mathcal{E}v_n(nx - Xz) \mathbb{1}_{P(Xz/n, 1/n)}(x).$$

Next, we show that for some skew-symmetric matrix $R_0 \in \mathbb{R}^{d \times d}_{\mathrm{skew}}$,

$$u_n \to (\lambda_\sigma(P) A_0 + R_0)x \qquad \text{in } \mathrm{L}^1(P; \mathbb{R}^d).$$

To see this, first observe

$$\left\| \frac{v_n(nx)}{n} \right\|_{\mathrm{L}^1(P;\mathbb{R}^d)} = \frac{1}{n} \|v_n\|_{\mathrm{L}^1(P;\mathbb{R}^d)} \quad \to \quad 0 \qquad \text{as } n \to \infty$$

by a change of variables. On the other hand,

$$\frac{1}{n}\left(q_1 \left\lfloor nx \cdot a + \frac{1}{2} \right\rfloor b + q_2 \left\lfloor nx \cdot b + \frac{1}{2} \right\rfloor a \right) \quad \to \quad \big[q_1 (b \otimes a) + q_2 (a \otimes b)\big] x$$

uniformly. The symmetric part of the matrix on the right hand side is $(q_1 + q_2) a \odot b = \lambda_\sigma(P) A_0$ and so the claim follows. Subtracting $R_0 x$ from v_n, v, we may even assume that $R_0 = 0$.

We can now use Proposition 2.30 (ii) again to get a sequence $(\tilde{u}_n) \subset \mathrm{LD}(P; \mathbb{R}^d)$ satisfying $\tilde{u}_n(x) = \lambda_\sigma(P) A_0 x$ on ∂P such that for all g as in the statement of the lemma,

$$\lim_{n\to\infty} \int_P g(\mathcal{E}\tilde{u}_n) \, \mathrm{d}x = \lim_{n\to\infty} \int_P g(\mathcal{E}u_n) \, \mathrm{d}x,$$

by using the fact that $(E\tilde{u}_n)$ and (Eu_n) generate the same (unnamed) Young measure. The boundary conditions of $\tilde{u}_n$ together with the quasiconvexity of g imply ($|P| = 1$)

$$g\big(\lambda_\sigma(P)A_0\big) \le \int_P g(\mathcal{E}\tilde{u}_n)\,\mathrm{d}x \qquad \text{for all } n \in \mathbb{N}.$$

This allows us to calculate

$$\begin{aligned}
\lambda_\sigma(P)g(A_0) &\le \lim_{n\to\infty} \int_P g(\mathcal{E}\tilde{u}_n)\,\mathrm{d}x = \lim_{n\to\infty} \int_P g(\mathcal{E}u_n)\,\mathrm{d}x \\
&= \lim_{n\to\infty} \sum_{z\in\{0,\ldots,n-1\}^d} \int_{P(Xz/n,1/n)} g\big(\mathcal{E}v_n(nx - Xz)\big)\,\mathrm{d}x \\
&= \lim_{n\to\infty} \sum_{z\in\{0,\ldots,n-1\}^d} \frac{1}{n^d} \int_P g(\mathcal{E}v_n)\,\mathrm{d}y \\
&= \lim_{n\to\infty} \int_P g(\mathcal{E}v_n)\,\mathrm{d}y = \big\langle\!\big\langle \mathbb{1}_P \otimes g, \sigma \big\rangle\!\big\rangle = \big\langle g, \nu_{x_0}^\infty \big\rangle \lambda_\sigma(P),
\end{aligned}$$

where the two last equalities follow from (3.16) in conjunction with (5.15). Hence we have also shown the claim in this case.

Case 3: $A_0 = q(a \odot a)$, where $a \in \mathbb{S}^{d-1}, q \in \mathbb{R}$.
This case follows analogously to the previous one, but using (rotated) cubes, of which we only prescribe one face normal a instead of a, b, and with

$$v(x) = v_0 + h(x \cdot a)a + Rx.$$

in place of (5.14) by Theorem 5.6 (iii). □

5.6 Lower semicontinuity and relaxation

The Jensen-type inequalities from the previous Theorem 5.11 can now be employed to show the main lower semicontinuity result in the space $\mathrm{BD}(\Omega)$:

Proof of Theorem 5.1. Let $u_j \overset{*}{\rightharpoonup} u$ in $\mathrm{BD}(\Omega)$ and consider u_j, u to be extended by zero to $\mathbb{R}^d$. Assume also, taking a subsequence if necessary, that $Eu_j \overset{\mathbf{Y}}{\to} \nu$ in $\mathbf{BDY}(\mathbb{R}^d)$. The operation of taking subsequences does not impede our aim to prove lower semicontinuity since we will show an inequality for all such subsequences, which then also holds for the original sequence.

For the barycenter of ν we have

$$[\nu] = Eu \llcorner \Omega + (u|_{\partial\Omega} \odot n_\Omega)\,\mathcal{H}^{d-1} \llcorner \partial\Omega.$$

Denote by λ_ν^* the singular part of λ_ν with respect to $|E^s u| + |u|\,\mathcal{H}^{d-1} \llcorner \partial\Omega$, i.e. λ_ν^* is concentrated in an $(|E^s u| + |u|\,\mathcal{H}^{d-1} \llcorner \partial\Omega)$-negligible set. We compute

$$\begin{aligned}
&\big\langle \mathrm{id}, \nu_x \big\rangle + \big\langle \mathrm{id}, \nu_x^\infty \big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x) = \frac{\mathrm{d}[\nu]}{\mathrm{d}\mathcal{L}^d}(x) = \begin{cases} \mathcal{E}u(x) & \text{for } \mathcal{L}^d\text{-a.e. } x \in \Omega, \\ 0 & \text{for } \mathcal{L}^d\text{-a.e. } x \in \mathbb{R}^d \setminus \Omega, \end{cases} \\
&\frac{\langle \mathrm{id}, \nu_x^\infty \rangle}{|\langle \mathrm{id}, \nu_x^\infty \rangle|} = \frac{\mathrm{d}[\nu]^s}{\mathrm{d}|[\nu]^s|}(x) = \begin{cases} \dfrac{\mathrm{d}E^s u}{\mathrm{d}|E^s u|}(x) & \text{for } |E^s u|\text{-a.e. } x \in \Omega, \\ \dfrac{u|_{\partial\Omega}(x) \odot n_\Omega(x)}{|u|_{\partial\Omega}(x) \odot n_\Omega(x)|} & \text{for } |u|\,\mathcal{H}^{d-1}\text{-a.e. } x \in \partial\Omega, \end{cases} \\
&\big\langle \mathrm{id}, \nu_x^\infty \big\rangle = 0 \qquad \text{for } \lambda_\nu^*\text{-a.e. } x \in \mathbb{R}^d, \\
&\big|\big\langle \mathrm{id}, \nu_x^\infty \big\rangle\big|\, \lambda_\nu^s = |E^s u| + \big|u|_{\partial\Omega} \odot n_\Omega\big|\, \mathcal{H}^{d-1} \llcorner \partial\Omega, \\
&\big\langle \mathrm{id}, \nu_x \big\rangle = 0 \qquad \text{for } x \in \mathbb{R}^d \setminus \overline{\Omega}, \\
&\lambda_\nu \llcorner (\mathbb{R}^d \setminus \overline{\Omega}) = 0.
\end{aligned}$$

Moreover, consider f to be extended to $\mathbb{R}^d \times \mathbb{R}^{d\times d}_{\mathrm{sym}}$ as follows: first extend f^∞ restricted to $\overline{\Omega} \times \partial\mathbb{B}^{d\times d}_{\mathrm{sym}}$ continuously to $\mathbb{R}^d \times \partial\mathbb{B}^{d\times d}_{\mathrm{sym}}$ (where $\mathbb{B}^{d\times d}_{\mathrm{sym}} := \mathbb{B}^{d\times d} \cap \mathbb{R}^{d\times d}_{\mathrm{sym}}$) and then set $f(x, A) := f^\infty(x, A)$ for $x \in \mathbb{R}^d \setminus \overline{\Omega}$ and $A \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$. Hence, the so extended f is still a Carathéodory function, f^∞ is jointly continuous and $f(x, 0) = 0$ for all $x \in \mathbb{R}^d \setminus \overline{\Omega}$. The extended representation result for generalized Young measures, Proposition 2.17 together with Theorem 5.11 yields

$$\begin{aligned}
\liminf_{j\to\infty} \mathcal{F}(u_j) &= \int \big\langle f(x, \,\bullet\,), \nu_x \big\rangle + \big\langle f^\infty(x, \,\bullet\,), \nu_x^\infty \big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x) \,\mathrm{d}x \\
&\qquad + \int \big\langle f^\infty(x, \,\bullet\,), \nu_x^\infty \big\rangle \,\mathrm{d}\lambda_\nu^s(x) \\
&\geq \int f\Big(x, \big\langle \mathrm{id}, \nu_x \big\rangle + \big\langle \mathrm{id}, \nu_x^\infty \big\rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x)\Big) \,\mathrm{d}x \\
&\qquad + \int f^\infty\big(x, \big\langle \mathrm{id}, \nu_x^\infty \big\rangle\big) \,\mathrm{d}\lambda_\nu^s(x) \\
&= \mathcal{F}(u).
\end{aligned}$$

Hence we have established lower semicontinuity. □

Remark 5.14. Symmetric-quasiconvexity is also necessary for weak* lower semicontinuity since it is already necessary for weak* lower semicontinuity of $\mathcal{F}$ restricted to $\mathrm{W}^{1,\infty}(\Omega; \mathbb{R}^d)$, which is a subspace of $\mathrm{BD}(\Omega)$.

Remark 5.15 (Recession functions). Notice that we needed to require the existence of the strong recession function f^∞ in the previous result and could not just use the generalized recession function $f^\#$. Unfortunately, this cannot be avoided as long as no Alberti-type theorem is available in BD. The reason is that for lower semicontinuity the *lower* generalized recession function

$$f_\#(x, A) := \liminf_{\substack{x'\to x \\ A'\to A \\ t\to\infty}} \frac{f(tA')}{t}, \qquad x \in \overline{\Omega},\ A \in \mathbb{R}^{d\times d}_{\mathrm{sym}}.$$

would be the natural choice of recession function since for $f_\#$ it still holds that

$$\liminf_{j\to\infty} \mathcal{F}(u_j) \geq \int \bigl\langle f(x,\,\cdot\,), \nu_x \bigr\rangle \,\mathrm{d}x + \int \bigl\langle f_\#(x,\,\cdot\,), \nu_x^\infty \bigr\rangle \,\mathrm{d}\lambda_\nu(x),$$

see Theorem 2.5 (iii) in [3] (recall that f is Lipschitz continuous by quasiconvexity). The problem with that choice, however, is that we cannot easily ascertain that $f_\#$ is symmetric-quasiconvex. For f such that we know a-priori that $f_\#$ is symmetric-quasiconvex, the above theorem also holds with $f_\#$ in place of f^∞, see Remark 4.14 for details.

Remark 5.16 (Dirichlet boundary conditions). The trace operator is not weakly* continuous in $\mathrm{BD}(\Omega)$, so boundary conditions are in general not preserved under this convergence and we need to switch to a suitable relaxed formulation of Dirichlet boundary conditions. However, since for linear growth integrands all parts of the symmetrized derivative may interact, this constraint is not easily formulated and is probably only meaningful in connection with concrete problems. Some results for special BD-functions can be found in [23], Chapter II.8 of [122] (also see Proposition II.7.2) treats the case where additionally divergences converge weakly; also see [86] for a related approach. Finally, Section 14 of [67] contains general remarks on boundary conditions for linear growth functionals.

We immediately also have the following relaxation theorem.

Corollary 5.17 (Relaxation). *Let $f \in \mathbf{R}(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}})$ be symmetric-quasiconvex in its second argument. Then, the lower semicontinuous envelope of the functional*

$$\int_\Omega f\bigl(x, \mathcal{E}u(x)\bigr) \,\mathrm{d}x + \int_{\partial\Omega} f^\infty\bigl(x, u(x) \odot n_\Omega(x)\bigr) \,\mathrm{d}\mathcal{H}^{d-1}(x), \qquad u \in \mathrm{LD}(\Omega),$$

with respect to weak convergence in* $\mathrm{BD}(\Omega)$ *is the functional* $\mathcal{F}$ *from Theorem 5.1.*

Of course, for $f \geq 0$, we again may omit the boundary term.

Proof. Denote the $\mathcal{G}$ the functional defined in the statement of the corollary and let $\mathcal{G}_*$ be its weakly* (sequentially) lower semicontinuous envelope. By Proposition 5.3, $\mathcal{F}$ is the $\langle\,\cdot\,\rangle$-strictly continuous extension of $\mathcal{G}$ to $\mathrm{BD}(\Omega)$, in particular $\mathcal{G}_* \leq \mathcal{F}$. On the other hand, $\mathcal{F}$ is weakly* lower semicontinuous, hence also $\mathcal{F} \leq \mathcal{G}_*$. □

Remark 5.18. Of course it would be desirable to have a relaxation theorem for integrands f that are not symmetric-quasiconvex. Then, the relaxed functional should be $\mathcal{F}$ from Theorem 5.1, but with f replaced by its symmetric-quasiconvex envelope $\mathrm{SQ}f$. However, we do not know whether $(\mathrm{SQ}f)^\infty$ exists, and without an Alberti-type theorem in BD we cannot show lower semicontinuity for the functional with $(\mathrm{SQ}f)^\infty$ replaced by $(\mathrm{SQ}f)^\#$ within our framework, see the remarks above.

Remark 5.19. Just like for the space BV (see Remark 4.16), it should be remarked that some parts of the proof could also be reformulated in a more elementary fashion, circumventing the machinery of Young measures. However, without the use of tangent Young measures and working with blow-up sequences directly, several arguments would require additional technical steps. Particularly the construction of "good" blow-ups through the "iterated blow-up" trick in Theorem 5.6 is not easily formulated with mere sequences instead of tangent Young measures. At the core of this lies the fact that in the blow-up technique, we are not primarily interested in the blow-up *limit*, but in the behavior of the blow-up *sequence*, just as represented in a (generalized) Young measure limit. This is precisely the idea behind the concept of tangent Young measures and the localization principles, Propositions 3.5 and 3.6.

5.7 Rigidity in 2D

We end this chapter by illustrating the rigidity arguments of Section 5.4 in a more concrete situation. In the following we give a complete analysis of solutions for the differential inclusion

$$\mathcal{E}u \in \operatorname{span}\{P\} \qquad \text{pointwise a.e.,} \qquad u \in \mathrm{LD}_{\mathrm{loc}}(\mathbb{R}^2), \tag{5.16}$$

for a *fixed* symmetric matrix $P \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}$. The results presented here are not needed anywhere else in this thesis, so for convenience we restrict our analysis to the space $\mathrm{LD}_{\mathrm{loc}}(\mathbb{R}^2)$ and omit extensions to $\mathrm{BD}_{\mathrm{loc}}(\mathbb{R}^2)$.

First notice that we may always reduce the above problem to an equivalent differential inclusion with P diagonal. Indeed, let $Q \in \mathbb{R}^{2\times 2}$ be an orthogonal matrix such that

$$QPQ^T = \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} =: \tilde{P}, \qquad \lambda_1, \lambda_2 \in \mathbb{R}.$$

Clearly, $u \in \mathrm{LD}_{\mathrm{loc}}(\mathbb{R}^2)$ solves (5.16) if and only if $\tilde{u}(x) := Qu(Q^T x)$ solves

$$\mathcal{E}\tilde{u} \in \operatorname{span}\{\tilde{P}\} \qquad \text{pointwise a.e.,} \qquad \tilde{u} \in \mathrm{LD}_{\mathrm{loc}}(\mathbb{R}^2),$$

hence we may always assume that P in (5.16) is already diagonal.

According to Lemma 5.4 we have three non-trivial cases to take care of, corresponding to the signs of the eigenvalues λ_1, λ_2; the trivial case $\lambda_1 = \lambda_2 = 0$, i.e. $P = 0$, was already settled in Lemma 5.2.

We will formulate our results on solvability of (5.16) in terms of conditions on $g \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^2)$ in the differential equation

$$\mathcal{E}u = Pg \quad \text{a.e.,} \qquad u \in \mathrm{LD}_{\mathrm{loc}}(\mathbb{R}^2).$$

With g as an additional unknown this is clearly equivalent to (5.16).

First, consider the situation that $\lambda_1, \lambda_2 \neq 0$ and that these two eigenvalues have opposite signs. Then, from (the proof of) Lemma 5.4, we know that $P = a \odot b$ ($a \neq b$) for

$$a := \begin{pmatrix} \gamma \\ 1 \end{pmatrix}, \qquad b := \begin{pmatrix} \lambda_1 \gamma^{-1} \\ \lambda_2 \end{pmatrix}, \qquad \text{where} \qquad \gamma := \sqrt{-\frac{\lambda_1}{\lambda_2}}.$$

The result about solvability of (5.16) for this choice of P is:

Proposition 5.20 (Rigidity for $P = a \odot b$). *Let $P = \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} = a \odot b$, where $\lambda_1, \lambda_2 \in \mathbb{R}$ have opposite signs. Then, there exists a function $u \in \mathrm{LD}_{\mathrm{loc}}(\mathbb{R}^2)$ solving the differential equation*

$$\mathcal{E}u = Pg \quad a.e.$$

if and only if $g \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^2)$ is of the form

$$g(x) = h_1(x \cdot a) + h_2(x \cdot b), \qquad x \in \mathbb{R}^2 \text{ a.e.},$$

where $h_1, h_2 \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$. In this case,

$$u(x) = u_0 + H_1(x \cdot a)b + H_2(x \cdot b)a + Rx, \qquad x \in \mathbb{R}^2 \text{ a.e.},$$

with $u_0 \in \mathbb{R}^2$, $R \in \mathbb{R}^{d \times d}_{\mathrm{skew}}$ and $H_1, H_2 \in \mathrm{W}^{1,1}_{\mathrm{loc}}(\mathbb{R})$ satisfying $H_1' = h_1$ and $H_2' = h_2$.

Proof. This follows directly from Lemma 5.9 and some elementary computations. □

In the case $\lambda_1 \neq 0$, $\lambda_2 = 0$, i.e. $P = \lambda_1(\mathrm{e}_1 \odot \mathrm{e}_1)$, one could guess by analogy to the previous case that if $u \in \mathrm{LD}_{\mathrm{loc}}(\mathbb{R}^2)$ satisfies $\mathcal{E}u = Pg$ for some $g \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$, then u and g should only depend on x_1 up to a rigid deformation. This, however, is *false*, as can be seen from the following counterexample.

Counterexample 5.21. Consider

$$P := \begin{pmatrix} 1 & \\ & 0 \end{pmatrix}, \qquad u(x) := \begin{pmatrix} 4x_1^3 x_2 \\ -x_1^4 \end{pmatrix}, \qquad g(x) := 12x_1^2 x_2.$$

Then, u satisfies $\mathcal{E}u = Pg$, but neither u nor g only depend on x_1.

One can, however, show the following result:

Proposition 5.22 (Rigidity for $P = a \odot a$). *Let $P = \begin{pmatrix} \lambda_1 & \\ & 0 \end{pmatrix} = \lambda_1(\mathrm{e}_1 \odot \mathrm{e}_1)$. Then, there exists a function $u \in \mathrm{LD}_{\mathrm{loc}}(\mathbb{R}^2)$ solving the differential equation*

$$\mathcal{E}u = Pg \quad a.e.$$

if and only if $g \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^2)$ is of the form

$$g(x) = h(x_1) + p(x_1)x_2, \qquad x \in \mathbb{R}^2 \text{ a.e.},$$

where $h, p \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$. In this case,

$$u(x) = u_0 + \lambda_1 \begin{pmatrix} H(x_1) + \mathcal{P}'(x_1)x_2 \\ -\mathcal{P}(x_1) \end{pmatrix} + Rx, \qquad x \in \mathbb{R}^2 \text{ a.e.},$$

with $u_0 \in \mathbb{R}^2$, $R \in \mathbb{R}^{d \times d}_{\mathrm{skew}}$ and $H \in \mathrm{W}^{1,1}_{\mathrm{loc}}(\mathbb{R})$, $\mathcal{P} \in \mathrm{W}^{2,1}_{\mathrm{loc}}(\mathbb{R})$ satisfying $H' = h$ and $\mathcal{P}'' = p$.

Proof. From the arguments in Section 5.4.3 we know that whenever $u \in \mathrm{LD}_{\mathrm{loc}}(\mathbb{R}^2)$ solves the differential equation $\mathcal{E}u = Pg$, then g (and hence also u) must have the form exhibited in the statement of the proposition. Conversely, it is elementary to check that u as defined above satisfies $\mathcal{E}u = Pg$. □

Finally, we consider the case where the eigenvalues λ_1 and λ_2 are non-zero and have the same sign. Then, $P \neq a \odot b$ for any $a, b \in \mathbb{R}^2$ by Lemma 5.4. Define the differential operator

$$\mathcal{A}_P := \lambda_2 \partial_{11} + \lambda_1 \partial_{22}$$

and notice that whenever a function $g \colon \mathbb{R}^2 \to \mathbb{R}$ satisfies $\mathcal{A}_P g \equiv 0$ distributionally, then by elliptic regularity (generalized Weyl's Lemma), we have that in fact $g \in \mathrm{C}^\infty(\mathbb{R}^2)$.

Proposition 5.23 (Rigidity for $P \neq a \odot b$). *Let $P = \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix}$, where $\lambda_1, \lambda_2 \in \mathbb{R}$ have the same sign. Then, there exists a function $u \in \mathrm{LD}_{\mathrm{loc}}(\mathbb{R}^2)$ solving the differential equation*

$$\mathcal{E}u = Pg \quad \textit{a.e.}$$

if and only if $g \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^2)$ satisfies

$$\mathcal{A}_P g \equiv 0.$$

Moreover, in this case both g and u are smooth.

Proof. First assume that $g \in \mathrm{C}^\infty(\mathbb{R}^2)$ satisfies $\mathcal{A}_P g \equiv 0$. Define

$$F := \nabla g \begin{pmatrix} 0 & \lambda_2 \\ -\lambda_1 & 0 \end{pmatrix} = (-\lambda_1 \partial_2 g, \lambda_2 \partial_1 g)$$

and observe (we use $\operatorname{curl}(h_1, h_2) = \partial_2 h_1 - \partial_1 h_2$)

$$\operatorname{curl} F = -\lambda_1 \partial_{22} g - \lambda_2 \partial_{11} g = -\mathcal{A}_P g \equiv 0.$$

Hence, there exists $f \in \mathrm{C}^\infty(\mathbb{R}^2)$ with $\nabla f = F$, in particular

$$\partial_1 f = -\lambda_1 \partial_2 g, \qquad \partial_2 f = \lambda_2 \partial_1 g. \tag{5.17}$$

Put

$$\mathcal{U} := \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} g + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} f.$$

We calculate (this time we apply the curl row-wise), using (5.17),

$$\operatorname{curl} \mathcal{U} = \begin{pmatrix} \operatorname{curl}(\lambda_1 g, -f) \\ \operatorname{curl}(f, \lambda_2 g) \end{pmatrix} = \begin{pmatrix} \lambda_1 \partial_2 g + \partial_1 f \\ \partial_2 f - \lambda_2 \partial_1 g \end{pmatrix} \equiv 0. \tag{5.18}$$

Let $u \in \mathrm{C}^\infty(\mathbb{R}^2; \mathbb{R}^2)$ be such that $\nabla u = \mathcal{U}$. Then, $\mathcal{E}u = Pg$.

For the other direction, it suffices to show that $\mathcal{E}u = Pg$ implies $\mathcal{A}_P g \equiv 0$, the smoothness of u, g follows from the first step. Notice further that by a mollification argument we may in fact assume that $u \in \mathrm{C}^\infty(\mathbb{R}^2;\mathbb{R}^2)$, $g \in \mathrm{C}^\infty(\mathbb{R}^2)$ since the conditions $\mathcal{E}u = Pg$ and $\mathcal{A}_P g \equiv 0$ are preserved under smoothing. Then split the gradient into its symmetric and skew-symmetric parts,

$$\nabla u = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} g + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} f$$

for some function $f \in \mathrm{C}^\infty(\mathbb{R}^2)$. As in (5.18), this implies the conditions (5.17) for $\nabla g, \nabla f$. Hence,

$$\nabla f = \nabla g \begin{pmatrix} 0 & \lambda_2 \\ -\lambda_1 & 0 \end{pmatrix} = (-\lambda_1 \partial_2 g, \lambda_2 \partial_1 g).$$

Since the curl of ∇f vanishes, we get

$$0 \equiv \operatorname{curl} \nabla f = -\lambda_1 \partial_{22} g - \lambda_2 \partial_{11} g = -\mathcal{A}_P g,$$

so g satisfies $\mathcal{A}_P g \equiv 0$. □

Remark 5.24 (Harmonic functions). By Lemma 5.4, the simplest matrix that cannot be written as a symmetric tensor product is the identity matrix $P = I_2 = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$. In this case $\mathcal{A}_P$ is the Laplacian and the differential equation $\mathcal{E}u = I_2 g$ is solvable in $\mathrm{LD}_{\mathrm{loc}}(\mathbb{R}^2)$ if and only if g is harmonic.

Remark 5.25 (Comparison to gradients). Proposition 5.23 should be compared to the corresponding situation for gradients. If $u \in \mathrm{W}^{1,1}_{\mathrm{loc}}(\mathbb{R}^2;\mathbb{R}^2)$ satisfies

$$\nabla u \in \operatorname{span}\{P\} \qquad \text{pointwise a.e.},$$

and $\operatorname{rank} P = 2$, then necessarily u is affine, this follows from the BV-Rigidity Lemma 4.4. Notice that this behavior for the gradient is in sharp contrast to the behavior for the symmetrized gradient, as can be seen from the following example.

Example 5.26. Let

$$P := \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}, \qquad u(x) := \begin{pmatrix} \mathrm{e}^{x_1} \sin(x_2) \\ -\mathrm{e}^{x_1} \cos(x_2) \end{pmatrix}, \qquad g(x) := \mathrm{e}^{x_1} \sin(x_2).$$

Then, one can check that g is harmonic and u satisfies $\mathcal{E}u = Pg$. So, the fact that P cannot be written as a symmetric tensor product does not imply that that every solution to the differential inclusion $\mathcal{E}u \in \operatorname{span}\{P\}$ must be affine.

Appendix A

Brief history of Young measures

Laurence Chisholm Young (⋆14 July 1905 †24 December 2000) originally introduced the objects that are now called Young measures as "generalized curves/surfaces" in the late 1930s and early 1940s [124–126] to treat problems in the Calculus of Variations and Optimal Control Theory that could not be solved using classical methods. At the end of the 1960s he also wrote a book [127] (2nd edition [128]) on the topic, which explained these objects and their applications in great detail (including philosophical discussions on the semantics of the terminology "generalized curve/surface").

The basic idea of Young measures—that is to superimpose several curves/surfaces, controls, or simply functions, to form generalized solutions to non-convex problems—has turned out to be very useful. In particular, this innovative paradigm allowed a variety of new applications, where traditional methods hit a natural barrier; we refer to Young's book [127] and McShane's 1940 article [90]. More recently, starting from the seminal work of Ball & James in 1987 [18], Young measures have provided a convenient framework to describe fine phase mixtures in the theory of microstructure, also see [45, 96] and the references cited therein. Further, and more related to the present work, relaxation formulas involving Young measures for non-convex integral functionals have been developed and can for example be found in Pedregal's book [104] and also in Chapter 11 of the recent text [11]; historically, Dacorogna's lecture notes [33] were also influential.

A second avenue of development—somewhat different from Young's original intention—is to use Young measures as a *technical tool* only (that is without them appearing in the final statement of a result). This approach is in fact quite old and was probably first adopted in a series of articles by McShane from the 1940s [90–92]. There, the author first finds a generalized Young measure solution to a variational problem, then proves additional properties of the obtained Young measures, and finally concludes that these properties entail that the generalized solution is in fact classical. This idea exhibits some parallels to the hugely successful approach of first proving existence of a weak solution to a PDE and then showing that this weak solution has (in some cases) additional regularity.

After Morrey had introduced quasiconvexity in the 1950s [93] (also see his very influential book [94]), much interest in the Calculus of Variations has been centered around vector-valued problems. This development also necessitated new tools in Young measure theory, which were subsequently provided by many researchers from the 1970s onward, including Berliocchi & Lasry [25], Balder [13], Ball [17] and Kristensen [81], 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 [75, 76]. Their result places gradient Young measures in duality with quasiconvex functions via Jensen-type inequalities (another work in this direction is Sychev's article [115]). Young measures also play a role in the theory of convex integration to construct highly irregular solutions to differential inclusions. This technique was originally pioneered in a geometric context by Nash [101] and Gromov [69, 70] (without Young measures), and then adapted to a PDE setting by Müller & Šverák [97] (with Young measures) at the beginning of the new millennium. Conversely, Young measures can also be used to show regularity, see the recent work by Dolzmann & Kristensen [44]. Carstensen & Roubíček [28] considered numerical approximations.

Another branch of Young measure theory started in the late 1970s and early 1980s, when Tartar [117,118,120] and Murat [98–100] 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 [42]. A key point of this strategy is to use the good compactness properties of Young measures to pass to limits easily 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*. For an overview of this and other "weak convergence methods" see for example Evans's book [49]. It is interesting to remark that some of these ideas were already present in McShane's work cited earlier. In the context of such compensated compactness problems, extensions of Young measures that allow to pass to the limit in quadratic expressions, have been developed by Tartar [119] under the name "H-measures" and, independently, by Gérard [64], who called them "micro-local defect measures", cf. the survey articles [60, 121].

After all these developments, by the end of the 1990s, the theory of Young measure in L^p- and $\mathrm{W}^{1,p}$-spaces had reached such a degree of maturity that nowadays it is widely considered to be a standard tool in the modern theory of the Calculus of Variations. Several monographs [29,104,109] have appeared that give a comprehensive overview over the theory.

For linear-growth problems, however, the theory only really started in 1987, when DiPerna and Majda [43] used a compactification of the target space to account for concentration effects in solutions to the Euler equation (in an L^2-setting). An L^1-formulation of this theory was given in the article [3] by Alibert & Bouchitté from 1997, and since then there have been many works on generalized Young measures or "DiPerna–Majda measures", a selection is [54, 74, 84, 85, 114, 116]. The crucial advantage of generalized Young measures over their classical counterparts is that they allow to construct meaningful Young measure limits for merely L^1-bounded functions, thereby extending the reach of Young measures to concentration effects instead of mere oscillations. Alternative approaches to model concentrations are varifolds [4–6, 59] and currents [51, 52, 65, 66]. Recently, time-dependent generalized Young measures have been used for quasistatic evolution in plasticity theory by Dal Maso, DeSimone, Mora & Morini [35–37]; also see [40], where Young measures are applied to evolutionary PDEs.

More distant from the topic of this thesis, Young measures are also used in Probability Theory and Optimal Control Theory, a comprehensive recent account of the theoretical aspects is the book [29]. In particular, these applications require the extension to much more general spaces than are useful in the Calculus of Variations, a subject to which Balder contributed many results [13–16].

At the present day and after more than 70 years of development, Young measure theory constitutes a rich field of research within Analysis, which is still rapidly developing and which seems to hold much promise for future applications.

Bibliography

[1] G. Alberti, *Rank one property for derivatives of functions with bounded variation*, Proc. Roy. Soc. Edinburgh Sect. A **123** (1993), 239–274.

[2] G. Alberti, M. Csörnyei, and D. Preiss, *Structure of null sets in the plane and applications*, Proceedings of the Fourth European Congress of Mathematics (Stockholm, 2004), European Mathematical Society, 2005, pp. 3–22.

[3] J. J. Alibert and G. Bouchitté, *Non-uniform integrability and generalized Young measures*, J. Convex Anal. **4** (1997), 129–147.

[4] W. K. Allard, *On the first variation of a varifold*, Ann. of Math. **95** (1972), 417–491.

[5] F. J. Almgren, Jr., *Plateau's problem: An invitation to varifold geometry*, Benjamin, 1966.

[6] ———, *Existence and regularity almost everywhere of solutions to elliptic variational problems among surfaces of varying topological type and singularity structure*, Ann. of Math. **87** (1968), 321–391.

[7] L. Ambrosio, A. Coscia, and G. Dal Maso, *Fine properties of functions with bounded deformation*, Arch. Ration. Mech. Anal. **139** (1997), 201–238.

[8] L. Ambrosio and G. Dal Maso, *On the relaxation in* $\mathrm{BV}(\Omega;\mathbf{R}^m)$ *of quasi-convex integrals*, J. Funct. Anal. **109** (1992), 76–97.

[9] L. Ambrosio and E. De Giorgi, *Un nuovo tipo di funzionale del calcolo delle variazioni*, Atti Acc. Naz. dei Lincei, Rend. Cl. Sc. Fis. Mat. Natur. **82** (1988), 199–210.

[10] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free-Discontinuity Problems*, Oxford Mathematical Monographs, Oxford University Press, 2000.

[11] H. Attouch, G. Buttazzo, and G. Michaille, *Variational Analysis in Sobolev and BV Spaces – Applications to PDEs and Optimization*, MPS-SIAM Series on Optimization, Society for Industrial and Applied Mathematics, 2006.

[12] A. Auslender and M. Teboulle, *Asymptotic Cones and Functions in Optimization and Variational Inequalities*, Springer Monographs in Mathematics, Springer, 2003.

[13] E. J. Balder, *A general approach to lower semicontinuity and lower closure in optimal control theory*, SIAM J. Control Optim. **22** (1984), 570–598.

[14] ———, *A general denseness result for relaxed control theory*, Bull. Austral. Math. Soc. **30** (1984), 463–475.

[15] ———, *Consequences of denseness of Dirac Young measures*, J. Math. Anal. Appl. **207** (1997), 536–540.

[16] ———, *New fundamentals of Young measure convergence*, Calculus of variations and optimal control (Haifa, 1998), Res. Notes Math., vol. 411, Chapman & Hall/CRC, 2000, pp. 24–48.

[17] J. M. Ball, *A version of the fundamental theorem for Young measures*, PDEs and continuum models of phase transitions (Nice, 1988), Lecture Notes in Physics, vol. 344, Springer, 1989, pp. 207–215.

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

[19] J. M. Ball, B. Kirchheim, and J. Kristensen, *Regularity of quasiconvex envelopes*, Calc. Var. Partial Differential Equations **11** (2000), 333–359.

[20] J. M. Ball and F. Murat, *$W^{1,p}$-quasiconvexity and variational problems for multiple integrals*, J. Funct. Anal. **58** (1984), 225–253.

[21] ———, *Remarks on Chacon's biting lemma*, Proc. Amer. Math. Soc. **107** (1989), 655–663.

[22] A. C. Barroso, I. Fonseca, and R. Toader, *A relaxation theorem in the space of functions of bounded deformation*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **29** (2000), 19–49.

[23] G. Bellettini, A. Coscia, and G. Dal Maso, *Compactness and lower semicontinuity properties in* $\mathrm{SBD}(\Omega)$, Math. Z. **228** (1998), 337–351.

[24] J. Bergh and J. Löfström, *Interpolation Spaces*, Grundlehren der mathematischen Wissenschaften, vol. 223, Springer, 1976.

[25] H. Berliocchi and J.-M. Lasry, *Intégrandes normales et mesures paramétrées en calcul des variations*, Bull. Soc. Math. France **101** (1973), 129–184.

[26] M. Bildhauer, *Convex variational problems*, Lecture Notes in Mathematics, vol. 1818, Springer, 2003.

[27] G. Bouchitté, I. Fonseca, and L. Mascarenhas, *A global method for relaxation*, Arch. Ration. Mech. Anal. **145** (1998), 51–98.

[28] C. Carstensen and T. Roubíček, *Numerical approximation of Young measures in non-convex variational problems*, Numer. Math. **84** (2000), 395–415.

[29] C. Castaing, P. R. de Fitte, and M. Valadier, *Young Measures on Topological Spaces: With Applications in Control Theory and Probability Theory*, Mathematics and its Applications, Kluwer, 2004.

[30] R. E. Chandler, *Hausdorff compactifications*, Lecture Notes in Pure and Applied Mathematics, vol. 23, Marcel Dekker, 1976.

[31] P. Ciarlet, *Mathematical Elasticity*, vol. 1: Three Dimensional Elasticity, North-Holland, 1988.

[32] S. Conti, D. Faraco, and F. Maggi, *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** (2005), no. 2, 287–300.

[33] B. Dacorogna, *Weak Continuity and Weak Lower Semicontinuity for Nonlinear Functionals*, Lecture Notes in Mathematics, vol. 922, Springer, 1982.

[34] ———, *Direct Methods in the Calculus of Variations*, 2nd ed., Applied Mathematical Sciences, vol. 78, Springer, 2008.

[35] G. Dal Maso, A. DeSimone, M. G. Mora, and M. Morini, *Time-dependent systems of generalized Young measures*, Netw. Heterog. Media **2** (2007), 1–36.

[36] ———, *Globally stable quasistatic evolution in plasticity with softening*, Netw. Heterog. Media **3** (2008), 567–614.

[37] ———, *A vanishing viscosity approach to quasistatic evolution in plasticity with softening*, Arch. Ration. Mech. Anal. **189** (2008), 469–544.

[38] C. De Lellis, *A note on Alberti's rank-one theorem*, Transport equations and multi-D hyperbolic conservation laws (Bologna, 2005), Lecture Notes of the Unione Matematica Italiana, vol. 5, Springer, 2008, pp. 61–74.

[39] S. Delladio, *Lower semicontinuity and continuity of functions of measures with respect to the strict convergence*, Proc. Roy. Soc. Edinburgh Sect. A **119** (1991), 265–278.

[40] S. Demoulini, *Young measure solutions for nonlinear evolutionary systems of mixed type*, Ann. Inst. H. Poincaré Anal. Non Linéaire **14** (1997), 143–162.

[41] N. Dinculeanu, *Vector measures*, International Series of Monographs in Pure and Applied Mathematics, vol. 95, Pergamon Press, 1967.

[42] R. J. DiPerna, *Compensated compactness and general systems of conservation laws*, Trans. Amer. Math. Soc. **292** (1985), 383–420.

[43] R. J. DiPerna and A. J. Majda, *Oscillations and concentrations in weak solutions of the incompressible fluid equations*, Comm. Math. Phys. **108** (1987), 667–689.

[44] G. Dolzmann and J. Kristensen, *Higher integrability of minimizing Young measures*, Calc. Var. Partial Differential Equations **22** (2005), 283–301.

[45] G. Dolzmann and S. Müller, *Microstructures with finite surface energy: the two-well problem*, Arch. Ration. Mech. Anal. **132** (1995), 101–141.

[46] F. Ebobisse, *On lower semicontinuity of integral functionals in $LD(\Omega)$*, Ricerche Mat. **49** (2000), no. 1, 65–76.

[47] ———, *A lower semicontinuity result for some integral functionals in the space SBD*, Nonlinear Anal. **62** (2005), 1333–1351.

[48] R. E. Edwards, *Functional analysis. Theory and applications*, Holt, Rinehart and Winston, 1965.

[49] L. C. Evans, *Weak convergence methods for nonlinear partial differential equations*, CBMS Regional Conference Series in Mathematics, vol. 74, Conference Board of the Mathematical Sciences, 1990.

[50] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, Studies in Advanced Mathematics, CRC Press, 1992.

[51] H. Federer, *Geometric measure theory*, Grundlehren der mathematischen Wissenschaften, vol. 153, Springer, 1969.

[52] H. Federer and W. H. Fleming, *Normal and integral currents*, Ann. of Math. **72** (1960), 458–520.

[53] G. B. Folland, *Real analysis*, 2nd ed., Pure and Applied Mathematics, Wiley, 1999.

[54] I. Fonseca and M. Kružík, *Oscillations and concentrations generated by $\mathcal{A}$-free mappings and weak lower semicontinuity of integral functionals*, ESAIM Control Optim. Calc. Var. **16** (2010), 472–502.

[55] I. Fonseca and G. Leoni, *Modern Methods in the Calculus of Variations: L^p Spaces*, Springer, 2007.

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

[57] ———, *Relaxation of quasiconvex functionals in* $\mathrm{BV}(\Omega, \mathbf{R}^p)$ *for integrands* $f(x, u, \nabla u)$, Arch. Ration. Mech. Anal. **123** (1993), 1–49.

[58] ———, *$\mathcal{A}$-quasiconvexity, lower semicontinuity, and Young measures*, SIAM J. Math. Anal. **30** (1999), no. 6, 1355–1390.

[59] I. Fonseca, S. Müller, and P. Pedregal, *Analysis of concentration and oscillation effects generated by gradients*, SIAM J. Math. Anal. **29** (1998), 736–756.

[60] G. A. Francfort, *An introduction to H-measures and their applications*, Variational problems in materials science (Trieste, 2004), Progress in Nonlinear Differential Equations and their Applications, vol. 68, Birkhäuser, 2006, pp. 85–110.

[61] M. Fuchs and G. Seregin, *Variational methods for problems from plasticity theory and for generalized Newtonian fluids*, Lecture Notes in Mathematics, vol. 1749, Springer, 2000.

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

[63] G. Gargiulo and E. Zappale, *A lower semicontinuity result in SBD*, J. Convex Anal. **15** (2008), 191–200.

[64] P. Gérard, *Microlocal defect measures*, Comm. Partial Differential Equations **16** (1991), 1761–1794.

[65] M. Giaquinta, G. Modica, and J. Souček, *Cartesian currents in the calculus of variations. I*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 37, Springer, 1998.

[66] ———, *Cartesian currents in the calculus of variations. II*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 38, Springer, 1998.

[67] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, Birkhäuser, 1983.

[68] L. Grafakos, *Classical Fourier Analysis*, 2nd ed., Graduate Texts in Mathematics, vol. 249, Springer, 2008.

[69] M. Gromov, *Partial differential relations*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 9, Springer, 1986.

[70] M.L. Gromov, *Convex integration of differential relations. I*, Izv. Akad. Nauk SSSR Ser. Mat. **37** (1973), 329–343.

[71] W. Han and B. D. Reddy, *Plasticity – Mathematical Theory and Numerical Analysis*, Interdisciplinary Applied Mathematics, vol. 9, Springer, 1999.

[72] D. Hilbert, *Mathematische Probleme – Vortrag, gehalten auf dem internationalen Mathematiker-Kongreß zu Paris 1900*, Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse (1900), 253–297.

[73] C. O. Horgan, *Korn's inequalities and their applications in continuum mechanics*, SIAM Rev. **37** (1995), 491–511.

[74] A. Kałamajska and M. Kružík, *Oscillations and concentrations in sequences of gradients*, ESAIM Control Optim. Calc. Var. **14** (2008), 71–104.

[75] D. Kinderlehrer and P. Pedregal, *Characterizations of Young measures generated by gradients*, Arch. Ration. Mech. Anal. **115** (1991), 329–365.

[76] ———, *Gradient Young measures generated by sequences in Sobolev spaces*, J. Geom. Anal. **4** (1994), 59–90.

[77] B. Kirchheim, *Rigidity and Geometry of Microstructures*, Lecture notes 16, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig, 2003.

[78] B. Kirchheim, S. Müller, and V. Šverák, *Studying nonlinear pde by geometry in matrix space*, Geometric analysis and nonlinear partial differential equations (S. Hildebrandt and H. Karcher, eds.), Springer, 2003, pp. 347–395.

[79] R. V. Kohn, *New estimates for deformations in terms of their strains*, Ph.D. thesis, Princeton University, 1979.

[80] ———, *New integral estimates for deformations in terms of their nonlinear strains*, Arch. Ration. Mech. Anal. **78** (1982), 131–172.

[81] J. Kristensen, *Lower semicontinuity in spaces of weakly differentiable functions*, Math. Ann. **313** (1999), 653–710.

[82] J. Kristensen and F. Rindler, *Characterization of generalized gradient Young measures generated by sequences in $W^{1,1}$ and BV*, Arch. Ration. Mech. Anal. **197** (2010), 539–598.

[83] ———, *Relaxation of signed integral functionals in BV*, Calc. Var. Partial Differential Equations **37** (2010), 29–62.

[84] M. Kružík and T. Roubíček, *Explicit characterization of L^p-Young measures*, J. Math. Anal. Appl. **198** (1996), 830–843.

[85] ———, *On the measures of DiPerna and Majda*, Math. Bohem. **122** (1997), 383–399.

[86] M. Kružík and J. Zimmer, *On an extension of the space of bounded deformations*, Z. Anal. Anwend., to appear (preprint 2009).

[87] C. J. Larsen, *Quasiconvexification in $W^{1,1}$ and optimal jump microstructure in BV relaxation*, SIAM J. Math. Anal. **29** (1998), 823–848.

[88] H. Matthies, G. Strang, and E. Christiansen, *The saddle point of a differential program*, Energy methods in finite element analysis, Wiley, 1979, pp. 309–318.

[89] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces*, Cambridge Studies in Advanced Mathematics, vol. 44, Cambridge University Press, 1995.

[90] E. J. McShane, *Generalized curves*, Duke Math. J. **6** (1940), 513–536.

[91] ———, *Necessary conditions in generalized-curve problems of the calculus of variations*, Duke Math. J. **7** (1940), 1–27.

[92] ———, *Sufficient conditions for a weak relative minimum in the problem of Bolza*, Trans. Amer. Math. Soc. **52** (1942), 344–379.

[93] C. B. Morrey, Jr., *Quasiconvexity and the semicontinuity of multiple integrals*, Pacific J. Math. **2** (1952), 25–53.

[94] ———, *Multiple Integrals in the Calculus of Variations*, Grundlehren der mathematischen Wissenschaften, vol. 130, Springer, 1966.

[95] S. Müller, *On quasiconvex functions which are homogeneous of degree* 1, Indiana Univ. Math. J. **41** (1992), 295–301.

[96] ———, *Variational models for microstructure and phase transitions*, Calculus of variations and geometric evolution problems (Cetraro, 1996), Lecture Notes in Mathematics, vol. 1713, Springer, 1999, pp. 85–210.

[97] S. Müller and V. Šverák, *Convex integration for Lipschitz mappings and counterexamples to regularity*, Ann. of Math. **157** (2003), 715–742.

[98] F. Murat, *Compacité par compensation*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **5** (1978), 489–507.

[99] ———, *Compacité par compensation. II*, Proceedings of the International Meeting on Recent Methods in Nonlinear Analysis (Rome, 1978), Pitagora, 1979, pp. 245–256.

[100] ———, *Compacité par compensation: condition nécessaire et suffisante de continuité faible sous une hypothèse de rang constant*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **8** (1981), 69–102.

[101] J. Nash, C^1 *isometric imbeddings*, Ann. of Math. **60** (1954), 383–396.

[102] T. O'Neil, *A measure with a large set of tangent measures*, Proc. Amer. Math. Soc. **123** (1995), 2217–2220.

[103] D. Ornstein, *A non-inequality for differential operators in the* L_1 *norm.*, Arch. Rational Mech. Anal. **11** (1962), 40–49.

[104] P. Pedregal, *Parametrized Measures and Variational Principles*, Progress in Nonlinear Differential Equations and their Applications, vol. 30, Birkhäuser, 1997.

[105] D. Preiss, *Geometry of measures in* $\mathbf{R}^n$*: distribution, rectifiability, and densities*, Ann. of Math. **125** (1987), 537–643.

[106] Y. G. Reshetnyak, *Weak convergence of completely additive vector functions on a set*, Siberian Math. J. **9** (1968), 1039–1045.

[107] F. Rindler, *Lower semicontinuity and Young measures in BV without Alberti's Rank-One Theorem*, Adv. Calc. Var., to appear (Ahead of Print: DOI 10.1515/ACV.2011.008; submitted February 2010).

[108] ———, *Lower semicontinuity for integral functionals in the space of functions of bounded deformation via rigidity and Young measures*, Arch. Ration. Mech. Anal. **202** (2011), 63–113.

[109] T. Roubíček, *Relaxation in Optimization Theory and Variational Calculus*, De Gruyter Series in Nonlinear Analysis & Applications, vol. 4, De Gruyter, 1997.

[110] W. Rudin, *Real and Complex Analysis*, 3rd ed., McGraw–Hill, 1986.

[111] P.-M. Suquet, *Existence et régularité des solutions des équations de la plasticité*, C. R. Acad. Sci. Paris Sér. A **286** (1978), 1201–1204.

[112] ———, *Un espace fonctionnel pour les équations de la plasticité*, Ann. Fac. Sci. Toulouse Math. **1** (1979), 77–87.

[113] V. Šverák, *Quasiconvex functions with subquadratic growth*, Proc. Roy. Soc. London Ser. A **433** (1991), 723–725.

[114] M. A. Sychev, *A new approach to Young measure theory, relaxation and convergence in energy*, Ann. Inst. H. Poincaré Anal. Non Linéaire **16** (1999), 773–812.

[115] ———, *Characterization of homogeneous gradient Young measures in case of arbitrary integrands*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **29** (2000), 531–548.

[116] L. Székelyhidi, Jr. and E. Wiedemann, *Young measures generated by ideal incompressible fluid flows*, submitted, arXiv:1101.3499, January 2011.

[117] L. Tartar, *Compensated compactness and applications to partial differential equations*, Nonlinear analysis and mechanics: Heriot-Watt Symposium, Vol. IV, Res. Notes in Math., vol. 39, Pitman, 1979, pp. 136–212.

[118] ———, *The compensated compactness method applied to systems of conservation laws*, Systems of nonlinear partial differential equations (Oxford, 1982), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., vol. 111, Reidel, 1983, pp. 263–285.

[119] ———, *H-measures, a new approach for studying homogenisation, oscillations and concentration effects in partial differential equations*, Proc. Roy. Soc. Edinburgh Sect. A **115** (1990), 193–230.

[120] ———, *On mathematical tools for studying partial differential equations of continuum physics: H-measures and Young measures*, Developments in partial differential equations and applications to mathematical physics (Ferrara, 1991), Plenum, 1992, pp. 201–217.

[121] ———, *Beyond Young measures*, Meccanica **30** (1995), 505–526.

[122] R. Temam, *Mathematical Problems in Plasticity*, Gauthier-Villars, 1985.

[123] R. Temam and G. Strang, *Functions of bounded deformation*, Arch. Ration. Mech. Anal. **75** (1980), 7–21.

[124] L. C. Young, *Generalized curves and the existence of an attained absolute minimum in the calculus of variations*, C. R. Soc. Sci. Lett. Varsovie, Cl. III **30** (1937), 212–234.

[125] ———, *Generalized surfaces in the calculus of variations*, Ann. of Math. **43** (1942), 84–103.

[126] ———, *Generalized surfaces in the calculus of variations. II*, Ann. of Math. **43** (1942), 530–544.

[127] ———, *Lectures on the calculus of variations and optimal control theory*, Saunders, 1969.

[128] ———, *Lectures on the calculus of variations and optimal control theory*, 2nd ed., Chelsea, 1980, (reprinted by AMS Chelsea Publishing 2000).

[129] K. Zhang, *A construction of quasiconvex functions with linear growth at infinity*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **19** (1992), 313–326.

[130] W. P. Ziemer, *Weakly differentiable functions*, Graduate Texts in Mathematics, vol. 120, Springer, 1989.

Index