Calculus of variations in the mixed smoothness setting

Adam Prosiński

St John's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2019

Non, Vicomte, jamais. Il faut vaincre ou périr.

La Marquise de Merteuil au Vicomte de Valmont

— Pierre Choderlos de Laclos, *Les Liaisons dangereuses*

Declaration of Authorship

I hereby declare that, to the best of my knowledge, the contents of this thesis are original and entirely my own work, except where otherwise indicated, cited, or commonly known, nor has any part of this thesis been submitted for a degree at another university.

....

Adam Prosiński

Acknowledgements

First and foremost I would like to express my gratitude towards my advisor, Professor Jan Kristensen. I am thankful for his invaluable help throughout my graduate studies; from the inception of the idea for the project, all the way to the writing of the present thesis, I could always count on his assistance. I am lucky to have worked on a problem that suited my interests from the very start, and which has exposed me to so many new areas, techniques, and questions. Looking back to when I was applying to the programme and writing my research proposal, I appreciate how much my mathematical horizons have broadened, and I have Jan to thank for that. Finally, outside of the purely mathematical side of the DPhil, he has taught me a lot about all the different aspects of academia and helped me navigate this new world. For all that, and more, I am forever grateful.

I would like to thank the entire Oxford PDE group, in particular Professor Gregory Seregin, Professor Sir John Ball FRS, and Professor Luc Nguyen for their helpful comments and for reviewing my work as part of the transfer and confirmation of status procedures. I also wish to express my gratitude towards all the other professors, mentors, and teachers who have helped me along the way. I thank all the faculty of École Polytechnique and Université Paris-Sud, particularly Professor Filippo Santambrogio (now at Université Claude Bernard Lyon 1), with whom I had the pleasure to work on my master's thesis, and later collaborate on my first academic publication. The three years I spent pursuing my bachelor's degree at the University of Warsaw were wonderful, and the scholarly atmosphere created by the faculty and students of the Faculty of Mathematics, Informatics and Mechanics has had a huge impact on my decision to pursue the academic path. Special thanks go to Dr Jerzy Konarski whom I first met as a mathematical analysis professor at Stanisław Staszic High School in Warsaw, and who later taught me at the University of Warsaw.

He has had a profound influence on how I feel and think about mathematics, and I vividly remember the impression his classes made on me. His tremendous dedication to teaching remains, in my mind, an unparalleled example, and I am thankful for all the help he has given me. Lastly, I want to thank all the friends and colleagues I have made amongst my fellow students, in particular my office mates: Simon, Shaun, and Alexei, with whom I have shared the ups and downs of the DPhil experience.

No words that I know of can express the gratitude I feel towards my parents, Teresa and Zygmunt. I wish to thank them for their love beyond measure, their unwavering support, and for providing me with every opportunity that I could ask for. I thank my brothers, Kuba and Michał, whom I have always looked up to, and who, now with their own families, continue to give me reasons to always look forward to visiting Poland. I am thankful to all the members of my close family, who contribute to the wonderful atmosphere I get to enjoy whenever I come home. I consider myself immensely lucky to have made so many friendships that have stood the test of time, and I am grateful to my friends for all the good times that we have shared, and for the memories yet to be made.

Last but not least, I wish to thank Dominika for her invaluable support and patience. You have stood by my side through thick and thin, and you have made even the most stressful periods so much more bearable. I will always cherish the time I have spent in Oxford, and it is all because of you.

Abstract

The present work constitutes a first step towards establishing a systematic framework for treating variational problems posed in Sobolev spaces of mixed smoothness. The crucial difference that separates the following from the existing body of work is that the functionals we consider here depend on the argument function through a mixture of derivatives of different orders in different directions. For a fixed vector $\mathbf{a} := (a_1, \ldots, a_N)$ with positive integer coordinates and a given function $u \colon \mathbb{R}^N \supset \Omega \to \mathbb{R}^n$ we denote by

$$\nabla_{\mathbf{a}} u := (\partial^\alpha u)_{\langle \alpha, \mathbf{a}^{-1} \rangle = 1}$$

the matrix whose i-th row is composed of derivatives $\partial^\alpha u^i$ of the i-th component of the map u, and where the multiindices α satisfy $\langle \alpha, \mathbf{a}^{-1} \rangle = \sum_{j=1}^N \frac{\alpha_j}{a_j} = 1$. We study functionals of the form

$$\mathrm{W}^{\mathbf{a},p}(\Omega; \mathbb{R}^n) \ni u \mapsto \int_\Omega F(\nabla_{\mathbf{a}} u(x)) \, \mathrm{d}x,$$

where $\mathrm{W}^{\mathbf{a},p}(\Omega; \mathbb{R}^n)$ is an appropriate Sobolev space of mixed smoothness and F is the integrand. We are interested in existence and regularity of minimisers of such functionals under prescribed Dirichlet boundary conditions. Throughout the thesis we focus on the difficulties that stem directly from the anisotropic nature of the problem and work with the model case of continuous and autonomous integrands.

The first contribution of this thesis is proving a generalisation of a result due to Chen and Kristensen [37] characterising coercivity of integral functionals posed on $\mathrm{W}^{\mathbf{a},p}(\Omega; \mathbb{R}^n)$ Dirichlet classes in terms of $\mathbf{a}$-quasiconvexity of the integrand, which is the natural adaptation, to the mixed smoothness setting, of Morrey's classical notion of quasiconvexity. We then go on to identify $\mathbf{a}$-quasiconvexity as equivalent to sequential weak lower semicontinuity of our functionals. Finally, we provide relaxation formulas in two scenarios — for integrands that satisfy a strong pointwise coercivity condition such as $F(\xi) \geqslant c|\xi|^p$ and are allowed to be valued in $[0, +\infty]$, and, secondly, for integrands that only satisfy a mean L^p coercivity condition, but that are assumed to be locally Lipschitz and of p-growth from above, i.e., with $0 \leqslant F(\xi) \leqslant C(|\xi|^p + 1)$. The proofs of these results rely heavily on the theory of Young measures generated

by weakly convergent sequences $\nabla_{\mathbf{a}} u_j$ of $\mathbf{a}$-gradients of functions in $\mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n)$, which is another technical contribution of this work. We develop the theory in parallel to what has been proven for $\mathrm{W}^{1,p}$-gradient Young measures and establish results on generation by equiintegrable sequences, a localisation principle, and an analogue of the Kinderlehrer-Pedregal characterisation.

Finally, in the last chapter we study partial regularity of minimisers under a strict $\mathbf{a}$-quasiconvexity assumption on the integrand. In the spirit of Evans' seminal paper [61] we prove that strictly $\mathbf{a}$-quasiconvex integrands of quadratic growth give rise to minimisers whose maximal derivatives are locally Hölder continuous, up to a set of zero Lebesgue measure. Furthermore, we identify the set of regular points through an appropriate smallness condition on the excess of the minimiser around such a point. Our approach is based on a careful local comparison with the solution of the linearised Euler-Lagrange system in a neighbourhood of each point of the domain, hence we also present a number of facts regarding linear quasielliptic systems of partial differential equations.

Contents

Chapter 1

Introduction

Calculus of variations deals, broadly stated, with describing optimal (in a given sense) configurations in problems that often have a physical motivation. As such it may be argued that the roots of calculus of variations may be traced back to ancient Greece. However, the story of what we now call calculus of variations probably begins with the works of Bernoulli, Euler, Lagrange, and Newton, amongst others. Indeed, the name itself stems from Euler's famous treatise 'Elementa calculi variationum', see [60]. The principal idea is that physically relevant scenarios often may be characterised as extrema of appropriate energies. Study of minimal surfaces lets us predict the shape a soap film takes, in general relativity null geodesics are extremals of the distance functional and they are the paths the light travels along, and minimising elastic energy determines (at least as a first approximation) the shape of an elastic material when acted upon by an external force, to name a few examples. Other than physics, economics has recently proven to be a rich source of variational problems. The natural strive for allocating limited resources in the best feasible way has motivated a number of scientists to work on optimal control and optimal transport problems. Furthermore, calculus of variations has proven to be a particularly convenient approach to certain classes of, notoriously difficult, nonlinear partial differential equations. As it turns out, such equations may arise as the Euler-Lagrange system of a given minimisation problem. We will return to this later and, for now, content ourselves with simply signalling the topic.

Calculus of variations has seen rapid development throughout the 20th century. Starting with the announcement of Hilbert's 23 problems (see [85]), the last of which was 'further development of the calculus of variations', followed by introduction of new convexity notions, establishing a plethora of sufficient conditions for existence of minimisers (the 20th problem), proving a number of positive and negative results on regularity of solutions (the 19th problem) and so forth. Whilst, as a community, we now know so much more than we did in 1900, there is still plenty that we do not yet understand, and so the field remains very active. New results are being constantly proven, while the old ones have been incorporated into a standard mathematical curriculum taught at universities around the world.

The present work is intended as a foundation for studying variational problems in which the energy depends on the input function through a selection of its derivatives of different orders in different directions, hence we relax the usual assumptions on the differential operators involved and extend the theory available for homogeneous operators into the mixed smoothness setting. To the best of the author's knowledge, the following is the first systematic approach to variational problems of this type. We tackle a wide range of questions — coercivity, lower semicontinuity, and regularity; along the way we build the necessary machinery, in particular the theory of Young measures generated by sequences in appropriate Sobolev spaces of mixed smoothness. In light of the foundational nature of the work, we focus our attention on the issues that stem directly from the mixed smoothness setting, rather than other technical difficulties. In particular, throughout the thesis we work with continuous and autonomous integrands acting on regular (in the appropriate sense) domains. To facilitate the presentation, we begin, in the introduction, with a selection of classical results of the vectorial calculus of variations. These serve as indicators for desirable results, as well as inspirations for their proofs. For that reason we include, in the introduction, a brief selection of some important contributions to the classical calculus of variations. The list is certainly far from being complete, instead we refer to books by Dacorogna, Giusti, Müller, and Rindler (see [40], [79], [133], and [148] respectively), the original papers cited in what follows, and finally further references in these.

Apart from purely academic interest, the principal motivations for our work are its possible applications to the theory of quasielliptic systems and the study of certain, recently introduced, metamaterials. Quasielliptic differential operators generalise the well-known class of elliptic operators, and are a subclass of hypoelliptic operators in the sense of Hörmander (see [88], [89]). A particular instance of a quasielliptic operator that is not elliptic in the usual sense is furnished by the operator $-\partial_{x_1}^2 + \partial_{x_2}^4$ acting on functions of two variables, with its associated energy

$$\int |\partial_{x_1} u|^2 + |\partial_{x_2}^2 u|^2 \,\mathrm{d}x. \tag{1.1}$$

Since the introduction of the quasielliptic class by Volevich in [174] and [175] this operator (and its modifications) has been studied, as a model example, by a number of authors, see for instance [171] by Triebel and [153] by Shevchik; see also [115] by Lions and Magenes. Contributions to the general theory of quasielliptic operators and related issues have been made by, for example, Barozzi ([15], [16]), Cavallucci ([36]), Demidenko ([50], [51]), Demidenko & Upsenskii ([52]), Friberg ([69]), Giusti ([77], [78]), Hile ([86]), Hile, Mawata, and Zhou ([87]), and Troisi ([172]), to name a few. We defer a more detailed discussion to the last chapter of this thesis. Instead, let us simply note that it is our hope that, similarly to the classical calculus of variations and elliptic systems, our study of variational problems of mixed smoothness will lead to further development of the theory of quasielliptic systems.

As for applications to material science, let us remark that the introduction of 3D printing has made experimenting with man-made materials much easier, and thus led to a rapid increase of interest in metamaterials. While the idea of using mathematical and physical theories to design composite materials that possess specific desired properties, not found in naturally occurring materials, is certainly not new, the amount of research being done in this direction has greatly risen since late XXth century. One instance of a recently introduced metamaterial, that we are particularly interested in, are the pantographic sheets. Pantographic sheets are designed to be a lightweight material that is highly resilient under directional extension and that can store a considerable amount of elastic energy. Proposed applications range from biomechanical prosthetics to impact shields (see [173] by Turco, Giorgio, Misra, and dell'Isola). The material is typically modelled as a set of long, parallel, and equispaced beams in one plane, with an identical set of beams in a parallel plane slightly below the initial one, and rotated by 90 degrees. Thus placed beams form a grid of sorts, and at each vertex of the grid (when viewed directly from above) one connects the corresponding beams with a small pivot that prevents relative translations but allows the beams to rotate freely with respect to one another. We refer the reader to [49] by dell'Isola, Lekszycki, Pawlikowski, Grygoruk, and Greco, for a more detailed description of the arrangement as well as instructive pictures. The theoretical pantographic sheet is then a homogenisation limit of such a configuration. The connection to our work is due to the fact that, in the limit, the deformation energy for the linearised model is effectively of the form (1.1). Rigorous study of this energy has recently been carried out by Eremeyev, dell'Isola, Boutin, and Steigmann in [59], but we believe that a more systematic framework, foundations of which we propose to develop in this thesis, will allow for a substantial simplification of the study. As a consequence, we hope that more detailed research on complex behaviour of pantographic sheets will follow, and ultimately lead to a quicker and more efficient introduction of this metamaterial as a solution to real world problems.

1.1 The Direct Method

The principal strategy for establishing existence of solutions to variational problems is the, so-called, direct method, the abstract formulation of which is often attributed to Tonelli, see [169] and [170]. The version we state here may be found, for example, in Proposition 1.2.2 in Buttazzo's book [30].

Theorem 1.1.1 (see Proposition 1.2.2 in [30])**.** *Let* (X, τ) *be a topological Hausdorff space. Let* $\mathrm{I}\colon X \to (-\infty; \infty]$ *be a functional that satisfies the two following assumptions:*

i) I *is* τ*-lower semicontinuous,*

ii) I *is* τ*-coercive, that is for every* $t \in \mathbb{R}$ *there exists a* τ*-compact set* $K_t \subset X$ *such that*

$$\{x \in X : \mathrm{I}(x) \leqslant t\} \subset K_t.$$

Then the minimisation problem $\inf_{x \in X} \mathrm{I}(x)$ *has at least one solution.*

At this level of abstraction the result, and its proof, are remarkably simple. In fact, the above could be seen as a variant of Weierstrass' theorem that a lower semicontinuous function defined on a compact set attains its infimum, since the coercivity assumption effectively allows us to restrict our attention to a compact set.

Proof. The case where I is identically equal to $+\infty$ is trivial, thus we may assume $\inf_{x \in X} \mathrm{I}(x) =: M < \infty$. Then the hypotheses imply that for every $M' > M$ the set $\{x \in X : \mathrm{I}(x) \leqslant M'\}$ is nonempty and τ-compact. We deduce that the set

$$\{x \in X : \mathrm{I}(x) \leqslant M\} = \bigcap_{M' > M} \{x \in X : \mathrm{I}(x) \leqslant M'\}$$

is nonempty, hence I admits a minimiser. □

Note that, in the abstract formulation of the problem 'minimise $\mathrm{I}(x)$ over $x \in X$', the set X and the functional I are assumed to be given. However, in practice one often finds that the initial space X is, in a sense, too small for the minimum to be achieved. For example, if one poses a variational problem on the space $C^1(\overline{\Omega}, \mathbb{R}^n)$ it may be very hard to find a sensible topology with enough compact sets for the coercivity assumption to be satisfied. This may often be remedied by passing instead to the larger space $\mathrm{W}^{1,p}(\Omega,\mathbb{R}^n)$ for some exponent p, the choice of which is again dictated by the functional. Then one needs to make sense of I as a functional acting on $\mathrm{W}^{1,p}(\Omega,\mathbb{R}^n)$ by some kind of extension. Finally, a minimiser over $\mathrm{W}^{1,p}(\Omega,\mathbb{R}^n)$ need not be of class $C^1(\overline{\Omega},\mathbb{R}^n)$. However, in some instances (although certainly not always), it may be possible to use the functional to deduce additional regularity of the minimiser once it is known to exist — we will return to this point later on.

Assume that the universe of feasible configurations X and the quantity one aims to minimise I are already decided upon. Observe that the topology τ on X, while irrelevant to the minimisation problem $\inf_{x\in X} \mathrm{I}(x)$, is a deciding factor in the question whether or not the direct method is applicable. It is up to the mathematician to decide on the appropriate topological structure to be introduced on X, and the choice needs to be balanced so that the two conditions are satisfied. On the one hand, the topology needs to be strong enough so that I is lower semicontinuous, but on the other hand, weaker topologies have more compact sets, which makes it easier for the second condition to be satisfied. In practice the choice of topology is usually quite clearly dictated by the coercivity properties of the functional in question, and the difficulties then concentrate around establishing lower semicontinuity.

Finally, we note that, as explained in Remark 1.2.3 in [30], the conditions in Theorem 1.1.1 may be weakened to sequential τ-lower semicontinuity and sequential τ-compactness of K_t. This version of the statement is the one that is used most often, and the one we shall have in mind in the rest of this work. Note that sequential τ-lower semicontinuity is again a topological concept — it is equivalent to lower semicontinuity with respect to τ_{seq}, where τ_{seq} is the strongest topology on X for which the convergent sequences are the τ-convergent sequences. Nevertheless, this viewpoint is not employed very often, since usually it is easier to work with the, more familiar, topology τ and just keep in mind that lower semicontinuity need only be tested along sequences. To put our discussion in a more concrete setting, let us briefly consider some toy examples.

In a typical situation X is some kind of a function space and I is an integral functional. A convenient (and very relevant!) first example would be $X := \mathrm{L}^p(\Omega;\mathbb{R}^n)$ for some $p \in (1,\infty)$ and a bounded domain Ω, and $\mathrm{I}(V) := \int_\Omega F(V(x))\,\mathrm{d}x$. To simplify things further let us assume that the integrand F is a continuous functions satisfying the p-growth bound, i.e. $|F(W)| \leqslant C|W|^p$ for some $C > 0$ and all W. To use the direct method we need to decide on what topology to use on $\mathrm{L}^p(\Omega;\mathbb{R}^n)$. The natural candidate would be the standard norm topology as, due to the growth bound, the functional is continuous with respect to the $\mathrm{L}^p(\Omega;\mathbb{R}^n)$ norm. However, there are very few (pre)compact sets in the strong topology on $\mathrm{L}^p(\Omega;\mathbb{R}^n)$, and so it is quite unlikely that the second condition of our theorem will be satisfied. A relatively common assumption that yields some compactness is that of pointwise coercivity, i.e. $F(W) \geqslant c|W|^p$ for some $c > 0$ and all W. Under this assumption all sublevel sets of I are norm-bounded, and norm-bounded sets are sequentially precompact in the weak topology of $\mathrm{L}^p(\Omega;\mathbb{R}^n)$, thanks to the Banach-Alaoglu Theorem, thus weak topology seems to be a good choice. What remains to be shown is that I is lower semicontinuous along weakly convergent sequences. This need not hold in general, thus we need an additional assumption on F, and the right one turns out to be convexity. If F is convex then so is I and thus a, relatively simple, argument based on Mazur's lemma yields (sequential) lower semicontinuity of I. In fact, in this situation, one may show that convexity is not only a sufficient, but also a necessary condition for lower semicontinuity. To do so one considers a sequence of highly oscillatory vector fields V_j that only take two values and weakly converge to a weighted average of these two values — we do not go into detail here, as the argument is well-known.

The above example is a good demonstration of how the direct method works — usually the space X with its topology is quite clearly determined by the coercivity properties of the functional, and the bulk of work is establishing lower semicontinuity. This will also be reflected in this thesis — we discuss coercivity and prove a theorem characterising it, however most of the part on existence will deal with lower semicontinuity under various assumptions. On the other hand, the downside to this example is that the space $X = \mathrm{L}^p(\Omega;\mathbb{R}^n)$ is, in a sense, too simple to be interesting. Indeed, the classical setting for a variational problem, highly relevant in applications, is that of Sobolev spaces $X := \mathrm{W}^{1,p}_{u_0}(\Omega;\mathbb{R}^n)$ with some fixed Dirichlet boundary condition u_0. The simplest integral functionals on this space take the form $\mathrm{I}(u) := \int_\Omega F(\nabla u)\,\mathrm{d}x$, which is what we would call the autonomous case. Clearly, if F is pointwise coercive and convex then the direct method is, again, applicable, but it is not obvious whether convexity is still a necessary condition. Moreover, the integrand could also depend not only on the spatial variable x, i.e., $F(x, \nabla u)$, but also on the lower order term u, i.e., it could have the form $F(x, u, \nabla u)$ in general. Furthermore, it is often necessary to relax the (p, p) growth assumptions of our toy model, as there are physically relevant energies that do not fall into such category (for an example see [9] by Ball). In what follows we will briefly discuss the available results (and their history), which will form the basis for our considerations of the mixed smoothness case. In the mixed smoothness case instead of the gradient $\nabla u = (\partial_{x_1} u, \dots, \partial_{x_N} u)$ we allow the integrand to depend on u through a mixture of derivatives of different orders in different directions, i.e., we aim to work (among others) with functionals of the form

$$\mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n) \ni u \mapsto \int_\Omega F(\partial^{a_1}_{x_1}u,\ldots,\partial^{a_N}_{x_N}u)\,\mathrm{d}x,$$

where $\mathbf{a} = (a_1,\ldots,a_N)$ is an arbitrary vector with positive integer coordinates, and $\mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n)$ is the corresponding Sobolev space. In fact, as we will see in the first chapter, control on all derivatives of the form $\partial^{a_i}_{x_i}u$ implies control on all derivatives $\partial^\alpha u$ for any multiindex α with $\langle \alpha, \mathbf{a}^{-1}\rangle = 1$, where the coordinates of $\mathbf{a}^{-1} = (a_1^{-1},\ldots,a_N^{-1})$ are just multiplicative inverses of the respective coordinates of $\mathbf{a}$. Thus, we define the $\mathbf{a}$-gradient of u through $\nabla_{\mathbf{a}}u := (\partial^\alpha u)_{\langle\alpha,\mathbf{a}^{-1}\rangle=1}$ and in all that follows we consider integral functionals of the form

$$\mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n) \ni u \mapsto \int_\Omega F(\nabla_{\mathbf{a}}u)\,\mathrm{d}x.$$

1.2 Quasiconvexity

The proof of the necessity of convexity for sequential weak lower semicontinuity of the functional I on $\mathrm{L}^p(\Omega;\mathbb{R}^n)$ was based on the fact that for any two matrices W_1, W_2 and any $\lambda \in (0,1)$ one may construct a vector field V that takes the value W_1 on a subset $\Omega_1 \subset \Omega$ of measure $\lambda|\Omega|$ and the value W_2 on the rest of the domain. However, if we require V to be the gradient of some Sobolev function $u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^n)$ this might turn out to be impossible. This suggests that there may exist non-convex integrands F that induce sequentially weakly lower semicontinuous functionals on $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^n)$. This is indeed the case and the classical example is given by:

$$\mathrm{I}_{\det}\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \ni u \mapsto \int_\Omega \det(\nabla u)\,\mathrm{d}x,$$

with $p > N$. First of all, let us observe that the functional $\mathrm{I}_{\det}$ is constant on $\mathrm{W}^{1,p}_{u_0}(\Omega;\mathbb{R}^N)$ for any fixed boundary condition u_0, which results immediately from the fact that the determinant may be written in divergence form. Furthermore, it is not difficult to show that, even without fixing the boundary condition, the functional is sequentially weakly lower semicontinuous on the entire space $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$, see for instance Dacorogna's book [40]. However, as soon as $N > 1$ determinant is not a convex function of ∇u, thus proving that convexity is not a necessary condition for sequential weak lower semicontinuity of integral functionals on Sobolev classes.

The correct convexity notion, quasiconvexity, has been introduced by Morrey in his seminal paper [129]. We say that a function $F\colon \mathbb{R}^{n\times N} \to \mathbb{R}$ is quasiconvex at $W \in \mathbb{R}^{n\times N}$ if

$$F(W) \leqslant \inf_{\varphi\in C_c^\infty(Q;\mathbb{R}^n)} \fint_Q F(W+\nabla\varphi(x))\,\mathrm{d}x, \tag{1.2}$$

where $Q \subset \mathbb{R}^N$ is the unit cube. Naturally, we say that F is quasiconvex if it is quasiconvex at every point. One way to look at quasiconvexity is that it prescribes that, under an affine boundary condition, the affine map is energetically optimal. In the aforementioned paper Morrey has shown that, for continuous and autonomous integrands, quasiconvexity is a sufficient condition for sequential weak* lower semicontinuity on $\mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^n)$ of the induced functional. It is not difficult to show that it is also necessary (also proven by Morrey), which is usually done using the Riemann-Lebesgue lemma (see, for example, Lemma 4.15 in [148]). It is also easy to see that, in general, quasiconvexity is weaker than convexity — again, det is an example. In fact, this example shows that quasiconvexity is strictly weaker than convexity as soon as $N, n \geqslant 2$, since it is enough to consider the determinant of the leading principal 2×2 minor of ∇u. In the scalar case $N = 1$ or $n = 1$ the two notions are equivalent, and we will explain this shortly.

From the above we know that, at least in certain cases, quasiconvexity is equivalent to lower semicontinuity of the induced functional, thus it is the 'correct' notion for studying the applicability of the direct method. One of the major drawbacks of quasiconvexity is that it is notoriously difficult to verify. The reason behind it is that the notion is non-local, as shown by Kristensen in [101]. Thus, two other, easier to use in practice, notions have been subsequently studied: polyconvexity and rank-one convexity.

Following Ball (see [9]) we say that a function $F\colon \mathbb{R}^{n\times N} \to [-\infty,\infty]$ is polyconvex if it may be written as a convex function of all the minors of the argument matrix. A particular instance of a polyconvex function would be the determinant mentioned earlier or, indeed, any convex function of the determinant. Ball introduced polyconvexity to deal with certain problems of nonlinear elasticity, in particular to avoid interpenetration of matter and incorporate the physical restriction $\det(\nabla u) > 0$ into the considerations.

The notion of rank-one convexity was already discussed by Morrey who noted that if F is quasiconvex then for any $V \in \mathbb{R}^{n\times N}$ and any $W \in \mathbb{R}^{n\times N}$ with $\operatorname{rank}(W) = 1$ the function $t \mapsto F(V + tW)$ is convex. We call any function that is convex along rank-one directions a rank-one convex function. When F is of class C^2 then it is rank-one convex if and only if

$$\sum_{i,j}\sum_{\alpha,\beta} F_{W^i_\alpha W^j_\beta}(W)a_\alpha a_\beta b^i b^j \geqslant 0, \tag{1.3}$$

for any $W \in \mathbb{R}^{n\times N}$, $a \in \mathbb{R}^n$, $b \in \mathbb{R}^N$. This inequality is known as the Legendre-Hadamard condition, and we shall return to it later in the context of regularity theory.

It is known (see, for instance, [40]) that the following chain of implications holds:

$$\text{convexity } \Rightarrow \text{polyconvexity } \Rightarrow \text{quasiconvexity } \Rightarrow \text{rank-one convexity.}$$

In the scalar case $N = 1$ or $n = 1$ clearly all directions are rank-one directions and rank-one convexity is the usual convexity, thus all the notions are equivalent. As for the reverse implications in the vectorial case $N, n \geqslant 2$ one of the most famous counterexamples is due to Alibert, Dacorogna, and Marcellini (see [5] and [42]). Let $F_\gamma\colon \mathbb{R}^{2\times 2} \to \mathbb{R}$ be defined by

$$F_\gamma(W) := |W|^2(|W|^2 - 2\gamma\det(W)).$$

Then F_γ is convex if and only if $|\gamma| \leqslant \frac{2\sqrt{2}}{3}$, polyconvex if and only if $|\gamma| \leqslant 1$, quasiconvex if and only if $|\gamma| \leqslant \gamma_{\mathrm{qc}}$ for some (unknown) $\gamma_{\mathrm{qc}} \in (1, \frac{2}{\sqrt{3}}]$, and rank-one convex if and only if $|\gamma| \leqslant \frac{2}{\sqrt{3}}$. Since we do not know if γ_{qc} is strictly smaller than $\frac{2}{\sqrt{3}}$, the example is inconclusive with respect to the last implication. It has been conjectured by Morrey (see [129]) that rank-one convexity does not imply quasiconvexity but we still do not know whether this is true in the case $N = n = 2$. On the other hand, in the case $N \geqslant 2$ and $n \geqslant 3$ the famous counterexample due to Šverák (see [157]) proved Morrey's conjecture. Very recently Grabovsky (see [83]) has found a new example of a family of functions in higher dimensions (8×2), that are rank-one convex, but not quasiconvex.

In this thesis we do not touch upon the question of an analogue of polyconvexity in the mixed smoothness setting, but we do discuss the possibility of extending the notion of rank-one convexity. The reason behind this is that rank-one convexity may be used to prove very useful regularity results regarding quasiconvex envelopes of functions such as inheriting growth bounds from the original function, local Lipschitz continuity, or even differentiability (see [12]). So far a satisfactory analogue is only available in certain cases (see [93]), and so we are led to introduce new proof strategies that allow us to circumvent this problem.

Let us note here that a whole family of various notions of quasiconvexity may be generated by changing the admissible test space in the defining inequality (1.2). Following Ball and Murat (see [13]) we say that F is $\mathrm{W}^{1,p}$-quasiconvex at W if

$$F(W) \leqslant \inf_{\varphi \in \mathrm{W}_0^{1,p}(Q;\mathbb{R}^n)} \fint_Q F(W + \nabla\varphi(x))\,\mathrm{d}x. \tag{1.4}$$

It is not difficult to see that for continuous integrands of p growth, i.e., satisfying $|F(W)| \leqslant C(1+|W|^p)$, the notions of quasiconvexity and $\mathrm{W}^{1,p}$-quasiconvexity coincide. This results from density of $C_c^\infty(Q;\mathbb{R}^n)$ in $\mathrm{W}^{1,p}(Q;\mathbb{R}^n)$ and strong $\mathrm{L}^p(\Omega;\mathbb{R}^n)$ continuity of $V \mapsto \fint_Q F(V(x))\,\mathrm{d}x$. However, in the absence of the p-growth condition, this continuity need not hold. Indeed, it has been shown in [13] that the notions of $\mathrm{W}^{1,p}$- and $\mathrm{W}^{1,q}$-quasiconvexity, in general, differ for $p \neq q$, although clearly $\mathrm{W}^{1,p}$-quasiconvexity implies $\mathrm{W}^{1,q}$-quasiconvexity for any $q > p$. The reason for introducing $\mathrm{W}^{1,p}$-quasiconvexity is that it is naturally better suited for studying lower semicontinuity on $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^n)$.

We remark that, typically, the definition of quasiconvexity excludes extended real-valued integrands, although being able to work with integrands that may take the value $+\infty$ allows one to incorporate hard constraints on the minimisers. For example, in the context of nonlinear elasticity the relevant condition is $\det \nabla u > 0$ (however, note that, as mentioned before, polyconvexity deals with this problem, although it only provides a sufficient, and not necessary, condition). There have been a few attempts at treating this issue, and some may be found in the works of Ball and Murat [13] (where the test space consists of periodic functions), Dacorogna and Fusco [41] (where the usual notion is used, but lower semicontinuity is only obtained along particular sequences), or Wagner [176], [177], and [178] (who considers integrands defined only on a fixed convex subset of the target space and extended by $+\infty$ elsewhere). For us, the most relevant notion is that of closed $\mathrm{W}^{1,p}$-quasiconvexity introduced by Pedregal in [140] and later studied by Kristensen in [102], whose results we generalise to the mixed smoothness framework in Section 4.3.3. At this stage, let us only say that the definition replaces the test space $\mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^n)$ in (1.4) by the space of homogeneous $\mathrm{W}^{1,p}$-gradient Young measures with barycentre 0, and postpone the discussion until we have rigorously introduced the theory of Young measures.

Lastly, let us note that, while the first order gradients case is the classical setting for modern calculus of variations, it is not the only one that has been studied. The first natural generalisation is to consider k-th order gradients, first considered by Meyers in [125], who extended Morrey's lower semicontinuity results from [129]. Ball, Currie, and Olver (see [11]) later elaborated on the problem by providing a characterisation of null Lagrangians (i.e., integrands that yield weakly continuous, as opposed to lower semicontinuous, functionals) of arbitrary order. As it turns out, the notions of k-th order quasiconvexity and ordinary (first order) quasiconvexity are remarkably similar. In this context it has been shown by Dal Maso, Fonseca, Leoni, and Morini (see [44]) that second order quasiconvexity effectively reduces to traditional quasiconvexity, which has been later extended by Cagnetti (see [32]) to k-th order quasiconvexity for an arbitrary k.

Another generalisation that has attracted much attention is $\mathcal{A}$-quasiconvexity. The idea behind it is that a vector field V is the gradient of some function if and only if it satisfies $\mathrm{curl} V = 0$. It is interesting to investigate functionals acting on vector fields satisfying a general differential constraint of the form $\mathcal{A}V = 0$. Such a framework is applicable to, for example, micromagnetics, through a mixture of div-free and curl-free conditions. The operator $\mathcal{A}$ is typically assumed to be a homogeneous k-th order differential operator with constant coefficients, and to satisfy Murat's constant rank condition (see [136]). Foundations for studying variational problems posed on $\mathcal{A}$-free vector fields were developed by a number of authors, such as Murat (see [135]), Tartar (see [164]), Dacorogna (see [38]), and many others. More recently, $\mathcal{A}$-quasiconvexity has been studied by Fonseca and Müller (see [67]), Fonseca, Leoni, and Müller (see [65]), Braides, Fonseca, and Leoni (see [26]), as well as the author (see [145]), to name a few.

1.3 Lower semicontinuity

We have already seen that various notions of convexity are intimately related to questions of sequential lower semicontinuity of integral functionals. Let us mention some examples of the available results under relatively general assumptions and briefly discuss their history. Let us note here that, as a shorthand, we will sometimes talk of 'lower semicontinuity' of functionals meaning 'sequential weak lower semicontinuity', but we shall only do so when there is no risk of confusion.

The story begins with Tonelli (see [170]), who has shown that (in dimension one) non-negative integrands $F(x, z)$ that are convex in z for almost every x induce functionals that are sequentially weakly lower semicontinuous on $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^n)$ for any $p \in (1, \infty)$. We have already seen that, in the scalar case, convexity is the optimal assumption, so let us move on to the vectorial case.

The first results using the quasiconvexity assumption are due to Morrey (see [129] and the book [131]). These were later generalised and improved by a number of authors, notably Acerbi and Fusco in [1] and Marcellini in [116]. The statements we give here are from [40].

In the case $p = \infty$:

Theorem 1.3.1 (see Theorem 8.8 in [40])**.** *Let $\Omega \subset \mathbb{R}^N$ be a bounded open set with a Lipschitz boundary and let $F\colon \Omega \times \mathbb{R}^n \times \mathbb{R}^{n\times N}$ be a Carathéodory function such that $W \mapsto F(x,u,W)$ is quasiconvex for almost every $x \in \Omega$ and every $u \in \mathbb{R}^n$. Assume that F satisfies*

$$|F(x,u,W)| \leqslant \beta(x) + \alpha(|u|,|W|),$$

where $\alpha, \beta \geqslant 0$, $\beta \in \mathrm{L}^1(\Omega;\mathbb{R})$ and α is continuous and increasing in each argument.

Let

$$\mathrm{I}(u,\Omega) := \int_\Omega F(x,u(x),\nabla u(x))\,\mathrm{d}x.$$

Then I is sequentially weakly lower semicontinuous on $\mathrm{W}^{1,\infty}(\Omega;\mathbb{R}^n)$.*

In the case $p < \infty$:

Theorem 1.3.2 (see Theorem 8.11 in [40])**.** *Let $\Omega \subset \mathbb{R}^N$ be a bounded open set with a Lipschitz boundary and let $F\colon \Omega \times \mathbb{R}^n \times \mathbb{R}^{n\times N}$ be a Carathéodory function such that $W \mapsto F(x,u,W)$ is quasiconvex for almost every $x \in \Omega$ and every $u \in \mathbb{R}^n$. Let $1 \leqslant p < \infty$ and assume that F satisfies, for almost every $x \in \Omega$ and every $(u,W) \in \mathbb{R}^n \times \mathbb{R}^{n\times N}$,*

$$-\alpha(|u|^r + |W|^q) - \beta(x) \leqslant F(x,u,W) \leqslant g(x,u)(1+|W|^p),$$

where $\alpha, \beta, g \geqslant 0$, $\beta \in \mathrm{L}^1(\Omega;\mathbb{R})$, $1 \leqslant q < p$, $1 \leqslant r < Np/(N-p)$ if $p < N$ and $1 \leqslant r < \infty$ if $p \geqslant N$ and $g\colon \Omega \times \mathbb{R}^n \to [0,\infty)$ is a Carathéodory function. In the case $p = 1$ we assume that $|F(x,u,W)| \leqslant \alpha(1+|W|)$ for some $\alpha \geqslant 0$.

Let

$$\mathrm{I}(u,\Omega) := \int_\Omega F(x,u(x),\nabla u(x))\,\mathrm{d}x.$$

Then I is sequentially weakly lower semicontinuous on $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^n)$.

The second result is the most relevant to what we shall do later on in the thesis, as we will confine our attention to the case $p \in (1,\infty)$. Note that introducing lower order terms, i.e., x and u, to the integrand, necessarily means the statement becomes more complicated, but the core of the result is still the behaviour of F with respect to its last argument. From above the growth of F is controlled by $|W|^p$ so that quasiconvexity is equivalent to $\mathrm{W}^{1,p}$-quasiconvexity and thus is the appropriate notion for studying lower semicontinuity on $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^n)$. On the other hand, the negative part of F is required to be bounded by $|u|^r + |\nabla u|^q$ with exponents r and q strictly below critical. Thanks to that, the negative part is equiintegrable along sequences bounded in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^n)$ and does not pose major problems.

In this thesis we identify **a**-quasiconvexity as a sufficient condition for lower semicontinuity in the following lemma:

Lemma 1.3.3. *Let Ω be a bounded open domain. Suppose that $F\colon \mathbb{R}^{n\times m} \to [0,\infty]$ is a closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvex integrand. Then the functional*

$$\mathrm{I}_F(u) := \int_\Omega F(\nabla_{\mathbf{a}} u)\,\mathrm{d}x$$

is sequentially weakly lower semicontinuous on $\mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n)$.

We refer to Section 2.2 and Section 4.3 for the precise definitions of the weak **a**-horn condition and of closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity respectively. In the case where F is continuous and of p-growth, closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity is equivalent to **a**-quasiconvexity (see Lemma 4.2.5) and thus the above is a direct generalisation of Theorem 1.3.2 to the mixed smoothness setting, at least for non-negative and autonomous functionals. The statement that we provide has the advantage that it can also deal with extended real-valued integrands at, essentially, no additional cost in the proof. This, in turn, is thanks to the fact that we take advantage of the modern, Young measures based, approach, and the difficulty is instead shifted onto establishing structural theory of these measures.

We defer the question of necessity of quasiconvexity until we move on to discuss relaxation results. However, before we do that, let us take a look at the other component necessary for existence of minimisers — coercivity.

1.4 Coercivity

Let us remark that lower semicontinuity results paired with some sort of coercivity assumption immediately yield existence. For example, under the assumptions of Theorem 1.3.2 if one assumes further that F satisfies a bound of the form

$$F(x,u,W) \geqslant \alpha|W|^p + \beta|u|^q + \gamma(x),$$

for some $\alpha > 0$, $\beta \in \mathbb{R}$, $\gamma \in \mathrm{L}^1(\Omega;\mathbb{R})$, $1 \leqslant q < p$ then for any $u_0 \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^n)$ the minimisation problem $\inf_{u\in \mathrm{W}^{1,p}_{u_0}(\Omega;\mathbb{R}^n)} \mathrm{I}(u)$, admits at least one solution (see Theorem

8.29 in [40]). This result is due to Acerbi and Fusco ([1]) and Marcellini ([116]), who have improved on results by Morrey ([129], see also the book [131]) and Meyers ([125]).

Historically coercivity of functionals in the calculus of variations has received considerably less attention than their lower semicontinuity. Indeed, existence results are often stated under pointwise coercivity assumption such as the one above. This assumption is convenient but not necessary, and indeed, there are a number of physically motivated functionals that are not pointwise coercive. A less restrictive notion has been suggested by Iwaniec and Lutoborski in [91] and is called L^p mean coercivity. In the language that we use here L^p mean coercivity of I on $\mathrm{W}^{1,p}_{u_0}(\Omega;\mathbb{R}^n)$ means that there exist constants C_1, C_2 with $C_1 > 0$ and such that

$$\mathrm{I}(u) \geqslant C_1 \|u\|^p_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^n)} + C_2 \text{ for all } u \in \mathrm{W}^{1,p}_{u_0}(\Omega,\mathbb{R}^n). \tag{1.5}$$

Observe that here we ask for coercivity on a given domain and with fixed boundary condition. In fact, one can easily check that if the constants C_i may be chosen uniformly with respect to the boundary condition, then the integrand has to be pointwise coercive. Finally, note that for applicability of the direct method we do not really need the lower bound to be of the form given in (1.5). It would be enough to know that the value of the functional goes to $+\infty$ along any sequence $u_j \in \mathrm{W}^{1,p}_{u_0}(\Omega;\mathbb{R}^n)$ such that $\|u_j\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^n)} \to \infty$, and this notion is called simply L^p coercivity.

Coercivity of functionals in the calculus of variations has been recently studied by Chen and Kristensen (see [37]). For continuous and autonomous integrands of p-growth they have shown that L^p coercivity and L^p mean coercivity are equivalent, and that coercivity is an intrinsic property of the integrand that may be characterised in terms of strict quasiconvexity of the integrand at a point. In Section 4.1 we show that this may be generalised to the mixed smoothness setting and prove the following:

Theorem 1.4.1. *Let $F\colon \mathbb{R}^{n\times m} \to \mathbb{R}$ be a continuous integrand satisfying the p-growth condition $|F(W)| \leqslant C(1+|W|^p)$. Then, for any $q \in [1,p]$, the following are equivalent:*

a) For any bounded open set $\Omega \subset \mathbb{R}^N$ and any boundary datum $g \in \mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N;\mathbb{R}^n)$ the functional $\mathrm{I}(\cdot,\Omega)$ is L^q coercive on $\mathrm{W}^{\mathbf{a},p}_g(\Omega;\mathbb{R}^n)$.

b) There exist a non-empty bounded open set $\Omega \subset \mathbb{R}^N$ and a boundary datum $g \in \mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N;\mathbb{R}^n)$ such that the functional $\mathrm{I}(\cdot,\Omega)$ is L^q coercive on $\mathrm{W}^{\mathbf{a},p}_g(\Omega;\mathbb{R}^n)$.

c) For any bounded open set $\Omega \subset \mathbb{R}^N$ and any boundary datum $g \in \mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N;\mathbb{R}^n)$ the functional $\mathrm{I}(\cdot,\Omega)$ is L^q mean coercive on $\mathrm{W}^{\mathbf{a},p}_g(\Omega;\mathbb{R}^n)$.

d) There exist a non-empty bounded open set $\Omega \subset \mathbb{R}^N$ *and a boundary datum* g *in* $\mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N;\mathbb{R}^n)$ *such that the functional* $\mathrm{I}(\cdot,\Omega)$ *is* L^q *mean coercive on* $\mathrm{W}^{\mathbf{a},p}_g(\Omega;\mathbb{R}^n)$.

e) For any bounded open set $\Omega \subset \mathbb{R}^N$ *and any boundary datum* $g \in \mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N;\mathbb{R}^n)$ *all* $\mathrm{W}^{\mathbf{a},p}_g(\Omega;\mathbb{R}^n)$ *minimising sequences for the functional* $\mathrm{I}(\cdot,\Omega)$ *are bounded in* $\mathrm{W}^{\mathbf{a},q}_g(\Omega;\mathbb{R}^n)$.

f) There exist a non-empty bounded open set $\Omega \subset \mathbb{R}^N$ *and a boundary datum* g *in* $\mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N;\mathbb{R}^n)$ *such that all* $\mathrm{W}^{\mathbf{a},p}_g(\Omega;\mathbb{R}^n)$ *minimising sequences for the functional* $\mathrm{I}(\cdot,\Omega)$ *are bounded in* $\mathrm{W}^{\mathbf{a},q}_g(\Omega;\mathbb{R}^n)$.

g) There exist a constant $c > 0$ *and a point* $W_0 \in \mathbb{R}^{n\times m}$ *such that the integrand* $W \mapsto F(W) - c|W|^q$ *is* $\mathbf{a}$*-quasiconvex at* W_0.

Let us note that the last point of the theorem is particularly useful in the anisotropic setting. Indeed, it is one of the important ingredients that allow us to prove a relaxation result in absence of an analogue of rank-one directions that has been mentioned previously.

1.5 Relaxation

When a functional lacks coercivity or lower semicontinuity, minimisers need not exist. It is therefore natural to look for ways to remedy the problem by changing the functional slightly and considering a 'relaxed' version of it. There are a number of ways to go about it, one of which is considering the lower semicontinuous envelope of a functional that need not be lower semicontinuous, to begin with.

Suppose that X is a reflexive and separable Banach space, let $\rightharpoonup$ denote weak convergence in X, and let $\mathrm{I} : X \to \mathbb{R}$ be an arbitrary functional. Clearly, I is not sequentially weakly lower semicontinuous at $u \in X$ precisely if there exists a sequence $u_j \rightharpoonup u$ with $\mathrm{I}(u) > \liminf_{j\to\infty} \mathrm{I}(u_j)$. Thus, it is tempting to define the relaxed functional $\bar{\mathrm{I}}\colon X \to [-\infty,\infty)$ by

$$\bar{\mathrm{I}}(u) := \inf_{u_j \rightharpoonup u} \left\{ \liminf_{j\to\infty} \mathrm{I}(u_j) \right\},$$

where the infimum is taken over all sequences u_j converging weakly to u. Alternatively, one could look at functionals that already have the desired lower semicontinuity property, and try to find one that best approximates I from below. This would mean defining

$$\widetilde{\mathrm{I}}(u) := \sup \{ \mathrm{J}(u)\colon \mathrm{J} \leqslant \mathrm{I} \text{ and } \mathrm{J} \text{ is sequentially weakly lower semicontinuous} \}.$$

Assuming that I satisfies the weak coercivity assumption of Theorem 1.1.1 and that X is a reflexive and separable Banach space the two definitions coincide (see Theorem 7.5 and Problem 7.5 in [148]): $\widetilde{\mathrm{I}} = \bar{\mathrm{I}}$ and the functional $\bar{\mathrm{I}}$ is weakly lower semicontinuous, admits at least one minimiser, and we have

$$\inf_{u \in X} \bar{\mathrm{I}}(u) = \inf_{u \in X} \mathrm{I}(u).$$

This justifies calling $\bar{\mathrm{I}}$ the relaxation of I and leads to the question — how to describe $\bar{\mathrm{I}}$ in a more explicit way?

It is reasonable to expect that a relaxation of an integral functional should again be an integral functional with a 'relaxed' integrand. This need not always hold but is a good place to start. Since the relaxed functional is lower semicontinuous, the relaxed integrand should be a quasiconvex function. Hence, the first place to start would be to introduce quasiconvex envelopes of functions, similarly to the well-known convex envelopes. Before we do that, let us note that in the scalar case it is well known that the relaxed problem is given by integration against the convexification of the original integrand. This has been shown in dimension one by Young in [180] (see also the book [183]) and generalised to the scalar case by several other authors — see Berliocchi and Lasry in [19], Ekeland in [57], Ekeland and Temam in their book [58], McShane in [122] and [123], and Marcellini and Sbordone in [119] and [120].

As far as the vectorial case is concerned, Dacorogna [38] was the first one to identify the envelope of relaxation as being given by integration against the quasiconvex envelope of the integrand. The quasiconvex envelope of a function may, similarly to the relaxation of a functional, be defined in two ways. One could set

$$\widetilde{\mathcal{Q}}F(W) := \sup \{G(W) \colon G \leqslant F \text{ and } G \text{ is quasiconvex}\}, \tag{1.6}$$

or

$$\mathcal{Q}F(W) := \inf_{\varphi \in C_c^\infty(Q)} \fint_Q F(W + \nabla\varphi(x)) \, \mathrm{d}x. \tag{1.7}$$

In [38] Dacorogna has shown that, for continuous, non-negative integrands F satisfying a p-growth condition $F(W) \leqslant C(1 + |W|^p)$ the two definitions are equivalent and give rise to a quasiconvex function. In fact, the expression in (1.7) is usually called the Dacorogna formula. In the same paper the author then goes on to show that, under the same hypotheses (although the growth conditions may be relaxed), the lower semicontinuous envelope of relaxation is given by the quasiconvex envelope of the integrand, i.e.,

$$\bar{\mathrm{I}}_F(u) = \mathrm{I}_{\mathcal{Q}F}(u).$$

This result has then been generalised to non-autonomous integrands by Acerbi and Fusco in [1], and we refer the reader to Dacorogna's book [40] for a more thorough summary of relaxation theorems of this kind. In this thesis we generalise Dacorogna's result for autonomous integrands into the mixed smoothness setting, and we consider several further improvements. On one hand we are able to remove the upper growth bound and deal with extended real-valued integrands, on the other we may keep the upper growth bound but relax pointwise coercivity to mean coercivity under further Lipschitz assumption on the integrand, and we prove the following:

Theorem 1.5.1. *Let Ω be a bounded open domain with $|\partial\Omega| = 0$ and satisfying the weak* **a***-horn condition. Suppose that $F\colon \mathbb{R}^{n\times m} \to (-\infty, \infty]$ is a continuous integrand satisfying $F(W) \geqslant C|W|^p - C^{-1}$ for some $C > 0$. Then the sequentially (with respect to approximate* **a***-p gradient convergence) weakly lower semicontinuous envelope of the functional I_F is given by*

$$\bar{\mathrm{I}}_F[V] := \inf_{V_j \to_{\mathbf{a}\text{-}p} V} \left\{ \liminf_{j\to\infty} \mathrm{I}_F[V_j] \right\} = \int_\Omega \overline{F}(V(x))\,\mathrm{d}x,$$

where the infimum is taken over all sequences V_j converging to V in the sense of approximate **a***-p gradient convergence, and $\overline{F}$ denotes the closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvex envelope of F.*

The notion of approximate **a**-p gradient convergence is defined in Section 4.3. It essentially means that $V - V_j$ is an **a**-gradient of some function, up to a small (in $\mathrm{L}^p(\Omega;\mathbb{R}^n)$ norm) perturbation that vanishes as $j \to \infty$. In the absence of upper growth bounds even small perturbations may greatly change the value of the functional, and thus we cannot simply reduce to exact **a**-gradients by projection if we wish to work with extended real-valued integrands. This is, of course, not an issue in the case of functionals of p-growth, as evidenced by the next result:

Theorem 1.5.2. *Let Ω be a bounded open domain with $|\partial\Omega| = 0$ and satisfying the weak* **a***-horn condition. Supppose that $F\colon \mathbb{R}^{n\times m} \to [0,\infty)$ is a continuous and L^p coercive integrand with $F(W) \leqslant C(|W|^p + 1)$. Assume furthermore that F satisfies*

$$F(W) \geqslant D|W|^p - D^{-1}$$

or

$$|F(W) - F(V)| \leqslant D(1 + |W|^{p-1} + |V|^{p-1})|W - V|.$$

Then the sequentially weakly lower semicontinuous envelope of the functional I_F *is given by*

$$\bar{\mathrm{I}}_F(u) := \inf_{u_j \rightharpoonup u} \left\{ \liminf_{j\to\infty} \mathrm{I}_F(u_j) \right\} = \int_\Omega \mathcal{Q}F(\nabla_{\mathbf{a}} u(x))\,\mathrm{d}x = \mathrm{I}_{\mathcal{Q}F}(u),$$

where the infimum is taken over all sequences u_j *converging to* u *weakly in* $\mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n)$.

Note that in the p-growth case we may also relax the lower growth bound and instead only work under the natural coercivity assumption. However, the price we pay for that is an additional assumption about local Lipschitz continuity of the integrand. This allows us to show that the **a**-quasiconvex envelope is again a continuous function (see Lemma 4.3.1), which need not always hold in the mixed smoothness setting. Let us remark here that our proof may easily be adapted to the $\mathcal{A}$-free framework to show that, under analogous coercivity and locally Lipschitz assumptions on the integrand, $\mathcal{A}$-quasiconvex envelopes are continuous, which is known to fail in general if the characteristic cone of the operator does not span the entire space, see [67] by Fonseca and Müller. It should be noted, however, that in the $\mathcal{A}$-free framework it is more natural to deal with such problems by restricting the domain of definition of the integrand in question — this is discussed in a recent preprint by Raiţă and the author (see [146]), and we shall not pursue it here.

Let us observe that the approach to relaxation we have decided to take is not the only possible. Instead of relaxing the integrand one could extend the functional to a larger space than the one it is initially posed on, and then check whether the extended space is big enough for the functional to attain its minimum, but on the other hand not much bigger than the original space, so that the two problems are still 'close' in a sense. One particularly fruitful instance of such a strategy is to extend the functional to the space of appropriate Young measures. This approach is due to Young (see [180] and the book [183]), see also the works of McShane, [122] and [123].

Another extension strategy that must be mentioned is the Lebesgue-Serrin approach. This is particularly relevant in situations where coercivity may only be ensured in a space larger than the natural one on which the functional is defined, for example in the so-called (p,q) growth framework. Let F be an integrand that satisfies the following growth bounds

$$|W|^p \lesssim F(W) \lesssim |W|^q. \tag{1.8}$$

The integral $\int_\Omega F(\nabla u)\,\mathrm{d}x$ is well-defined for $u \in \mathrm{W}^{1,q}(\Omega;\mathbb{R}^n)$, which is the natural space of definition for the functional. However, a priori, the growth condition (1.8) ensures coercivity only in the space $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^n)$. Thus, one needs to appropriately extend the functional to $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^n)$ and look for minimisers there. Observe that the minimiser, if found, is not guaranteed to belong to $\mathrm{W}^{1,q}(\Omega;\mathbb{R}^n)$, and the infima over $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^n)$ and $\mathrm{W}^{1,q}(\Omega;\mathbb{R}^n)$ may differ, which is connected to the so-called Lavrentieff gap phenomenon (see [113]). We do not discuss this approach further as we will not use it in the thesis — we only note that it was first used by Lebesgue in [114], adapted by Serrin in [151] and [152], and later studied by many others, for example Marcellini in [117], or Fonseca and Malý in [66].

1.6 Regularity

The last property of minimisers that we investigate in this thesis is their regularity. Let us start with the classical example of the Dirichlet energy. It is well-known that minimisers of the problem

$$\inf_{u\in\mathrm{W}^{1,2}_{u_0}(\Omega;\mathbb{R}^n)}\int_\Omega |\nabla u|^2\,\mathrm{d}x$$

satisfy

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

and are C^∞ smooth inside Ω. This leads to the question of whether all 'regular' variational problems have regular solutions.

To begin with, let us recall that critical points (minimisers in particular) of integral functionals of the form $\int F(x,u,\nabla u)\,\mathrm{d}x$ solve the Euler-Lagrange system

$$\mathrm{div} F_W(x,u,\nabla u) = F_v(x,u,\nabla u),$$

where v and W denote, respectively, the second and third argument of F and F_v, F_W are the respective derivatives. Under suitable convexity assumptions on the integrand the associated Euler-Lagrange system is elliptic, albeit nonlinear in the general case. Thus, all the positive results regarding regularity of elliptic equations immediately translate into positive results on regularity of minimisers, but not the other way around — it is important to note that being a global minimiser is, in general, more restrictive than solving the associated Euler-Lagrange system. Let us also briefly remark that in the mixed smoothness setting we use the Euler-Lagrange system similarly to the way it has been done in the gradient case, although the system we get is not elliptic, but quasielliptic. Quasielliptic operators are a particular instance of hypoelliptic operators (which yield smooth solutions for smooth data) and have been studied, for instance, by Hörmander in [88] and [89], and by Giusti in [77] and [78];

we defer further discussion of quasielliptic systems to Section 5.1, where the reader can also find more references.

There are plenty of very good sources for the theory of elliptic equations, here let us just mention the books by Ladyzhenskaya and Ural'tseva [112], Bensoussan and Frehse [18], and by Gilbarg and Trudinger [76]. Our story begins with the celebrated papers by De Giorgi (see [45]) and Nash (see [137]) who have, independently and using different approaches, proven Hölder continuity of solutions to elliptic equations in divergence form, which is what the Euler-Lagrange equation is for convex integrands, thus settling the regularity problem in the scalar case. Let us remark here that De Giorgi's method turned out to be very versatile, in particular it could be applied to the minimisers directly (see the works of Frehse [68] and Giaquinta & Giusti [74]), and without going through the Euler-Lagrange equation, thus without loss of information.

In the vectorial case the Euler-Lagrange equation becomes a system and the regularity drastically deteriorates. An example of a functional with discontinuous coefficients that admits discontinuous local minimisers in dimensions $N, n \geqslant 3$ was constructed by De Giorgi in [47]; a similar example was independently found by Maz'ya (see [121]). De Giorgi's example leads to a system that is linear but has discontinuous coefficients. As it turns out, however, the coefficients can be analytic if they are allowed to depend on the solution itself (as in [80] by Giusti and Miranda). Further counterexamples to regularity were found by Nečas in [138] who has constructed an autonomous functional which admits a minimiser that is Lipschitz, but not C^1. Following that, Šverák and Yan have constructed a functional that admits an unbounded minimiser. Let us remark that the examples we have mentioned so far only work in sufficiently high dimension, in particular the question of regularity in dimension $N = 3$, $n = 2$, remained open for a long time, until very recently Mooney and Savin (see [128]) constructed an example of a regular variational integral in dimensions $N \geqslant 3$, $n \geqslant 2$ that admits a non-Lipschitz minimiser.

The examples mentioned above prove that one cannot hope for global smoothness of minimisers of variational problems in the vectorial case. However, Morrey (see [130]) has shown that weak solutions of non-linear elliptic systems are regular outside of a null subset of their domain, thus opening the way for the so-called partial regularity and Evans' seminal paper (see [61]) on partial (up to a subset of Lebesgue measure 0) $C^{1,\alpha}$ regularity of minimisers under the assumption of strict quasiconvexity. Improvements on Evans' result were made notably by Acerbi and Fusco in [2] and [3], by Carozza, Fusco, and Mingione in [35], by Evans and Gariepy in [62], and by Giaquinta and Modica in [75].

The idea that minimising an energy is, in general, stronger than solving the associated Euler-Lagrange system has been reinforced by Müller and Šverák in [134], who provided an example of a strictly quasiconvex integrand that induces an Euler-Lagrange system which admits solutions that are nowhere C^1. This has been subsequently extended to polyconvex integrands by Székelyhidi in [162]. However, partial regularity results may be shown to hold for some, albeit not all, local minimisers — this has been done by Kristensen and Taheri in [108].

In this thesis we extend the partial regularity results to the case of variational problems posed in the mixed smoothness setting, and prove the following:

Theorem 1.6.1. *Let $\Omega \subset \mathbb{R}^N$ be a bounded domain satisfying the weak $\mathbf{a}$-horn condition and suppose that $F\colon \mathbb{R}^{n\times m} \to \mathbb{R}$ is a strictly $\mathrm{W}^{\mathbf{a},2}$-quasiconvex integrand of class C^2. Assume furthermore that $|F(W)| \leqslant C_1(1+|W|^2)$ and $|F''(W)| \leqslant C_2$ for some $C_1, C_2 > 0$ and all $W \in \mathbb{R}^{n\times m}$. Let u be a minimiser of the induced functional $\mathrm{I}_F(u) := \int_\Omega F(\nabla_{\mathbf{a}} u)\,\mathrm{d}x$ over the class $\mathrm{W}^{\mathbf{a},2}_{u_b}(\Omega;\mathbb{R}^n)$ for some fixed boundary datum $u_b \in \mathrm{W}^{\mathbf{a},2}(\mathbb{R}^N;\mathbb{R}^n)$. Then, for any $\alpha < 1$, there exists an open set $\Omega_g \subset \Omega$ with $|\Omega \setminus \Omega_g| = 0$ and such that, on Ω_g, $\nabla_{\mathbf{a}} u$ is locally Hölder continuous with exponent α with respect to the metric $\delta_{\mathbf{a}}$.*

We refer to Section 5.2 for the definition of strict $\mathrm{W}^{\mathbf{a},2}$-quasiconvexity, and to Section 2.6 for the definition of the anisotropic metric $\delta_{\mathbf{a}}$. The proof is based on Campanato-type estimates obtained by locally approximating the minimiser by solutions of the linearised Euler-Lagrange system, which in our case is a quasielliptic system of partial differential equations. Due to the anisotropic nature of the problem we are naturally led to working with Campanato spaces with respect to the metric $\delta_{\mathbf{a}}$, which is adapted to the scaling corresponding to our hyperplane of homogeneity. The result may be rephrased in terms of the standard Euclidean metric, but then the Hölder exponent is different in each direction, i.e., we may show that $\nabla_{\mathbf{a}} u$ is Hölder in direction x_i with exponent $\frac{a_i}{\max_j a_j}\alpha$, where a_i is the order of the maximal pure derivative in direction x_i. This is an interesting feature of the mixed smoothness setting and, in a sense, it is to be expected, since what the result means is that we get less regularity in 'expensive' directions, i.e., ones that allow fewer derivatives.

1.7 Young measures

The main technical tool we use to study existence of solutions to variational problems is the theory of Young measures. These were introduced by Young in [180], [181],

[182], see also the book [183]. The classical Young measures are objects that may be associated with subsequences of weakly convergent $\mathrm{L}^p(\Omega;\mathbb{R}^n)$ vector fields that preserve information on whether, and how, the sequence oscillates around its weak limit. They are especially useful when verifying lower semicontinuity of integral functionals as, essentially, they allow one to break the proof down into two steps, where one first passes to the Young measure limit and then checks whether a Jensen-like inequality is verified. Furthermore, Young measures are great at localising problems, as they measure the oscillations in a point-by-point manner. By now they are a classical tool of calculus of variations, and a very good source on how to use Young measures to approach variational problems is the recent book [148] by Rindler.

As mentioned above, Young measures were introduced by Young ([180], [181], [182]), who called them generalised curves or surfaces. They were since studied and applied by many authors, such as McShane ([122]), Berliocchi and Lasry ([19]), Balder ([8]), Ball ([10]), Kristensen ([100]), and Kinderlehrer & Pedregal ([94], [95]) among others. A classical reference for Young measures that contains a much more complete bibliography is Pedregal's book [141].

One of the most important, from our point of view, results in the theory of Young measures is the Kinderlehrer-Pedregal characterisation of gradient Young measures ([94] and [95], see also [161] by Sychev). The result completely characterises Young measures that are generated by sequences of gradients, rather than arbitrary vector fields. This is immensely important, as it essentially establishes a duality between Young measures generated by gradients and quasiconvex functions, describing one by the other. In this thesis we prove an analogous result for Young measures generated by the derivatives of interest in the mixed smoothness setting as follows:

Theorem 1.7.1. *Fix a bounded open domain Ω satisfying the weak* **a***-horn condition. Let $\{\nu_x\}_{x\in\Omega}$ be a weak* measurable family of probability measures on $\mathbb{R}^{n\times m}$. Then there exists a $\mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n)$ bounded sequence $\{v_n\}\subset \mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n)$ with $\{\nabla_{\mathbf{a}} v_n\}$ generating the oscillation Young measure $\{\nu_x\}_{x\in\Omega}$ if and only if the following conditions hold:*

i) there exists $v\in \mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n)$ such that

$$\nabla_{\mathbf{a}} v(x) = \langle \nu_x, \mathrm{Id}\rangle \text{ for a.e. } x\in\Omega;$$

ii)

$$\int_\Omega\int_{\mathbb{R}^{n\times m}} |W|^p \,\mathrm{d}\nu_x(W)\,\mathrm{d}x < \infty;$$

iii) for a.e. $x \in \Omega$ and all $g \in \mathcal{E}_p$ we have

$$\langle \nu_x, g \rangle \geqslant \mathcal{Q}g(\langle \nu_x, \mathrm{Id} \rangle).$$

We refer to Chapter 3 for explanation of the notation and precise definitions. At this stage let us only say that the essence of the result is contained in the third point above. Roughly translated, it means that we require that all $\mathrm{W}^{\mathbf{a},p}$-quasiconvex functions of p growth satisfy Jensen's inequality with respect to ν_x for almost every x, which in turn essentially means that each ν_x is a (homogeneous) $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure on its own. This is particularly important for the aforementioned localisation with Young measures.

In this thesis we deduce existence results using only the classical (oscillation) Young measures. We can do so, because we restrict to non-negative integrands in the reflexive setting $p \in (1, \infty)$, and thus we do not need to worry about concentration effects. When this is not the case, for example when one wishes to work with $p = 1$, concentration needs to be taken into account, and thus one needs a more accurate tool. Generalised Young measures do just that, they keep track of concentration as well as oscillations. They were introduced by DiPerna and Majda in [54] in connection to fluid mechanics, and later developed by a number of authors, such as Alibert and Bouchitté ([4]), Kristensen and Rindler ([107]), Fonseca and Kružík ([63]), Kałamajska and Kružík ([92]), Kružík and Roubiček ([109] and [110]), Székelyhidi and Wiedemann ([163]), among others. The main idea behind generalised Young measures is that one may choose an appropriate compactification of the target space and consider Young measures on this compactified space, which allows one to preserve some information on what happens at infinity — the choice of the compactification determines what sort of information is preserved. We do not go into details here as we do not use generalised Young measures in this thesis. Nevertheless, to lay the groundwork for future research, we prove a number of facts on them alongside our study of the oscillation Young measures in Chapter 3, to which we refer for details. Finally, let us also note that there are also other extensions of Young measures available in the literature. For example, a variant suitable for treating quadratic expressions has been developed by Tartar (see [167]) and, independently, by Gérard (see [71]), under the names 'H-measures' and 'micro-local defect measures', respectively. These have recently been combined with generalised Young measures in the concept of microlocal compactness forms, introduced by Rindler in [147].

1.8 Spaces of mixed smoothness

The classical Sobolev space $\mathrm{W}^{k,p}(\Omega;\mathbb{R}^n)$ is determined by two scalar parameters $k \in \mathbb{N}$ and $p \in [1,\infty]$ and is defined as the space of (classes of equivalences of) functions such that all their (weak) partial derivatives of order less than or equal to k may be represented by functions in $\mathrm{L}^p(\Omega;\mathbb{R}^n)$. These spaces are by now classical and have found widespread use throughout analysis. However, problems that exhibit different behaviour with respect to different variables call for some generalisations. In this work we are interested in functions that posses different number of derivatives in different directions. Thus, instead of the scalar parameter k we introduce the vector $\mathbf{a} = (a_1, \ldots, a_N) \in \mathbb{N}^N$, and we are interested in the space $\mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n)$ of functions for which all the distributional derivatives $\partial_{x_i}^{a_i} u$ may be represented by functions in $\mathrm{L}^p(\Omega;\mathbb{R}^n)$.

The first question that we need to answer is what other smoothness and integrability properties follow when a function u is assumed to be L^p integrable, together with all its derivatives $\partial_{x_i}^{a_i} u$ mentioned above. The theory of embeddings of spaces of mixed smoothness was largely developed by Nikolskii, who started the study with his 1951 paper [139]. Since then a number of authors have made important contributions to the theory of spaces of mixed smoothness, among which we recall Besov ([20], [21]), Boman ([25]), Burenkov ([28]), Il'in ([90]), Kolyada ([97]), Kolyada & Pérez ([98]), Pełczyński & Senator ([143], [144]), Slobodeckii ([154], [155]), and Solonnikov ([156]).

Let us note here that the list is very far from being complete. Instead we refer the reader to the two-volume book by Besov, Il'in, and Nikolskii (see [23] and [24]). In fact we will follow the books' exposition in most of the technical preliminaries on the structure of spaces of mixed smoothness. Let us note that in their work the authors call the spaces we work with anisotropic Sobolev spaces. However, in this thesis we have opted for the name Sobolev spaces of mixed smoothness to avoid confusion as to the nature of the anisotropy present in the problems we consider. Indeed, variational problems with different growth properties in different directions are often called anisotropic in the existing literature.

We supplement the material from [23] and [24] with a few additional results that are intimately tied to the techniques that we use in this thesis. First of all, quasiconvexity arguments often make use of piecewise affine (in our case piecewise polynomial of appropriate degrees in each variable) approximations. While the two books mentioned above provide some approximation results of this type, they are not very well suited for our purposes, thus we use the result of Dupont and Scott from [56]. Secondly, to be able to perform truncations without spoiling the derivative structure we need a projection theorem along the lines of Helmholtz decomposition for gradients. Such projection has been constructed by Pełczyński in [142]

Finally, let us remark that furthering the study of the structure of Sobolev spaces of mixed smoothness is outside the scope of this thesis. In Chapter 2 we recall the results that we employ throughout, and in all that follows we focus on building the machinery directly related to the variational side of the problem.

Chapter 2

Spaces of mixed smoothness

We begin by introducing the function spaces that we will set our variational problems in. The purpose is to collect the basic facts about Sobolev spaces of mixed smoothness $W^{\mathbf{a},p}$ which, while available in the literature, might not be as classical and well-known as the corresponding theory of the usual Sobolev spaces $W^{k,p}$. Let us note that what we call Sobolev spaces of mixed smoothness is often referred to as 'anisotropic Sobolev spaces' in the literature that we cite. We have opted for the name 'mixed smoothness', as the term 'anisotropic variational problem' is already widespread and used to describe problems where the integrand and the input functions exhibit different growth properties in derivatives in different directions. That is, 'anisotropic variational problems' usually refer to problems posed in the space $W^{k,\mathbf{p}}$ with a vector parameter $\mathbf{p}$, rather than $W^{\mathbf{a},p}$ with vector parameter $\mathbf{a}$, that we are interested in here. Nevertheless, our setting is certainly anisotropic, and we shall use this term on occasion, particularly when talking about scaling in a way adapted to our framework.

We introduce the language, describe the framework, and define the appropriate Sobolev spaces of mixed smoothness in Section 2.1. We then recall embedding results in Section 2.2, that we will use, in particular, to justify cut-off type arguments throughout the thesis. Section 2.3 is devoted to canonical projections that will allow us to truncate the maximal derivatives of our functions whilst preserving the $\mathbf{a}$-gradient structure. The anisotropic scaling that we have briefly mentioned above is introduced in Section 2.4, whilst Section 2.5 contains results on approximating functions in $W^{\mathbf{a},p}$ by piecewise polynomials (of appropriate degrees), that is an analogue of piecewise affine approximations for $W^{1,p}$ functions. Finally, in Section 2.6 we introduce anisotropic version of Campanato and Hölder spaces, which we will use to investigate regularity of minimisers in Chapter 5.

2.1 Preliminaries

First we establish the language for describing the mixed smoothness setting. Any partial differentiation operator $\partial^\alpha := \partial_1^{\alpha_1}\partial_2^{\alpha_2}\dots\partial_N^{\alpha_N}$ acting on functions mapping a subdomain of $\mathbb{R}^N$ to $\mathbb{R}^n$ may be identified with the multiindex $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{Z}_+^N$, where $\mathbb{Z}_+^N$ denotes the set of points in $\mathbb{R}^N$ with non-negative integer coordinates.

Specifying a set of derivatives $A \subset \mathbb{Z}_+^N$ and their desired integrability defines a Sobolev-like space — for example the classical $\mathrm{W}^{k,p}$ Sobolev spaces correspond to $A := \{\alpha \in \mathbb{Z}_+^N \colon |\alpha| \leqslant k\}$. Here we are interested in more general situations, but we must nevertheless impose some structural conditions.

Assumption 1. *If $\alpha, \beta \in \mathbb{Z}_+^N$ are such that $\alpha \in A$ and $\beta \leqslant \alpha$ (coordinate-wise) then also $\beta \in A$.*

This assumption may be phrased by saying that we assume A to be a smoothness (see for example [143]). In particular, whenever A is non-empty, which is naturally a standing assumption, this implies that $0 \in A$, thus we impose integrability condition on the function itself. Therefore, elements of our spaces of mixed smoothness will be classes of equivalence of certain L^p functions rather than classes of equivalence up to a certain polynomial, which is the case, e.g., in homogeneous Sobolev spaces. Furthermore, we assume p-integrability not only on the function and its maximal derivatives (given by maximal elements of A with respect to the coordinate-wise $\leqslant$ relation), but also on all the derivatives in between. In many situations, in particular in the cases considered in this work, this is a consequence of the p-integrability of the maximal derivatives and the function itself and follows from an appropriate variant of the Poincaré inequality. Such inequalities depend on the underlying domain, and although we will always work with domains adapted to the smoothness at hand, we have decided to include the control on the intermediate derivatives in the assumptions to stress its importance.

Assumption 2. *Let A_{max} denote the set of maximal elements of A. We assume that all elements of A_{max} lie on a common hyperplane and that this hyperplane intersects all positive semi-axes.*

This assumption stems from the paper [93] by Kazaniecki, Stolyarov, and Wojciechowski, where the authors call the hyperplane in question a pattern of homogeneity. The motivation behind this assumption is that it provides a certain homogeneity structure that allows one to rescale functions whilst preserving the distribution of the values of all the maximal derivatives. To make this more precise observe that if γ denotes the unit outwards normal to the pattern of homogeneity then there exists some positive real number k such that for all $\alpha \in A_{\max}$ we have $\langle \alpha, \gamma \rangle = k$. Thus, if for a given function u and a number $r > 0$, we let $u_r(x) := r^{-k} u((r^{\gamma_i} x_i)_i)$ then, for all $\alpha \in A_{\max}$, we have

$$(\partial^\alpha u_r)_\# \left(\frac{\mathcal{L}^N \llcorner Q_{r^{-1}}}{\mathcal{L}^N(Q_{r^{-1}})} \right) = (\partial^\alpha u)_\# \left(\frac{\mathcal{L}^N \llcorner Q}{\mathcal{L}^N(Q)} \right),$$

where $\mathcal{L}^N(Q)$ denotes the N-dimensional Lebesgue measure of the cube $Q \subset \mathbb{R}^N$, $\mathcal{L}^N \llcorner Q$ denotes the restriction of the N-dimensional Lebesgue measure to Q, $f_\# \mu$ denotes the pushforward of the measure μ via the function f, and finally Q_r denotes the appropriately scaled cube Q — we will return to this later on when discussing the anisotropic scaling in more detail. Finally, let us remark that the assumption that the pattern of homogeneity intersect each positive semi-axis is very natural, as we want A to be finite and we do not want to impose infinite differentiability in any direction.

Assumption 3. *We assume that the pattern of homogeneity of A_{max} intersects each positive semi-axis at an integer point and that A includes all the points of this plane with non-negative integer coordinates.*

This last assumption is the strictest and it implies the previous one, however we decided to first present the initial two assumptions for clarity of exposition and to explain the purposes they serve. We impose this condition in order for our Sobolev spaces of mixed smoothness to have structure similar to the one of classical Sobolev spaces, in particular we want to ensure that the control on maximal derivatives, paired with control on the function (or with a boundary condition) yields control on all the intermediate derivatives. We will see later that this is indeed true, however the 'intermediate' derivatives are not necessarily all the ones that satisfy $\beta \leqslant \alpha$ for some $\alpha \in A_{\max}$ (which would be the goal), but the ones given by indices in the convex hull of $A_{\max} \cup \{0\}$. If the pattern of homogeneity intersects one of the axes, say x_1 at a non-integer point then no points on this axis belong to the convex hull of $A_{\max} \cup \{0\}$. Thus, we have no information, for instance, on the regularity of derivatives of functions that only depend on the x_1 variable, which is clearly a degenerate situation that we wish to avoid, hence the assumption.

A model example of a pattern of homogeneity and the associated smoothness, mentioned in Equation (1.1) in the Introduction, is furnished by

$$A_{\max} := \{(1,0),(0,2)\} = \{\alpha \in \mathbb{Z}^2_+ \colon \langle \alpha, (1, \frac{1}{2}) \rangle = 1\},$$

with

$$A := \{\alpha \in \mathbb{Z}_+^2 \colon \langle \alpha, (1, \frac{1}{2}) \rangle \leqslant 1\} = \{1,0),(0,2),(0,1),(0,0)\}.$$

Naturally, an even simpler example may be given by simply recasting the case of k-th order derivatives, which would correspond to

$$A_{\max} := \{\alpha \in \mathbb{Z}_+^N \colon |\alpha| = k\} = \{\alpha \in \mathbb{Z}_+^N \colon \langle \alpha, (1,\dots,1)\rangle = k\},$$

with

$$A := \{\alpha \in \mathbb{Z}_+^N \colon |\alpha| \leqslant k\} = \{\alpha \in \mathbb{Z}_+^N \colon \langle \alpha, (1,\dots,1)\rangle \leqslant k\},$$

although this one is isotropic in nature, thus perhaps less instructive for what we wish to study in this thesis. Let us note that, throughout the thesis, we work with vector-valued functions $u\colon \mathbb{R}^N \supset \Omega \to \mathbb{R}^n$. Thus, for any fixed $\alpha \in A$, by the partial derivative $\partial^\alpha u$ we mean the (column) vector of partial derivatives of components of u, i.e.,

$$\partial^\alpha u = (\partial^\alpha u^1, \dots, \partial^\alpha u^n)^T.$$

Observe that this implies that we impose the same differentiability on all components of the functions considered. In general it would be interesting to allow for different smoothnesses for different components, this is, however, outside of the scope of the present thesis.

With these assumptions in place, we may now proceed to define the Sobolev spaces of mixed smoothness that we will work with. Let us fix a pattern of homogeneity satisfying the above assumptions and suppose that it intersects the x_i axis at the point $a_i e_i$, where $\{e_i\}$ is the standard basis or $\mathbb{R}^N$ and $a_i \in \mathbb{N}$. We denote by $\mathbf{a}$ the vector $(a_1, a_2, \dots, a_N)$ and by $\mathbf{a}^{-1}$ the vector $(a_1^{-1}, a_2^{-1}, \dots, a_N^{-1})$. Note that, clearly, any choice of a vector $\mathbf{a} \in \mathbb{N}^N$ uniquely determines an admissible pattern of homogeneity through the equation $\langle \alpha, \mathbf{a}^{-1}\rangle = 1$, with corresponding smoothness given by $\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1$. When talking about derivatives of functions given by a multiindex with only one non-zero entry, say $a_i e_i$, we will often write

$$\partial_i^{a_i} u = \partial_{x_i}^{a_i} u := \partial^{(0,\dots,0,a_i,0,\dots,0)} u = \partial^{a_i e_i} u.$$

Here, and in all that follows, Ω is a bounded open subset of $\mathbb{R}^N$ with $|\partial\Omega| = 0$, where $|A|$ denotes the N-dimensional Lebesgue measure of the set A. We denote by $C_c^\infty(\Omega, \mathbb{R}^n)$ the space of smooth and compactly supported functions $\varphi\colon \Omega \to \mathbb{R}^n$. We will often omit the target space when it is clear from the context, and simply write $C_c^\infty(\Omega)$. We also fix an exponent $p \in (1, \infty)$.

Definition 2.1.1. *For a bounded open set* $\Omega \subset \mathbb{R}^N$ *the Sobolev space* $\mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n)$ *is defined as the completion of* $C^\infty(\Omega;\mathbb{R}^n) \cap \{\varphi\colon \sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \|\partial^\alpha\varphi\|_{\mathrm{L}^p(\Omega;\mathbb{R}^n)} < \infty\}$ *with respect to the norm*

$$\|u\|_{\mathrm{W}^{\mathbf{a},p}} := \sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \|\partial^\alpha u\|_{\mathrm{L}^p(\Omega;\mathbb{R}^n)}.$$

We often omit the target space $\mathbb{R}^n$ *and write simply* $\mathrm{W}^{\mathbf{a},p}(\Omega)$ *or even* $\mathrm{W}^{\mathbf{a},p}$ *if the domain* Ω *is clear from the context. We also denote by* $\mathrm{W}^{\mathbf{a},p}_0(\Omega;\mathbb{R}^n)$ *the completion of* $C^\infty_c(\Omega;\mathbb{R}^n)$ *in the same norm.*

Proposition 2.1.2 (see [23]). *For* $p \in [1,\infty)$ *the space* $\mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n)$ *coincides with the space of functions* $u \in \mathrm{L}^p(\Omega;\mathbb{R}^n)$ *with distributional derivatives* $\partial^\alpha u \in \mathrm{L}^p(\Omega;\mathbb{R}^n)$ *for all* $\langle\alpha,\mathbf{a}^{-1}\rangle \leqslant 1$. *The spaces* $\mathrm{W}^{\mathbf{a},p}(\Omega)$ *and* $\mathrm{W}^{\mathbf{a},p}_0(\Omega)$ *are both separable Banach spaces. For* $p \in (1,\infty)$ *the two spaces are also reflexive.*

Proof. The first two assertions have been established in Section 9 of [23]. The only part we need to prove is reflexivity. This results immediately from the fact that our spaces may, in a canonical way, be seen as subspaces of $\bigoplus_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \mathrm{L}^p(\Omega;\mathbb{R}^n)$. The $\mathrm{W}^{\mathbf{a},p}$ norm then coincides with the norm of the space $\bigoplus_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \mathrm{L}^p(\Omega;\mathbb{R}^n)$, thus $\mathrm{W}^{\mathbf{a},p}(\Omega)$ and $\mathrm{W}^{\mathbf{a},p}_0(\Omega)$ are both closed subspaces. Finally, $\bigoplus_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \mathrm{L}^p(\Omega;\mathbb{R}^n)$ is reflexive, and a closed subspace of a reflexive space is reflexive, thus we are done. □

2.2 Embeddings of Sobolev spaces of mixed smoothness

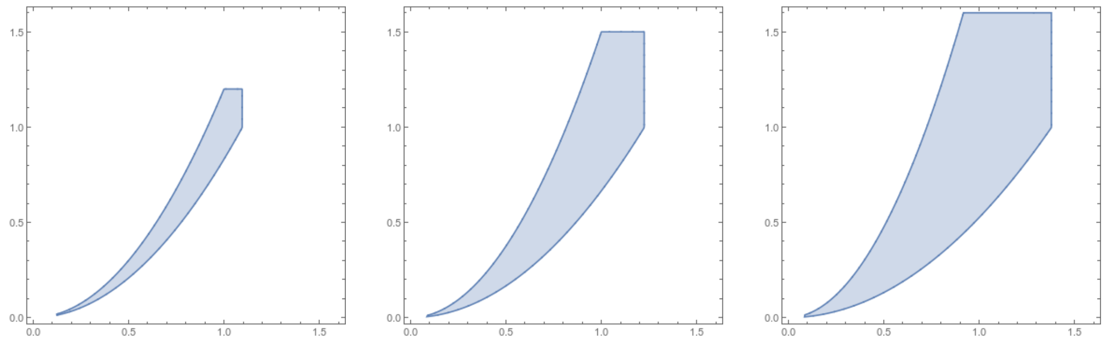
In what follows we will often make use of continuous and compact embeddings of our Sobolev space of mixed smoothness into other spaces. In particular, we need to ensure that weak convergence of u_j in $\mathrm{W}^{\mathbf{a},p}(\Omega)$ implies strong L^p convergence of $\partial^\alpha u_j$ for all α's strictly below the hyperplane of homogeneity, i.e., with $\langle\alpha,\mathbf{a}^{-1}\rangle < 1$. Clearly, this need not hold for a general smoothness; however, in our case, it is guaranteed by the assumption that we have a maximal 'pure' derivative in each direction, since the hyperplane of homogeneity intersect each axis at an integer point. The embeddings that we mention here are shown in the previously mentioned book by Besov, Il'in and Nikolskii (see [23], in particular section 9), to which we refer for proofs.

Definition 2.2.1. *Let* $b \in \mathbb{R}^N$ *be a vector with non-zero coordinates. Fix* $h \in (0,\infty)$ *and* $\varepsilon \in (0,\infty)$. *The set*

$$V(b,h,\varepsilon) := \bigcup_{0<v<h} \left\{ x \in \mathbb{R}^N\colon \frac{x_i}{b_i} > 0, v < (\frac{x_i}{b_i})^{a_i} < (1+\varepsilon)v \text{ for all } i \in \{1,2\ldots,N\} \right\}$$

is called an **a***-horn of radius* h *and opening* ε*.*

Let us remark that in the isotropic case (i.e., $a_i = a_j$ for all i, j) the **a**-horn is in fact a cone and the appropriate conditions we will state shortly are equivalent to the, more familiar, cone conditions. In the genuinely anisotropic scenario the **a**-horn looks like a curved cone. For the reader's convenience we include below a few sample pictures of **a**-horns with $\mathbf{a} = (2, 1)$, $b = (1, 1)$, $h = 1$ and $\varepsilon = 0.2$, 0.5, and 0.9 respectively.



Definition 2.2.2. *Let* $\Omega \subset \mathbb{R}^N$ *be open and let* $K \in \mathbb{N}$*. Suppose that for* $k \in \{1, 2, \ldots, K\}$ *there exist open sets* Ω_k *and* **a***-horns* V_k *(with coefficients* b_k, h_k, ε_k *depending on* k*) such that*

$$\Omega = \bigcup_{k=1}^{K} \Omega_k = \bigcup_{k=1}^{K} (\Omega_k + V_k).$$

Then we say that Ω *satisfies the weak* **a***-horn condition.*

Definition 2.2.3. *We say that* $\Omega \subset \mathbb{R}^N$ *satisfies the strong* **a***-horn condition if* Ω *satisfies the weak* **a***-horn condition with some open sets* Ω_k *and furthermore*

$$\Omega = \bigcup_{k=1}^{K} \Omega_k^{[\delta]} \text{ for some } \delta > 0,$$

where

$$\Omega_k^{[\delta]} := \{x \in \Omega_k \colon \operatorname{dist}(x, \Omega \setminus \Omega_k) > \delta\}.$$

Domains satisfying the **a**-horn condition were first studied by Besov and Il'in (see [22]). The idea behind the condition is simply that the respective embeddings are proven using integral representations, and the horns carry the support of these integral representations.

The following is crucial for us and may be found in [23] or [70].

Lemma 2.2.4. *Any set of the form* $(l_1, r_1) \times \ldots \times (l_N, r_N) \subset \mathbb{R}^N$ *for some* $l_i, r_i \in \mathbb{R}$ *satisfies the strong* **a***-horn condition.*

We note that the **a**-horn condition is not trivial and there are some very 'nice' domains that do not satisfy the above conditions. For example (see [23]) the two-dimensional disc only satisfies the weak **a**-horn condition if $\frac{1}{2}a_1 \leqslant a_2 \leqslant 2a_1$ and does not satisfy the strong **a**-horn condition unless $a_1 = a_2$. Finally, let us remark that in the particular case of sets of the above form (which are often called open boxes, but here we will reserve the term 'box' for a more specific situation) a bounded extension to the whole $\mathbb{R}^N$ may be constructed using the Hestenes' method as noted by Burenkov and Fain in [29], and clearly the existence of a bounded extension operator, together with embeddings for the whole space, implies existence of embeddings on domains as well.

Theorem 2.2.5 (see Theorem 9.5 in [23]). *Suppose that an open set* $\Omega \subset \mathbb{R}^N$ *satisfies the weak* **a***-horn condition and let* $p \in (1, \infty)$*. Then there exists a real number* $h_0 \in (0, \infty)$ *depending on* Ω *and a constant* C *such that for all* $h \in (0, h_0)$ *and all* $f \in \mathrm{W}^{\mathbf{a},p}(\Omega)$ *one has, for all multiindices* $\beta \in \mathbb{Z}_+^N$ *with* $\langle \beta, \mathbf{a}^{-1} \rangle \leqslant 1$*, that*

$$\|\partial^\beta f\|_p \leqslant C \left(h^{1-\langle \beta, \mathbf{a}^{-1} \rangle} \sum_{i=1}^{N} \|\partial_i^{a_i} f\|_p + h^{-\langle \beta, \mathbf{a}^{-1} \rangle} \|f\|_p \right).$$

Before we proceed, let us note the following:

Proposition 2.2.6. *Let* $\Omega \subset \mathbb{R}^N$ *be a bounded open set with a Lipschitz boundary. Then the embedding of* $\mathrm{W}^{\mathbf{a},p}(\Omega)$ *into* $\mathrm{L}^p(\Omega)$ *is compact.*

Proof. For any **a** one has the inclusion $\mathrm{W}^{\mathbf{a},p}(\Omega) \subset \mathrm{W}^{1,p}(\Omega)$ into the standard Sobolev space. Since the inclusion $\mathrm{W}^{1,p}(\Omega) \subset \mathrm{L}^p(\Omega)$ is compact the proof is finished. □

The two above results easily yield the following

Lemma 2.2.7. *Suppose that a bounded open set* $\Omega \subset \mathbb{R}^N$ *with a Lipschitz boundary satisfies the weak* **a***-horn condition and let* $p \in (1, \infty)$*. Then for any* β *with* $\langle \beta, \mathbf{a}^{-1} \rangle < 1$ *the mapping* $\mathrm{W}^{\mathbf{a},p}(\Omega) \hookrightarrow \mathrm{L}^p(\Omega)$ *given by* $u \mapsto \partial^\beta u$ *is completely continuous, i.e., if* $u_j \rightharpoonup u$ *in* $\mathrm{W}^{\mathbf{a},p}(\Omega)$ *then* $\partial^\beta u_j \to \partial^\beta u$ *in* $\mathrm{L}^p(\Omega)$.

Proof. Considering $u_j - u$ instead of u_j we may assume that $u = 0$. Fix a β with $\langle \beta, \mathbf{a}^{-1} \rangle < 1$ and note that Theorem 2.2.5 shows that there exists an $h_0 > 0$ such that, for all $h \in (0, h_0)$, one has

$$\|\partial^\beta u_j\|_p \leqslant C \left(h^{1-\langle \beta, \mathbf{a}^{-1} \rangle} \left(\sum_{\langle \alpha, \mathbf{a}^{-1} \rangle = 1} \|\partial^\alpha u\|_p \right) + h^{-\langle \beta, \mathbf{a}^{-1} \rangle} \|u_j\|_p \right).$$

Proposition 2.2.6 implies that u_j converges strongly to 0 in L^p. Hence, there exists a sequence $h_j \in (0, h_0)$ with $h_j \to 0$ and $Ch_j^{-\langle\beta,\mathbf{a}^{-1}\rangle}\|u_j\|_p \to 0$ as $j \to \infty$. Finally, because u_j is bounded in $\mathrm{W}^{\mathbf{a},p}$ we know that $\left(\sum_{\langle\alpha,\mathbf{a}^{-1}\rangle=1} \|\partial^\alpha u\|_p\right)$ is bounded, and since $Ch_j^{1-\langle\beta,\mathbf{a}^{-1}\rangle} \to 0$ we conclude that $\|\partial^\beta u_j\|_p \to 0$, which ends the proof. □

Another consequence of Theorem 2.2.5 is

Proposition 2.2.8. *Let $\Omega \subset \mathbb{R}^N$ satisfy the weak* **a***-horn condition. Then all of the following*

$$\begin{aligned} \|u\| &:= \|u\|_p + \sum_{i=1}^{N} \|\partial_i^{a_i} u\|_p, \\ \|u\| &:= \|u\|_p + \sum_{\langle\alpha,\mathbf{a}^{-1}\rangle=1} \|\partial^\alpha u\|_p, \\ \|u\| &:= \sum_{\langle\beta,\mathbf{a}^{-1}\rangle\leqslant 1} \|\partial^\beta u\|_p, \end{aligned} \tag{2.1}$$

yield equivalent norms on $\mathrm{W}^{\mathbf{a},p}(\Omega)$.

As in the usual case, membership in the Sobolev space of mixed smoothness yields higher integrability on non-maximal derivatives, as shown in the following result:

Theorem 2.2.9 (see Theorem 10.2 in [23])**.** *Suppose that an open set G satisfies the weak* **a***-horn condition and fix a multiindex α. Let $1 \leqslant p \leqslant q \leqslant \infty$ be such that, with χ defined as $\chi := \langle\alpha, \mathbf{a}^{-1}\rangle + (\frac{1}{p} - \frac{1}{q})|\mathbf{a}^{-1}|$, we have $\chi \leqslant 1$. For $\chi = 1$ suppose additionally that either $1 < p < q < \infty$ or $p = 1$ and $q = \infty$. Then ∂^α continuously maps $\mathrm{W}^{\mathbf{a},p}(G)$ into $\mathrm{L}^q(G)$. That is, for any $f \in \mathrm{W}^{\mathbf{a},p}(G)$ its distributional derivative $\partial^\alpha f$ belongs to $\mathrm{L}^q(G)$ and there exist numbers $h_0 > 0$ and $C > 0$ such that*

$$\|\partial^\alpha f\|_{\mathrm{L}^q(G)} \leqslant Ch^{1-\chi} \sum_{i=1}^{N} \|\partial_i^{a_i} f\|_{\mathrm{L}^p(G)} + Ch^{-\chi}\|f\|_{\mathrm{L}^p(G)},$$

where the constant C is independent of f and $h \in (0, h_0)$.

Finally, we state the following extension result:

Theorem 2.2.10 (see Theorem 9.6 in [23])**.** *Let Ω be an open set satisfying the strong* **a***-horn condition and let $p \in (1, \infty)$. Then the space $\mathrm{W}^{\mathbf{a},p}(\Omega)$ coincides with the restriction of the space $\mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N)$ to Ω. Furthermore, there exists a bounded linear operator extending function in $\mathrm{W}^{\mathbf{a},p}(\Omega)$ to $\mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N)$.*

2.3 Canonical Projection

For $u \in \mathrm{W}^{\mathbf{a},p}$ we write $\nabla_{\mathbf{a}} u$ for the **a**-gradient of u given by $\nabla_{\mathbf{a}} u := (\partial^\alpha u)_{\langle \alpha, \mathbf{a}^{-1}\rangle = 1}$. For future use let us denote the cardinality of the set $\{\alpha \in \mathbb{Z}_+ \colon \langle \alpha, \mathbf{a}^{-1}\rangle = 1\}$ by m, so that for $u \in \mathrm{W}^{\mathbf{a},p}(\Omega, \mathbb{R}^n)$ the **a**-gradient is a map defined on Ω with values in $\mathbb{R}^{n\times m}$. In what follows we will often need to carry out certain operations, for example truncations, on **a**-gradients of various functions. These are easy to do on mappings $\Omega \to \mathbb{R}^{n\times m}$, but they need not preserve the **a**-gradient structure, and to remedy that we turn to Canonical Sobolev Projections following Pełczyński's work in [142]. The aim is to obtain an analogue of the Helmholtz decomposition for the mixed smoothness setting. We then wish to use it for regularizing generating sequences of certain Young measures similarly to what has been done by Fonseca and Müller in [67] in the case of $\mathcal{A}$-free vector fields and $\mathcal{A}$-quasiconvexity. In this context see also a recent article [145] where the ideas of [67] have been expanded to deal with extended real valued integrands.

As mentioned before, there is a canonical embedding

$$\mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N;\mathbb{R}^n) \to \bigoplus_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \mathrm{L}^p(\mathbb{R}^N;\mathbb{R}^n)$$

given by

$$u \mapsto (\partial^\alpha u)_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1},$$

but it is not surjective. With $p = 2$ one may define the canonical projection of the target space onto the image of this embedding, i.e.,

$$\mathrm{P}_{\mathbf{a}} \colon \bigoplus_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \mathrm{L}^2(\mathbb{R}^N;\mathbb{R}^n) \to \mathrm{Im}\left(\mathrm{W}^{\mathbf{a},2}(\mathbb{R}^N;\mathbb{R}^n) \to \bigoplus_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \mathrm{L}^2(\mathbb{R}^N;\mathbb{R}^n) \right). \tag{2.2}$$

Let us look at what this means in the concrete case of our model example $N = 2$ and $\mathbf{a}^{-1} = (1, \frac{1}{2})$, i.e. $\left\{\alpha \in \mathbb{Z}_+^2 \colon \langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1\right\} = \{(0,0),(1,0),(0,1),(0,2)\rangle\}$. The canonical embedding is then the mapping

$$\mathrm{W}^{\mathbf{a},2}(\mathbb{R}^2;\mathbb{R}^n) \ni u \mapsto (u, \partial_{x_1} u, \partial_{x_2} u, \partial^2_{x_2} u) \in \bigoplus_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \mathrm{L}^2(\mathbb{R}^2;\mathbb{R}^n).$$

Thus the image of this embedding is the space

$$\mathrm{Im}\left(\mathrm{W}^{\mathbf{a},2}(\mathbb{R}^2;\mathbb{R}^n) \to \bigoplus_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \mathrm{L}^2(\mathbb{R}^2;\mathbb{R}^n) \right) =$$

$$\left\{(v^1, v^2, v^3, v^4)\colon v^i \in \mathrm{L}^2(\mathbb{R}^2;\mathbb{R}^n),\ v^2 = \partial_{x_1} v^1,\ v^3 = \partial_{x_2} v^1,\ v^4 = \partial^2_{x_2} v^1\right\}.$$

Clearly not all vector fields $v = (v^1, v^2, v^3, v^4)$ with $v^i \in \mathrm{L}^2(\mathbb{R}^2;\mathbb{R}^n)$ satisfy the relations specified above, and this is where the projection comes into play.

It has been shown in [142] (see Corollary 5.1 therein) that the projection from (2.2) is of strong type (p, p) for $1 < p < \infty$, thus one can extend it by continuity from

$$\bigoplus_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \mathrm{L}^2(\mathbb{R}^N;\mathbb{R}^n) \to \bigoplus_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \mathrm{L}^2(\mathbb{R}^N;\mathbb{R}^n)$$

to

$$\bigoplus_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \mathrm{L}^p(\mathbb{R}^N;\mathbb{R}^n) \to \bigoplus_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \mathrm{L}^p(\mathbb{R}^N;\mathbb{R}^n).$$

Lemma 2.3.1. *Fix any $p \in (1, \infty)$ and denote by $\mathrm{P}_{\mathbf{a}}$ the extension of the canonical projection discussed above. Then:*

i) the map $\mathrm{P}_{\mathbf{a}}$ is a bounded linear operator on $\bigoplus_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \mathrm{L}^p(\mathbb{R}^N;\mathbb{R}^n)$;

ii) for any $u \in \bigoplus_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \mathrm{L}^p$ we have $\mathrm{P}_{\mathbf{a}}(\mathrm{P}_{\mathbf{a}} u) = \mathrm{P}_{\mathbf{a}} u$;

iii) if the family $\{u_j\} \subset \bigoplus_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \mathrm{L}^p$ is p-equiintegrable then so is $\{\mathrm{P}_{\mathbf{a}} u_j\}$.

Proof. The first assertion is the content of Corollary 5.1 in [142]. Point ii) is, by definition, true in $\bigoplus_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \mathrm{L}^2(\mathbb{R}^N;\mathbb{R}^n)$. For a general exponent p let us fix $u \in \bigoplus_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \mathrm{L}^p(\mathbb{R}^N;\mathbb{R}^n)$ and a family $u_j \subset \bigoplus_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \mathrm{L}^2(\mathbb{R}^N;\mathbb{R}^n)$ with $\|u - u_j\|_p \to 0$. By continuity of $\mathrm{P}_{\mathbf{a}}$ (and thus of $\mathrm{P}_{\mathbf{a}} \circ \mathrm{P}_{\mathbf{a}}$) one has

$$\mathrm{P}_{\mathbf{a}} u_j \to \mathrm{P}_{\mathbf{a}} u \ \text{ and } \ \mathrm{P}_{\mathbf{a}}(\mathrm{P}_{\mathbf{a}} u_j) \to \mathrm{P}_{\mathbf{a}}(\mathrm{P}_{\mathbf{a}} u) \ \text{ in } \ \mathrm{L}^p(\mathbb{R}^N;\mathbb{R}^n),$$

and since $\mathrm{P}_{\mathbf{a}} u_j = \mathrm{P}_{\mathbf{a}}(\mathrm{P}_{\mathbf{a}} u_j)$ for each j the claim is proven.

For the last part consider the standard truncation τ_k given by

$$\tau_k(X) := \begin{cases} X & \text{if } |X| \leqslant k, \\ k\frac{X}{|X|} & \text{if } |X| > k. \end{cases} \tag{2.3}$$

Then fix any sequence u_j satisfying the assumptions of point iii). Since $\{\tau_k(u_j)\}$ is bounded in L^∞ and in L^p we know, by continuity of $\mathrm{P}_{\mathbf{a}}$ as a map from L^q to L^q, that $\{\mathrm{P}_{\mathbf{a}} \tau_k(u_j)\}$ is bounded in any L^q with $p \leqslant q < \infty$, so that this family is equiintegrable in L^p. Then again, p-equiintegrability of $\{u_j\}$ itself yields

$$\lim_{k\to\infty} \sup_j ||u_j - \tau_k(u_j)||_p = 0,$$

so again continuity of $\mathrm{P}_{\mathbf{a}} : \mathrm{L}^p \to \mathrm{L}^p$ gives

$$\lim_{k\to\infty} \sup_j ||\mathrm{P}_{\mathbf{a}}(u_j - \tau_k(u_j))||_p = 0,$$

hence $\{\mathrm{P}_{\mathbf{a}}(u_j)\}$ is p-equiintegrable as claimed.

□

Remark 1. Note that in the above we abuse notation slightly by treating $\mathrm{P}_{\mathbf{a}}$ as an operator acting on $\bigcup_{p\in(1,\infty)} \mathrm{L}^p$, however this is justified by the fact that, in the Fourier space, the operator is given by a multiplier independent of p, see [142].

2.4 Anisotropic scaling

As mentioned before, we will often need to use a specific anisotropic scaling. Thus, for a real number $R > 0$ and a vector $v = (v_1, v_2, \ldots, v_N) \in \mathbb{R}^N$ we let $R \odot v := (R^{1/a_1}v_1, R^{1/a_2}v_2, \ldots, R^{1/a_N}v_N)$. Here, and in all that follows, $Q \subset \mathbb{R}^N$ will, unless otherwise specified, denote the unit (in the ∞-norm) cube centred at 0. We let $Q_R(x_0) \subset \mathbb{R}^N$ be the open box centred at $x_0 \in \mathbb{R}^N$ and scaled according to the rule just described, so that

$$Q_R(x_0) := \{x \in \mathbb{R}^N \colon |x^i - x_0^i|^{a_i} < R \text{ for all } 1 \leqslant i \leqslant N\}.$$

Equivalently, we could write $Q_R(x_0) = x_0 + R \odot Q$. We call R the anisotropic radius of the box Q_R and x_0 its centre. This is a slight abuse of language, as we will see later that the boxes discussed here correspond to balls in an appropriate anisotropic metric, but with radius $R^{1/\max_i a_i}$, however we have decided to stick to the notation dictated by the scaling. From now on we will always understand a 'box' to mean a set of the form above, i.e., an anisotropically scaled and translated unit cube.

Observe that, unless the scaling is in fact isotropic (that is, $a_i = a_j$ for all i, j), our family of boxes is not of bounded eccentricity, i.e., there does not exist a constant $c > 0$ such that each box $Q_R(x_0)$ in our collection is contained in some (Euclidean) ball B with $|Q_R(x_0)| \geqslant c|B|$. Therefore, it is not obvious if standard results such as the Vitali covering lemma, or the Lebesgue differentiation theorem, hold for balls replaced by anisotropically scaled boxes. Note that even sharper statements of the Vitali covering lemma, such as the one in [149] by Saks (which only requires the eccentricity to be bounded along fixed sequences converging to a given point), or the one in [124] by Mejlbro and Topsøe (where the condition on the eccentricity is in integral form), are not directly applicable. Nevertheless, it turns out that the aforementioned results still hold, and the purpose of this section is to prove it. We follow the approach given in [17] for the case of (Euclidean) cubes. The argument is nearly the same but, for the sake of completeness, we include a short presentation of how to obtain an elementary version of the Vitali covering lemma adapted to our setting. With the help of this result the Lebesgue differentiation theorem follows easily.

We remark here that the results of the present subsection may be deduced from a more general work by Calderón and Torchinsky (see [33]). Nevertheless, we have decided to present a more straightforward approach to the particular case that we consider here, as the argument is not very long, and the Lebesgue differentiation theorem then follows immediately.

Lemma 2.4.1 (Elementary anisotropic Vitali covering lemma)**.** *Let A be a Lebesgue measurable subset of $\mathbb{R}^N$ with finite Lebesgue measure. For any covering of A by a family $\{Q_i\}$ of anisotropic boxes there exists a finite, pairwise disjoint, subfamily $\{Q_1, \ldots, Q_K\}$ such that*

$$|A| \leqslant 3^{|\mathbf{a}^{-1}|} \sum_{j=1}^{K} |Q_j|.$$

Proof. Without loss of generality assume that A is compact. Since we are using open boxes there is a finite subfamily $\mathcal{Q} := \{\widetilde{Q_1}, \ldots, \widetilde{Q_k}\}$ that covers A. Let Q_1 be an element of $\mathcal{Q}$ with the largest radius (arbitrary if there is more than one). Assume that $Q_1, \ldots, Q_j$ have been chosen and let Q_{j+1} be an element of

$$\{\widetilde{Q} \in \mathcal{Q} \colon \widetilde{Q} \cap Q_i = \emptyset \text{ for all } 1 \leqslant i \leqslant j\}$$

with the largest radius. Finish the construction when the set

$$\{\widetilde{Q} \in \mathcal{Q} \colon \widetilde{Q} \cap Q_i = \emptyset \text{ for all } 1 \leqslant i \leqslant j\}$$

is empty. We claim that the family $\{3 \odot Q_1, \ldots, 3 \odot Q_K\}$ covers A, where $3 \odot Q_i$ denotes a box with the same centre as Q_i but triple the anisotropic radius.

Fix an arbitrary $x \in A$. If $x \in \bigcup_{i=1}^{K} Q_i$ then there is nothing to show, thus assume otherwise. There has to exist an element $\widetilde{Q}$ of $\mathcal{Q}$ such that $x \in \widetilde{Q}$, since $\mathcal{Q}$ covers A. Let R be the radius of $\widetilde{Q}$. Since $\widetilde{Q} \notin \{Q_1, \ldots Q_K\}$ we must have $\widetilde{Q} \cap \bigcup_{i=1}^{K} Q_i \neq \emptyset$. Let Q_{i_0} be an element of those among $\{Q_1, \ldots Q_K\}$ that intersect $\widetilde{Q}$ and suppose that Q_{i_0} has maximal radius, say R_0. Then $R_0 \geqslant R$, as otherwise $\widetilde{Q}$ would have been chosen instead of Q_{i_0}. Let x_0 be the centre of Q_{i_0}. For any $1 \leqslant j \leqslant N$ we must have

$$|x_0^j - x^j| \leqslant R_0^{a_j} + 2R^{a_j} \leqslant 3R_0^{a_j} \leqslant (3R_0)^{a_j},$$

since $x \in \widetilde{Q}$ and $Q_{i_0} \cap \widetilde{Q} \neq \emptyset$. However, $|x_0^j - x^j| \leqslant (3R_0)^{a_j}$ for all $1 \leqslant j \leqslant N$, which means precisely that $x \in 3 \odot Q_{i_0}$. Finally, since $|3 \odot Q_{i_0}| = 3^{|\mathbf{a}^{-1}|}|Q_{i_0}|$ the proof is complete. □

Definition 2.4.2. *Let $f \in \mathrm{L}^1_{loc}(\mathbb{R}^N)$. The anisotropic Hardy-Littlewood maximal function $M(f)$ of f is defined as*

$$M(f)(x) := \sup_{R>0} \frac{1}{|Q_R(x)|} \int_{Q_R(x)} |f(y)| \, \mathrm{d}y,$$

for $x \in \mathbb{R}^N$.

As a consequence of the anisotropic Vitali covering lemma we immediately obtain the following two results, which may be proven exactly as in [17]:

Lemma 2.4.3 (Anisotropic Hardy-Littlewood lemma)**.** *For any $f \in \mathrm{L}^1_{loc}(\mathbb{R}^N)$ and for any $t > 0$ one has*

$$|\{x \in \mathbb{R}^N \colon M(f)(x) > t\}| \leqslant \frac{3^{|\mathbf{a}^{-1}|}}{t} \int_{\mathbb{R}^N} |f(y)| \, \mathrm{d}y.$$

Theorem 2.4.4 (Anisotropic Lebesgue differentiation theorem)**.** *Let $f \in \mathrm{L}^1_{loc}(\mathbb{R}^N)$. For Lebesgue almost every $x_0 \in \mathbb{R}^N$ one has*

$$\limsup_{R\to 0} \frac{1}{|Q_R(x_0)|} \int_{Q_R(x_0)} |f(x) - f(x_0)| \, \mathrm{d}x = 0.$$

2.5 Polynomial approximation

Recall that Q, unless otherwise specified, is the unit (in the ∞-norm) cube centered at 0, i.e., $Q = \{x \in \mathbb{R}^N \colon |x^i| < 1 \text{ for all } 1 \leqslant i \leqslant N\}$, whereas $Q_R(x_0)$ is a box obtained from Q through a composition of a translation (so that the center is at x_0) and anisotropic scaling, i.e.,

$$Q_R(x_0) := \{x \in \mathbb{R}^N \colon |x^i - x_0^i|^{a_i} < R \text{ for all } 1 \leqslant i \leqslant N\}.$$

For a function $f \in \mathrm{L}^1(Q_R)$ we denote by $(f)_{Q_R(x_0)}$ its average over Q_R, i.e.,

$$(f)_{Q_R(x_0)} := \fint_{Q_R(x_0)} f(x) \, \mathrm{d}x.$$

Lemma 2.5.1. *There exists a constant C such that, for any $r > 0$ and any $\sigma \in (0, \frac{1}{2})$, there exists a cut-off function $\eta \in C_c^\infty(Q_r; [0,1])$ which is identically equal to 1 on $Q_{(1-\sigma)r}$ and satisfies*

$$\|\partial^\beta \eta\|_{\mathrm{L}^\infty} \leqslant C r^{-\langle \beta, \mathbf{a}^{-1} \rangle} \sigma^{-|\beta|},$$

for all multi-indices β with $\langle \beta, \mathbf{a}^{-1} \rangle \leqslant 1$.

Proof. First of all, note that it is enough to consider the case $r = 1$, as the general result will then follow by our anisotropic scaling. With $r = 1$ observe that the distance between the faces of Q and $Q_{1-\sigma}$ along the x_i axis is equal to $1 - (1-\sigma)^{1/a_i}$. There exists a constant $C > 0$ such that for all $\sigma \in (0, \frac{1}{2})$ we have

$$1 - (1-\sigma)^{1/a_i} \geqslant C\sigma,$$

for all i. This follows simply by observing that we have equality with $\sigma = 0$, and for sufficiently small C the function $C\sigma + (1-\sigma)^{1/a_i}$ is decreasing, which may be easily checked by differentiation. Now it is enough to construct one dimensional cut-off functions $\eta^i(x_i)$ that realise the desired cut-off along the particular axes and satisfy $\|\partial^k \eta^i\|_{\mathrm{L}^\infty} \leqslant 2(C\sigma)^{-k}$, and then consider η of the form $\eta(x) := \prod_{i=1}^N \eta^i(x_i)$. □

We recall the following version of the Poincaré inequality in $\mathrm{W}^{\mathbf{a},p}$ proven by Dupont and Scott in [56], where, instead of requiring zero boundary values, we allow for correction in terms of the kernel of the operator $\nabla_{\mathbf{a}}$.

Proposition 2.5.2 (see Theorem 4.2 in [56])**.** *There exists a constant C such that for any function $f \in \mathrm{W}^{\mathbf{a},p}(Q)$ with $p \in [1,\infty)$ there exists a polynomial $P_f \in C^\infty(Q)$ with $\nabla_{\mathbf{a}} P_f \equiv 0$ satisfying*

$$\|f - P_f\|_{\mathrm{W}^{\mathbf{a},p}(Q)} \leqslant C \|\nabla_{\mathbf{a}} f\|_{\mathrm{L}^p(Q)}.$$

Observe that, in the above, Q is fixed to be the unit cube, and we do not assert anything about approximations on other domains. However, we will only ever use this result on anisotropic boxes, and it is easy to see how to adjust the constant to scaling, as shown in the following:

Corollary 2.5.3. *There exists a constant C such that for any function $f \in \mathrm{W}^{\mathbf{a},p}(Q_r)$ with $p \in [1,\infty)$ there exists a polynomial $P_f \in C^\infty(Q_r)$ with $\nabla_{\mathbf{a}} P_f \equiv (\nabla_{\mathbf{a}} f)_{Q_r}$ such that for any β with $\langle \beta, \mathbf{a}^{-1}\rangle \leqslant 1$ we have*

$$r^{-1+\langle \beta, \mathbf{a}^{-1}\rangle} \|\partial^\beta (f - P_f)\|_{\mathrm{L}^p(Q_r)} \leqslant C \|\nabla_{\mathbf{a}} (f - P_f)\|_{\mathrm{L}^p(Q_r)}.$$

The constant C does not depend on the function f nor the radius r.

Proof. First of all, note that using our anisotropic rescaling we may reduce to the case $r = 1$. This also determines the scaling, i.e., the $r^{-1+\langle \beta, \mathbf{a}^{-1}\rangle}$ factor. Secondly, it is clearly enough to prove the result for $f \in C^\infty(Q)$, as the general case then follows from density of smooth functions in $\mathrm{W}^{\mathbf{a},p}(Q)$. Observe that by considering

$$\widetilde{f}(x) := f(x) - \sum_{\langle \alpha, \mathbf{a}^{-1}\rangle = 1} x^\alpha \left(\partial^\alpha f\right)_Q,$$

we reduce our task to finding a polynomial $\widetilde{P}_{\widetilde{f}}$ with $\nabla_{\mathbf{a}}\widetilde{P}_{\widetilde{f}} \equiv 0$ and such that

$$\|\partial^\beta(\widetilde{f} - \widetilde{P}_{\widetilde{f}})\|_{\mathrm{L}^p(Q)} \leqslant C\|\nabla_{\mathbf{a}}\widetilde{f}\|_{\mathrm{L}^p(Q)}$$

for all β with $\langle\beta, \mathbf{a}^{-1}\rangle < 1$, as the case $\langle\beta, \mathbf{a}^{-1}\rangle = 1$ is trivial. The existence of such a $\widetilde{P}_{\widetilde{f}}$ is the content of Proposition 2.5.2, which completes the proof.

□

Proposition 2.5.4 (see Theorem 10.16 in [43]). *For any bounded open domain Ω, any $u \in \mathrm{W}^{\mathbf{a},p}(\Omega)$ and any $\varepsilon > 0$ there exist a function $u_\varepsilon \in \mathrm{W}^{\mathbf{a},p}_u(\Omega)$ and a finite family of disjoint boxes $\{Q_{\varepsilon,i}\}_i$ such that $\nabla_{\mathbf{a}}u_\varepsilon$ is constant on each $Q_{\varepsilon,i} \subset \Omega$, $|\Omega \setminus \bigcup_i Q_{\varepsilon,i}| < \varepsilon$, and $\|u - u_\varepsilon\|_{\mathrm{W}^{\mathbf{a},p}(\Omega)} < \varepsilon$.*

Proof. Fix an arbitrary $u \in \mathrm{W}^{\mathbf{a},p}(\Omega)$ and a parameter $\tau \in (0, 1)$ to be determined later. Decompose Ω, up to a set of measure zero, into a countable family of disjoint, open boxes $\{Q_{\tau,i}\}$ of radii equal, or smaller than, τ. Select a finite subfamily of I boxes covering Ω up to a set of measure less than $\varepsilon/2$ and relabel the elements so that $\left|\Omega \setminus \bigcup_{i=1}^I Q_{\tau,i}\right| < \varepsilon/2$. From now on we focus our attention only on these boxes $Q_{\tau,i}$ with $1 \leqslant i \leqslant I$. Let $P_{\tau,i}$ denote the polynomial approximating u on $Q_{\tau,i}$ given by Corollary 2.5.3. Let $\sigma \in (0, 1/2)$ be a parameter to be determined later. For every $Q_{\tau,i}$ take a cut-off function $\eta_{\tau,i} \in C_c^\infty(Q_{\tau,i})$ identically equal to 1 on $(1 - \sigma) \odot Q_{\tau,i}$, as in Lemma 2.5.1.

Define

$$v(x) := u(x) + \sum_{i=1}^{I} \eta_{\tau,i}(x)\left(P_{\tau,i}(x) - u(x)\right),$$

so that $\nabla_{\mathbf{a}}v$ is constant on each $(1-\sigma)\odot Q_{\tau,i}$ and $v \in \mathrm{W}^{\mathbf{a},p}_u(\Omega)$. We may now calculate

$$\|u - v\|^p_{\mathrm{W}^{\mathbf{a},p}(\Omega)} = \sum_{\langle\beta,\mathbf{a}^{-1}\rangle\leqslant 1} \int_\Omega |\partial^\beta(u - v)|^p \,\mathrm{d}x =$$

$$\sum_{\langle\beta,\mathbf{a}^{-1}\rangle\leqslant 1} \sum_{i=1}^{I} \int_{Q_{\tau,i}} \left|\partial^\beta\left(\eta_{\tau,i}(x)\left(u(x) - P_{\tau,i}(x)\right)\right)\right|^p \mathrm{d}x \leqslant$$

$$\sum_{\langle\beta,\mathbf{a}^{-1}\rangle\leqslant 1} \sum_{i=1}^{I} \int_{Q_{\tau,i}} \sum_{0\leqslant\gamma\leqslant\beta} |\partial^\gamma\eta_{\tau,i}(x)|^p \left|\partial^{\beta-\gamma}\left(u(x) - P_{\tau,i}(x)\right)\right|^p \mathrm{d}x.$$

Using the bounds on the derivatives of η we may write

$$\|u - v\|^p_{\mathrm{W}^{\mathbf{a},p}(\Omega)} \leqslant C \sum_{\langle\beta,\mathbf{a}^{-1}\rangle\leqslant 1} \sum_{0\leqslant\gamma\leqslant\beta} \sum_{i=1}^{I} \sigma^{-p|\beta|}\tau^{-p\langle\gamma,\mathbf{a}^{-1}\rangle} \int_{Q_{\tau,i}} \left|\partial^{\beta-\gamma}\left(u(x) - P_{\tau,i}(x)\right)\right|^p \mathrm{d}x.$$

Using the bound from Corollary 2.5.3 on each $Q_{\tau,i}$ now yields

$$\|u-v\|^p_{\mathrm{W}^{\mathbf{a},p}(\Omega)} \leqslant C \sum_{\langle\beta,\mathbf{a}^{-1}\rangle\leqslant 1}\sum_{0\leqslant\gamma\leqslant\beta}\sum_{i=1}^{I}\sigma^{-p|\beta|}\tau^{-p\langle\gamma,\mathbf{a}^{-1}\rangle}\tau^{p-p\langle\beta-\gamma,\mathbf{a}^{-1}\rangle}\|\nabla_{\mathbf{a}}(u-P_{\tau,i})\|^p_{\mathrm{L}^p(Q_{\tau,i})}.$$

Thus,

$$\|u-v\|^p_{\mathrm{W}^{\mathbf{a},p}(\Omega)} \leqslant C \sum_{\langle\beta,\mathbf{a}^{-1}\rangle\leqslant 1}\sum_{0\leqslant\gamma\leqslant\beta}\sum_{i=1}^{I}\sigma^{-p|\beta|}\tau^{p-p\langle\beta,\mathbf{a}^{-1}\rangle}\|\nabla_{\mathbf{a}}(u-P_{\tau,i})\|^p_{\mathrm{L}^p(Q_{\tau,i})},$$

and finally, since $\langle\beta,\mathbf{a}^{-1}\rangle\leqslant 1$, $\sigma<1/2$, and $\tau<1$, we may write

$$\|u-v\|^p_{\mathrm{W}^{\mathbf{a},p}(\Omega)} \leqslant C\sigma^{-p\max_j a_j}\sum_{i=1}^{I}\|\nabla_{\mathbf{a}}(u-P_{\tau,i})\|^p_{\mathrm{L}^p(Q_{\tau,i})}.$$

Now it is time to choose the parameters σ and τ. Observe that $\frac{|(1-\sigma)\odot Q_{\tau,i}|}{|Q_{\tau,i}|} = (1-\sigma)^{|\mathbf{a}^{-1}|}$ for any τ and any i. Thus, choosing σ small enough ensures that, with any τ, we will have $|\Omega\setminus(\bigcup_i(1-\sigma)\odot Q_{\tau,i})|<\varepsilon$. With σ fixed it is enough to choose τ small enough so that

$$\sum_{i=1}^{I}\|\nabla_{\mathbf{a}}(u-P_{\tau,i})\|^p_{\mathrm{L}^p(Q_{\tau,i})} \leqslant C^{-1}\sigma^{p\max_j a_j}\varepsilon,$$

which is possible, as due to the anisotropic Lebesgue's differentiation theorem (Theorem 2.4.4), one can approximate $\nabla_{\mathbf{a}}u$ in the L^p norm by its averages over a grid of boxes of sufficiently small radii. Using the corresponding v as u_ε ends the proof. □

2.6 Anisotropic Campanato and Hölder spaces

For future use in the study of regularity of minimisers of anisotropic variational problems we present here a selection of facts about anisotropic Campanato and Hölder spaces. This subsection is based on Giusti's exposition from [77], to which we refer for proofs.

To easily define the anisotropic Hölder spaces we let

$$\delta_{\mathbf{a}}(x,y) := \sup_{1\leqslant j\leqslant N}|x_j-y_j|^{a_j/\max_i a_i}$$

be the anisotropic ∞-metric that corresponds to the anisotropic scaling that we use. The factor $\max_i a_i$ is there to make the function a metric, as it is necessary for the triangle inequality to hold. Balls in this metric correspond to our anisotropic boxes, albeit balls of radius r in the metric $\delta_{\mathbf{a}}$ correspond to boxes with anisotropic radii of $r^{(\max_i a_i)}$.

Definition 2.6.1. *The anisotropic Campanato space* $\mathfrak{L}^{p,\theta}_{\mathbf{a}}(\Omega)$ *is defined as the subset of those functions* $u \in \mathrm{L}^p(\Omega)$ *for which we have*

$$[u]^p_{\mathfrak{L}^{p,\theta}_{\mathbf{a}}(\Omega)} := \sup_{x\in\overline{\Omega},\rho>0} |I(x,\rho)|^{-\theta} \int_{I(x,\rho)} |u(y) - u_{I(x,\rho)}|^p \,\mathrm{d}y < \infty,$$

where $I(x,\rho) = \Omega \cap Q_\rho(x)$ *and* $u_{I(x,\rho)} = \fint_{I(x,\rho)} u(y)\,\mathrm{d}y$. *Equipping* $\mathfrak{L}^{p,\theta}_{\mathbf{a}}(\Omega)$ *with the norm*

$$\|u\|_{\mathfrak{L}^{p,\theta}_{\mathbf{a}}(\Omega)} := \|u\|_{\mathrm{L}^p(\Omega)} + [u]_{\mathfrak{L}^{p,\theta}_{\mathbf{a}}(\Omega)}$$

makes it a Banach space.

Definition 2.6.2. *An open domain* $\Omega \subset \mathbb{R}^N$ *is said to be of type (A) with respect to the metric* δ *if there exists a constant* $C > 0$ *such that for all* $x \in \overline{\Omega}$ *and for all* $r \in (0, diam(\Omega))$ *one has* $|I(x,r)| \geqslant C|B(x,r;\delta)|$, *where* $B(x,r;\delta)$ *denotes the ball in metric* δ *centred at* x *and of radius* r.

Remark 2. Clearly any box with edges parallel to coordinate axes is of type (A) with respect to $\delta_{\mathbf{a}}$.

Lemma 2.6.3 (see Definition 2.IV in [77])**.** *If* Ω *is of type (A) with respect to* $\delta_{\mathbf{a}}$ *then, for* $\theta > 1$, *the space* $\mathfrak{L}^{p,\theta}_{\mathbf{a}}(\Omega, \delta_{\mathbf{a}})$ *is canonically isomorphic to the Hölder space* $C^{0,\chi}(\Omega, \delta_{\mathbf{a}})$ *with* $\chi = \frac{\max_i a_i |\mathbf{a}^{-1}|}{p}(\theta - 1)$.

Lemma 2.6.4 (see Observation 2 in [77])**.** *The space* $C^{0,\chi}(\Omega, \delta_{\mathbf{a}})$ *coincides with the space of functions that are Hölder continuous with respect to the* x_i *variable with Hölder exponent* $\chi_i := \frac{a_i}{\max_j a_j}\chi$, *when considered in the usual Euclidean metric.*

As a final remark let us say that in subsequent parts of the thesis we are only interested in the variational side of the problems considered, and we do not attempt to prove them under optimal conditions on the underlying domain Ω. Therefore, unless otherwise specified, we will always assume that Ω is a bounded open domain with Lipschitz boundary of measure zero, and satisfying the weak **a**-horn condition.

Chapter 3

Young measures

The main technical tool that we will use in studying lower semicontinuity of integral functionals defined on Sobolev spaces of mixed smoothness is the theory of Young measures. The purpose of this chapter is to introduce the principal technical results that we will need. As mentioned in the introduction, Young measures were introduced by Young (see [180], [181], [182], and [183]) and they are objects that describe the behaviour of weakly converging sequences more accurately than just their weak limits. When dealing with nonlinear expressions, it is often difficult to justify passing to the limit over a sequence if it only converges in the weak topology. Working in the space of measures facilitates that, as measures enjoy good compactness properties. Heuristically, the major advantage of the Young measures approach to lower semicontinuity of integral functionals is that it allows to split the problem into two parts. The first step is to pass to the limit in the functional and get an expression describing the action of the limiting Young measure on the integrand. Secondly, one needs to check if that limiting expression satisfies a certain kind of Jensen's inequality to check whether it is greater than the value of the functional evaluated at the (traditional) weak limit. Another benefit of the approach is that Young measures describe the oscillations of the sequence separately at each point of the domain, which further simplifies the problem through localisation.

We begin this chapter by introducing the classical (oscillation) Young measures in Section 3.1, in particular we recall the Fundamental Theorem of Young Measures (see Theorem 3.1.1). In Section 3.2 we prepare for the study of Young measures generated by **a**-gradients, by proving that weakly convergent sequences of **a**-gradients may be decomposed into an equiintegrable part that carries the oscillation and a remainder, converging to 0 in measure, that carries the concentration (see Proposition 3.2.3). This lets us obtain 'better' (here p-equiintegrable) generating sequences for the Young measures under consideration. We briefly discuss generalised (or DiPerna-Majda)

Young measures in Section 3.3, and we prove a localisation result in Proposition 3.4.2 of the following Section. We introduce **a**-quasiconvexity in Section 3.5, and we prove an initial formula for **a**-quasiconvex envelopes in Lemma 3.5.3. We then use it to investigate the structure of Young measures generated by **a**-gradients of functions in $\mathrm{W}^{\mathbf{a},p}$ in Sections 3.6 and 3.7. This culminates in Theorems 3.7.1 and 3.7.2 where we establish duality between such Young measures and **a**-quasiconvex functions, in the spirit of the famous Kinderlehrer-Pedregal characterisation of gradient Young measures.

3.1 Oscillation Young measures

We begin with a general overview of the theory of oscillation Young measures. Throughout this work we will always work with non-negative integrands and in reflexive spaces, so that oscillation Young measures provide enough information to study lower semicontinuity of our functionals. Therefore, we largely focus on them, although in the process we will also mention a few facts about generalised DiPerna-Majda Young measures, that we hope to use in future work in settings where concentration effects must be taken into account.

We denote by $\mathcal{M}(\mathbb{R}^{n\times m})$ the set of all Radon measures on $\mathbb{R}^{n\times m}$, and by $\mathcal{P}(\mathbb{R}^{n\times m}) \subset \mathcal{M}(\mathbb{R}^{n\times m})$ the set of all probability measures on $\mathbb{R}^{n\times m}$. Recall (see, for example, Chapter 6 of [64]) that we say that a function $F\colon \Omega \times \mathbb{R}^{n\times m} \to (-\infty, \infty]$ is a normal integrand if F is Borel measurable and, for every fixed $x \in \Omega$, the function $W \mapsto f(x, W)$ is lower semicontinuous. Similarly, we say that a function $F\colon \Omega \times \mathbb{R}^{n\times m} \to \mathbb{R}$ is Carathéodory if both F and $-F$ are normal integrands. Finally, a map $\nu\colon \Omega \to \mathcal{M}(\mathbb{R}^{n\times m})$ is said to be weak*-measurable if $x \mapsto \langle \nu_x, \varphi\rangle$ is (Lebesgue) measurable for any continuous and compactly supported function $\varphi\colon \mathbb{R}^{n\times m} \to \mathbb{R}$.

We begin with the following version of the Fundamental Theorem of Young Measures which may be found, for example, in Pedregal's book, see [141], followed by a simple result on translations, that may be found in the same book.

Theorem 3.1.1 (see [141])**.** *Let $\Omega \subset \mathbb{R}^N$ be a measurable set of finite measure and $V_j\colon \Omega \to \mathbb{R}^{n\times m}$ be a bounded sequence of L^p functions for some $p \in [1, \infty]$. Then there exists a subsequence V_{j_k} and a weak*-measurable map $\nu\colon \Omega \to \mathcal{M}(\mathbb{R}^{n\times m})$ such that the following hold:*

i) every ν_x is a probability measure;

ii) if $F\colon \Omega\times\mathbb{R}^{n\times m}\to\mathbb{R}\cup\{\infty\}$ *is a normal integrand bounded from below, then*

$$\liminf_{j\to\infty}\int_\Omega F(x,V_{j_k}(x))\,\mathrm{d}x \geqslant \int_\Omega \overline{F}(x)\,\mathrm{d}x,$$

where

$$\overline{F}(x) := \langle \nu_x, F(x,\cdot)\rangle = \int_{\mathbb{R}^{n\times m}} F(x,y)\,\mathrm{d}\nu_x(y);$$

iii) if $F\colon \Omega\times\mathbb{R}^{n\times m}\to\mathbb{R}\cup\{\infty\}$ *is Carathéodory and bounded from below, then*

$$\lim_{j\to\infty}\int_\Omega F(x,V_{j_k}(x))\,\mathrm{d}x = \int_\Omega \overline{F}(x)\,\mathrm{d}x < \infty$$

if and only if $\{F(\cdot,V_{j_k}(\cdot))\}$ *is equiintegrable (in the usual,* L^1 *sense). In this case*

$$F(\cdot,V_{j_k}(\cdot)) \rightharpoonup \overline{F} \text{ in } \mathrm{L}^1(\Omega)\cdot$$

The family $\{\nu_x\}_{x\in\Omega}$ *is called the (oscillation) Young measure generated by* V_{j_k}*. If there exists some* $x_0\in\Omega$ *such that* $\nu_x=\nu_{x_0}$ *for almost every* $x\in\Omega$ *then we say that* ν *is a homogeneous (oscillation) Young measure and often identify the family* $\{\nu_x\}$ *with the single measure* ν_{x_0} *if there is no risk of confusion.*

Proposition 3.1.2 (see [141])**.** *If* $\{V_j\}$ *generates an oscillation Young measure* ν *and if* $W_j\to W$ *in measure, then* $\{V_j+W_j\}$ *generates the translated Young measure*

$$\widetilde{\nu}_x := \delta_{W(x)} * \nu_x,$$

where

$$\langle \delta_U * \mu, \varphi\rangle = \langle \mu, \varphi(\cdot + U)\rangle$$

for $U\in\mathbb{R}^{n\times m}$ *and* $\varphi\in C_0(\mathbb{R}^{n\times m})$*. In particular, if* $W_j\to 0$ *in measure, then* $\{V_j+W_j\}$ *still generates* ν*. Similarly, if* $\|V_j-W_j\|_p\to 0$ *for some* $p\in[1,\infty]$ *then both* V_j *and* W_j *generate the same Young measure.*

3.2 Decomposition

This section lets us obtain 'better' generating sequences for the Young measures we intend to be dealing with. The core result here is Proposition 3.2.3 which lets us decompose a given weakly convergent sequence into a p-equiintegrable oscillation part supported away from the boundary, and a concentration part that converges to 0 in measure. The strategy is based on Jan Kristensen's lecture notes (see [103]), however we need to adapt the proofs to the mixed smoothness setting, thus we present them in full.

In this section we will make use of the canonical projection from Lemma 2.3.1. Naturally, we mostly care about projecting the maximal derivatives that make up the **a**-gradient, and the projection could be restricted to the coordinates that correspond to those derivatives only. However, then one loses the information regarding the lower derivatives and the image of the restriction of our projection would, in a way, correspond to equivalence classes of functions in $\mathrm{W}^{\mathbf{a},p}$ up to polynomials with $\nabla_{\mathbf{a}} P \equiv 0$. However, for what we do here, it is significantly more convenient to work with the full information on a given function rather than just its **a**-gradient $\nabla_{\mathbf{a}} u$. Thus, for the purpose of this section, we denote by $\widetilde{\nabla_{\mathbf{a}}} u$ the full gradient of u given by $\widetilde{\nabla_{\mathbf{a}}} u := (\partial^\alpha u)_{\langle \alpha, \mathbf{a}^{-1} \rangle \leqslant 1}$ and by $\widehat{\nabla_{\mathbf{a}}} u$ its lower gradient $\widehat{\nabla_{\mathbf{a}}} u := (\partial^\alpha u)_{\langle \alpha, \mathbf{a}^{-1} \rangle < 1}$, so that $\widetilde{\nabla_{\mathbf{a}}} u = \widehat{\nabla_{\mathbf{a}}} u \oplus \nabla_{\mathbf{a}} u$.

Definition 3.2.1. *For $p \in (1, \infty)$ we say that a family of functions $\{V_k\} \subset \mathrm{L}^p(\Omega, \mathbb{R}^d)$ is p-equiintegrable if the family $\{|V_k|^p\} \subset \mathrm{L}^1(\Omega; \mathbb{R})$ is equiintegrable in the usual sense.*

The following lemma shows how to cut-off a given weakly convergent sequence so that it becomes compactly supported in Ω. This is important for our canonical projection, as functions in $\mathrm{W}^{\mathbf{a},p}_0(\Omega)$ may always be extended by 0 to $\mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N)$, which allows us to work with domains that need not satisfy the strong **a**-horn condition, i.e., ones that need not be $\mathrm{W}^{\mathbf{a},p}$ extension domains.

Lemma 3.2.2. *Let $\Omega \subset \mathbb{R}^N$ be a bounded open set satisfying the weak **a**-horn condition. Then, for any sequence $u_j \rightharpoonup 0$ in $\mathrm{W}^{\mathbf{a},p}(\Omega)$, there exists a sequence $v_j \in C_c^\infty(\Omega)$ such that the sequence $(\widetilde{\nabla_{\mathbf{a}}} u_j - \widetilde{\nabla_{\mathbf{a}}} v_j)$ converges to 0 in measure. In particular, if $\nabla_{\mathbf{a}} u_j$ (or, equivalently, $\widetilde{\nabla_{\mathbf{a}}} u_j$) generates some oscillation Young measure ν then $\nabla_{\mathbf{a}} v_j$ (respectively $\widetilde{\nabla_{\mathbf{a}}} v_j$) also generates ν. Furthermore, if $\{\widetilde{\nabla_{\mathbf{a}}} u_j\}$ is p-equiintegrable then so is $\{\widetilde{\nabla_{\mathbf{a}}} v_j\}$.*

Proof. Take a family of smooth, open sub-domains $\Omega_k \Subset \Omega_{k+1} \Subset \Omega$ (here $\Subset$ denotes compact inclusion) with $\bigcup_{k=1}^\infty \Omega_k = \Omega$. Fix a family of cut-off functions $\varphi_k \in C_c^\infty(\Omega; [0,1])$ with $\varphi_k \equiv 1$ on Ω_k and denote $M_k := ||\widetilde{\nabla_{\mathbf{a}}} \varphi_k||_{\mathrm{L}^\infty(\Omega)} < \infty$.

For any k, j we may write $u_j = \varphi_k u_j + (1 - \varphi_k) u_j$, and the goal is to show that the second term is small. We have

$$\int_\Omega |\widetilde{\nabla_{\mathbf{a}}}((1 - \varphi_k) u_j)| \, \mathrm{d}x = \int_{\Omega \setminus \Omega_k} |\widetilde{\nabla_{\mathbf{a}}}((1 - \varphi_k) u_j)| \, \mathrm{d}x,$$

because on Ω_k the integrand is identically equal to 0. Differentiating the product we distinguish between the case where all the derivatives fall on u_j and the one where we also differentiate $(1-\varphi_k)$ and this yields

$$\int_\Omega |\widetilde{\nabla_{\mathbf{a}}}((1-\varphi_k)u_j)|\,\mathrm{d}x \leqslant \int_{\Omega\setminus\Omega_k} (1-\varphi_k)|\widetilde{\nabla_{\mathbf{a}}}u_j|\,\mathrm{d}x + \int_{\Omega\setminus\Omega_k} M_k|\widehat{\nabla_{\mathbf{a}}}u_j|\,\mathrm{d}x.$$

Since the family $\{|\widetilde{\nabla_{\mathbf{a}}}u_j|\}$ is bounded in L^p, it is uniformly integrable in L^1. Adding the fact that $|1-\varphi_k| \leqslant 1$ and $|\Omega \setminus \Omega_k| \to 0$ we deduce that the first term in our inequality converges to 0 with $k \to \infty$, uniformly in j. On the other hand, Lemma 2.2.7 shows that the lower gradients $\widehat{\nabla_{\mathbf{a}}}u_j$ converge to 0 strongly in L^p, thus in particular in L^1. Therefore, there exists a sequence $k_j \to \infty$ such that $\int_{\Omega\setminus\Omega_{k_j}} M_{k_j}|\widehat{\nabla_{\mathbf{a}}}u_j|\,\mathrm{d}x \to 0$ with $j \to \infty$. Combining the two we see that $\widetilde{\nabla_{\mathbf{a}}}((1-\varphi_{k_j})u_j \to 0$ strongly in L^1.

Repeating the above reasoning with L^1 norm replaced by L^p we get

$$\int_\Omega |\widetilde{\nabla_{\mathbf{a}}}((1-\varphi_k)u_j)|^p\,\mathrm{d}x \leqslant C\int_{\Omega\setminus\Omega_k} (1-\varphi_k)^p|\widetilde{\nabla_{\mathbf{a}}}u_j|^p\,\mathrm{d}x + C\int_{\Omega\setminus\Omega_k} M_k^p|\widehat{\nabla_{\mathbf{a}}}u_j|^p\,\mathrm{d}x, \quad (3.1)$$

with some absolute constant C. The first term is now bounded uniformly in k and j, whereas for the second one we may select a sequence k'_j such that it converges to 0. Hence, adjusting the first sequence k_j (i.e., slowing it down if necessary) we obtain that $(1-\varphi_{k_j})u_j$ is bounded in $\mathrm{W}^{\mathbf{a},p}$.

Combining the two we deduce that $(1-\varphi_{k_j})u_j \rightharpoonup 0$ in $\mathrm{W}^{\mathbf{a},p}$, since the sequence is bounded and the only possible limit is 0, as the full gradient converges to 0 in L^1. By construction, the sequence also converges to 0 in measure, thus setting $v_j := \varphi_{k_j}u_j$ ends the proof of the first part of the statement.

For equiintegrability it is enough to notice that if $\{\widetilde{\nabla_{\mathbf{a}}}u_j\}$ is p-equiintegrable then the first term in (3.1) converges to 0 with $k \to \infty$, uniformly in j (as $(1-\varphi_k)$ is bounded in L^∞ and $|\Omega \setminus \Omega_k| \to 0$). Thus, in this case, $\widetilde{\nabla_{\mathbf{a}}}u_j - \widetilde{\nabla_{\mathbf{a}}}v_j$ converges to 0 strongly in L^p, which proves p-equiintegrability of $\widetilde{\nabla_{\mathbf{a}}}v_j$. □

The following result is the key point of this section. It shows that (up to a subsequence) one may decompose a weakly convergent sequence into a p-equiintegrable part that carries the oscillation and a part converging to 0 in measure, which carries the concentration.

Proposition 3.2.3. *Let $\Omega \subset \mathbb{R}^N$ be a bounded open domain satisfying the weak* **a***-horn condition and let $p \in (1,\infty)$. Suppose that $u_j \rightharpoonup u$ in $\mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n)$. Then, there exists a subsequence u_{j_k} and sequences $\{g_k\} \subset C_c^\infty(\Omega;\mathbb{R}^n)$ and $\{b_k\} \subset \mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n)$, both weakly convergent to 0 in $\mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n)$, and such that the family $\widetilde{\nabla_{\mathbf{a}}}g_k$ is p-equiintegrable, $\widetilde{\nabla_{\mathbf{a}}}b_k \to 0$ in measure, and $u_{j_k} = u + g_k + b_k$.*

Proof. By considering $\{u_j - u\}$ instead of $\{u_j\}$ we may limit ourselves to the case $u \equiv 0$. Furthermore, by considering a subsequence if necessary, we may assume that $\widetilde{\nabla}_{\mathbf{a}} u_j$ generates some oscillation Young measure ν. Lemma 3.2.2 shows that we may also take $u_j = v_j + b_j^1$ with $b_j^1 \to 0$ in measure and $v_j \in \mathrm{W}_0^{\mathbf{a},p}(\Omega)$. Thus, in what follows we focus on decomposing v_j, remembering that $\widetilde{\nabla}_{\mathbf{a}} v_j$ generates the same Young measure ν.

For $l \in \mathbb{N}$ recall the standard truncation $\tau_l \colon \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times m}$ given by

$$\tau_l(W) := \begin{cases} W & \text{if } |W| \leqslant l, \\ l\frac{W}{|W|} & \text{if } |W| > l. \end{cases}$$

Since the truncation is bounded and continuous we get

$$\lim_{l\to\infty}\lim_{j\to\infty}\int_\Omega |\tau_l(\widetilde{\nabla}_{\mathbf{a}} v_j)|^p \,\mathrm{d}x = \lim_{j\to\infty}\int_\Omega\int_{\mathbb{R}^{n\times m}} |\tau_l(W)|^p \,\mathrm{d}\nu_x \,\mathrm{d}x = \int_\Omega\int_{\mathbb{R}^{n\times m}} |\cdot|^p \,\mathrm{d}\nu_x\,\mathrm{d}x,$$

where the first equality is due to Theorem 3.1.1 and the second one is an application of the Monotone Convergence Theorem. We deduce that one can extract a sequence $j_l \to \infty$ such that

$$\lim_{l\to\infty}\int_\Omega |\tau_l(\widetilde{\nabla}_{\mathbf{a}} v_{j_l})|^p \,\mathrm{d}x = \int_\Omega\int_{\mathbb{R}^{n\times m}} |\cdot|^p \,\mathrm{d}\nu_x\,\mathrm{d}x,$$

and from Theorem 3.1.1 we deduce that the family $\{\tau_l(\widetilde{\nabla}_{\mathbf{a}} v_{j_l}\}$ is p-equiintegrable. That is because L^p boundedness of $\widetilde{\nabla}_{\mathbf{a}} v_{j_l}$ implies equiintegrability in L^q for $q \in [1,p)$, hence $\tau_l(\widetilde{\nabla}_{\mathbf{a}} v_{j_l}) - \widetilde{\nabla}_{\mathbf{a}} v_{j_l} \to 0$ strongly in L^q, and thus the two generate the same oscillation Young measure ν.

In this way we have constructed a sequence $w_l := \tau_l(\widetilde{\nabla}_{\mathbf{a}} v_{j_l})$ that is p-equiintegrable and generates ν. The last thing we need to take care of is the fact that w_l is not necessarily a full gradient of a $\mathrm{W}_0^{\mathbf{a},p}$ function. To remedy that, extend w_l by 0 to the whole of $\mathbb{R}^N$ and do the same with v_{j_l}, keeping the same notation for the extensions. Since we had $v_{j_l} \in \mathrm{W}_0^{\mathbf{a},p}(\Omega)$, its extension by 0 is in the space $\mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N)$. Apply the canonical projection $P_{\mathbf{a}}$ to w_l to get the decomposition

$$w_l = \widetilde{\nabla}_{\mathbf{a}} g_l + r_l.$$

We have

$$\begin{aligned}\|r_l\|_{\mathrm{L}^q} =& \|w_l - \widetilde{\nabla}_{\mathbf{a}} g_l\|_{\mathrm{L}^q} \leqslant \|w_l - \widetilde{\nabla}_{\mathbf{a}} v_{j_l}\|_{\mathrm{L}^q} + \|\widetilde{\nabla}_{\mathbf{a}} g_l - \widetilde{\nabla}_{\mathbf{a}} v_{j_l}\|_{\mathrm{L}^q} = \\ & \|w_l - \widetilde{\nabla}_{\mathbf{a}} v_{j_l}\|_{\mathrm{L}^q} + \|P_{\mathbf{a}}(w_l - \widetilde{\nabla}_{\mathbf{a}} v_{j_l})\|_{\mathrm{L}^q} \leqslant C\|w_l - \widetilde{\nabla}_{\mathbf{a}} v_{j_l}\|_{\mathrm{L}^q} \to 0,\end{aligned}$$

where we have used the fact that $P_{\mathbf{a}}(\widetilde{\nabla_{\mathbf{a}}} v_{j_l}) = \widetilde{\nabla_{\mathbf{a}}} v_{j_l}$ and the L^q continuity of $P_{\mathbf{a}}$. In particular, this yields $r_l \to 0$ in measure. Furthermore, Lemma 2.3.1 shows that p-equiintegrability of $\{w_l\}$ yields the same for $\{\widetilde{\nabla_{\mathbf{a}}} g_l\}$.

Lastly, we restrict g_l to Ω to get $g_l \in \mathrm{W}^{\mathbf{a},p}(\Omega)$ and apply the cut-off argument from Lemma 3.2.2, to end the proof. □

The following is a simple corollary of the above and we will often use it in the subsequent parts of the present work.

Corollary 3.2.4. *Let Ω be a bounded open domain satisfying the weak* **a***-horn condition. Let ν be a $\mathrm{W}^{\mathbf{a},p}$-gradient oscillation Young measure on Ω. Then there exists a sequence $u_j \in \mathrm{W}^{\mathbf{a},p}(\Omega)$ generating ν and such that the family $\{\nabla_{\mathbf{a}} u_j\}$ is p-equiintegrable. Furthermore, if the barycentre of ν is 0 at all points of Ω, then the functions u_j may be chosen in the space $C_c^\infty(\Omega)$.*

Remark 3. It seems plausible that one could generalise Zhang's truncation lemma (see [184]) to the mixed smoothness setting, which would then allow us to extend our results to $\mathrm{W}^{\mathbf{a},\infty}$-gradient Young measures. The goal would be to show that, under the hypotheses of Corollary 3.2.4, if the Young measure ν is additionally assumed to satisfy

$$\mathrm{supp}(\nu_x) \subset K \text{ for almost every } x \in \Omega,$$

where $\mathrm{supp}(\nu_x)$ denotes the support of ν_x and K is a compact set, then the generating sequence u_j can be chosen to be bounded in $\mathrm{W}^{\mathbf{a},\infty}(\Omega)$ with

$$\sup_j \|u_j\|_{\mathrm{W}^{\mathbf{a},\infty}(\Omega)} \leqslant C(K),$$

where the constant $C(K)$ depends only on the set K. We do not have a proof yet, as we lack results on extending functions in the class $\mathrm{W}^{\mathbf{a},\infty}(\Omega')$, but we aim to tackle this problem in future work.

3.3 DiPerna-Majda Young measures

The purpose of this section is to introduce the DiPerna-Majda Young measures (often called generalised Young measures) for oscillation and concentration. We refer the reader to [54] for the original exposition, however here we follow the approach from a course given by Jan Kristensen at the University of Oxford (see [103]).

First, we introduce an appropriate test space that consists of continuous functions that admit a continuous p-recession function. Here the exponent $p \in (1, \infty)$ is fixed.

Definition 3.3.1. *For a bounded open set* $\Omega \subset \mathbb{R}^N$ *we define* $\mathrm{E}_p = \mathrm{E}_p(\Omega, \mathbb{R}^{n\times m})$ *to be the space of those continuous function* $\psi \in C(\overline{\Omega} \times \mathbb{R}^{n\times m}; \mathbb{R})$ *for which the limit* $\lim_{t\to\infty} \frac{\psi(x,tW)}{t^p}$ *exists uniformly in* $(x, W) \in \overline{\Omega \times \partial \mathrm{B}^{n\times m}}$*, where* $\mathrm{B}^{n\times m}$ *is the (Euclidean) unit ball in* $\mathbb{R}^{n\times m}$*.*

For $\psi \in \mathrm{E}_p$ *the* p*-recession integrand* $\psi_p^\infty \colon \overline{\Omega} \times \mathbb{R}^{n\times m} \to \mathbb{R}$ *is defined by*

$$\psi_p^\infty(x, W) := \lim_{t\to\infty} \frac{\varphi(x, tW)}{t^p}.$$

Remark 4. From the definition of E_p one immediately sees that for any $\psi \in \mathrm{E}_p$ the p-recession integrand ψ_p^∞ is continuous and positively p-homogeneous.

Proposition 3.3.2 (see [103])**.** *The space* E_p *equipped with the norm*

$$\|\psi\|_{\mathrm{E}_p} := \sup_{(x,W)\in\overline{\Omega}\times\mathbb{R}^{n\times m}} \frac{|\psi(x, W)|}{(1 + |W|)^p}$$

is a separable Banach space.

Definition 3.3.3. *A generalised* p*-Young measure* ν *is a triple* $((\nu_x)_{x\in\Omega}, \lambda, (\nu_x^\infty)_{x\in\overline{\Omega}})$*, where*

- $(\nu_x) \colon \Omega \to \mathcal{P}(\mathbb{R}^{n\times m})$ *is weak* Lebesgue measurable,*
- $\lambda \in \mathcal{M}^+(\overline{\Omega})$ *is a non-negative Radon measure on* $\overline{\Omega}$*,*
- $(\nu_x^\infty) \colon \overline{\Omega} \to \mathcal{P}(\partial \mathrm{B}^{n\times m})$ *is weak** λ *measurable,*
- $\int_\Omega \int_{\mathbb{R}^{n\times m}} |W|^p \,\mathrm{d}\nu_x(W)\,\mathrm{d}x < \infty.$

Observe that the maps $x \mapsto \nu_x$ and $x \mapsto \nu_x^\infty$ are only defined Lebesgue and λ almost everywhere respectively. Moreover, the exponent p only enters in the last condition. We will often say that ν is a generalised Young measure and omit the exponent p, as it will be fixed throughout the chapter.

Here (ν_x) is the classical oscillation Young measure as introduced before. We call λ the concentration measure and (ν_x^∞) the concentration-angle measure. Lastly, analogously to the Lebesgue case in the first point, we say that $(\nu_x^\infty)_x$ is weak* λ measurable if $x \mapsto \langle \nu_x^\infty, \varphi\rangle$ is λ measurable for any continuous $\varphi \colon \partial \mathrm{B}^{n\times m} \to \mathbb{R}$.

The set of generalised p-Young measures is denoted Y^p. Observe that Y^p may be seen as a subset of E_p^* if we set, for $\nu \in \mathrm{Y}^p$ and $\psi \in \mathrm{E}_p$,

$$\langle\langle \nu, \psi\rangle\rangle := \int_\Omega \langle \nu_x, \psi(x, \cdot)\rangle \,\mathrm{d}x + \int_{\overline{\Omega}} \langle \nu_x^\infty, \psi_p^\infty(x, \cdot)\rangle \,\mathrm{d}\lambda.$$

We remark that Y^p is in fact a proper subset of E_p^*, but we do not elaborate on it further and instead refer to [103] for a detailed characterisation. However, we recall the following useful result (see also [104]):

Proposition 3.3.4 (see [103])**.** *The space of generalised p-Young measures* Y^p *is a convex weakly* closed subset of* E_p^* *on which the relative sequential weak* topology is metrizable with the Kantorovich* L^p *metric when restricted to* Y^p*.*

We recall that the Kantorovich L^p distance between two probability measures μ and ν with finite p-th moments is defined as the infimum $\inf \mathbb{E}|X - Y|^p$, where $|\cdot|$ denotes the standard Euclidean norm, $\mathbb{E}$ is the expectation with respect to the joint distribution of (X, Y), and the infimum is taken over all pairs of random variables X and Y with distributions μ and ν respectively. We do not elaborate on this point further, as all we will use in what follows is that the topology is metrisable — the particular choice of a metric is not important. Instead, we refer the reader to Santambrogio's book [150] for more details on the Kantorovich distance and related issues.

Definition 3.3.5. *We say that a sequence* $V_j \in \mathrm{L}^p(\Omega; \mathbb{R}^{n\times m})$ *generates the generalised Young measure* ν *if for every* $\psi \in \mathrm{E}_p$ *we have*

$$\lim_{j\to\infty} \int_\Omega \psi(x, V_j(x))\,\mathrm{d}x = \langle\langle \nu, \psi \rangle\rangle.$$

Observe that, if for $V \in \mathrm{L}^p(\Omega; \mathbb{R}^{n\times m})$ we set $\varepsilon_V := ((\delta_{V(x)}), 0, \mathrm{n/a})$ to be the corresponding elementary DiPerna-Majda Young measure, then a sequence V_j generates ν if and only if ε_{V_j} converges weakly* to ν in E_p^*. Moreover, taking $g(x, W) := 1 + |W|^p \in \mathrm{E}_p$ we easily see that $\|\varepsilon_V\|_{\mathrm{E}_p^*} = 1 + \|V\|_{\mathrm{L}^p}$. This, together with Proposition 3.3.4 implies the following:

Lemma 3.3.6. *Suppose that* $V_j \rightharpoonup V$ *in* $\mathrm{L}^p(\Omega; \mathbb{R}^{n\times m})$*. Then there exists a subsequence* V_{j_k} *that generates some generalised Young measure* ν*.*

Proof. Since Proposition 3.3.4 tells us that Y^p is weak* closed with metrizable topology all that we need is to show that, under our assumptions, the sequence ε_{V_j} is bounded in E^*. This, however, results immediately from the fact that the L^p norm of V_j is bounded. □

Definition 3.3.7. *If* $\nu \in \mathrm{Y}^p$ *is generated by a sequence* $\nabla_{\mathbf{a}} u_j$ *for some* $u_j \rightharpoonup u$ *in* $\mathrm{W}^{\mathbf{a},p}$ *then we say that* ν *is a generalised* $\mathrm{W}^{\mathbf{a},p}$*-gradient Young measure.*

3.4 Localisation

Definition 3.4.1. *We say that τ is a tangent measure to μ at a point x_0, denoted by $\tau \in Tan(\mu, x_0)$, if there exists a sequence of $r_j \searrow 0$ such that, with μ^{x_0,r_j} defined by*

$$\mu^{x_0,r_j}(A) := \frac{\mu(x_0 + r_j \odot A)}{\mu(Q_{r_j}(x_0))}$$

for $A \in \mathcal{B}(Q)$, one has $\mu^{x_0,r_j} \overset{*}{\rightharpoonup} \tau$ in $C(\overline{Q})^*$.

Observe that the above definition differs from the usual one (where one uses Euclidean balls or cubes) — as all the previous arguments, this had to be adapted to our anisotropic scaling.

The following is, by now, a classical result in the theory of Young measures under differential constraints. We will be particularly interested in the part of the result that isolates the information for oscillation Young measures (see Corollary 3.4.3), but we include and prove the full result for future use. In our proof we follow the approach taken by Kristensen in his lecture notes (see [103]) and we adapt it to the mixed smoothness setting.

Proposition 3.4.2. *Let $\Omega \subset \mathbb{R}^N$ be an open and bounded domain. Fix $1 < p < \infty$ and let $\nu = ((\nu_x), \lambda, (\nu_x^\infty))$ be a generalised $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure on Ω. Write $\lambda = \frac{\mathrm{d}\lambda}{\mathrm{d}x}\mathcal{L}^n\big|_\Omega + \lambda^s$ for the Radon-Nikodym decomposition of λ with respect to the Lebesgue measure on Ω. Let Q denote the unit cube. Then*

1. *for $\mathcal{L}^n$-a.e. $x_0 \in \Omega$ the triple $((\nu_{x_0})_{y\in Q}, \frac{\mathrm{d}\lambda}{\mathrm{d}x}(x_0)\mathcal{L}^n\big|Q, (\nu_{x_0})_{y\in Q})$ is a homogeneous generalised $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure. Its barycentre is $\overline{\nu_{x_0}}\mathbb{1}_Q$,*
2. *for λ^s-a.e. $x_0 \in \Omega$ and each $\mu \in Tan(\lambda^s, x_0)$ the triple $((\delta_0)_{y\in Q}, \mu, (\nu_{x_0}^\infty)_{y\in\overline{Q}})$ is a generalised $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure. Its barycentre is 0.*

Proof. Let the generalised $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure ν be as in the statement. Let $\nabla_{\mathbf{a}} u_j$ be its generating sequence with $u_j \rightharpoonup u$ in $\mathrm{W}^{\mathbf{a},p}(\Omega)$. Fix $x_0 \in \Omega$ and an $r > 0$ such that $Q_r(x_0) \subset \Omega$ with $\lambda(\partial Q_r(x_0)) = 0$, which must hold for (Lebesgue) almost all small R, as λ is finite. Fix $\varphi \in C(\overline{Q})$ and $\psi \in C(\mathbb{R}^{n\times m})$ such that $\varphi \otimes \psi \in \mathrm{E}_p(Q)$. Define $\tau^{x_0,r} \colon \overline{Q}_r(x_0) \to \overline{Q}$ by $\tau^{x_0,r}(x) := (1/r) \odot (x - x_0)$. Then $(\varphi \circ \tau^{x_0,r}) \otimes \psi \in \mathrm{E}_p(Q_r(x_0), \mathbb{R}^{n\times m})$. Thus, by definition,

$$\lim_{j\to\infty} \int_{Q_r(x_0)} [\varphi \circ \tau^{x_0,r}](x)\psi(\nabla_{\mathbf{a}} u_j)\,\mathrm{d}x = \int_{Q_r(x_0)} \varphi \circ \tau^{x_0,r}\langle \nu_x, \psi\rangle\,\mathrm{d}x + \int_{Q_r(x_0)} \varphi \circ \tau^{x_0,r}\langle \nu_x^\infty, \psi_p^\infty\rangle\,\mathrm{d}\lambda.$$

Put

$$u_j^{x_0,r}(y) := r^{-1}u_j(x_0 + r \odot y) - P_j^{x_0,r}, \tag{3.2}$$

where $P_j^{x_0,r}$ is the correction polynomial for $r^{-1}u_j(x_0 + r \odot y)$ given by Proposition 2.5.2. The polynomial $P_j^{x_0,r}$ satisfies $\nabla_{\mathbf{a}} P_j^{x_0,r} = 0$, we only put it there so that $\|u_j^{x_0,r}\|_{\mathrm{W}^{\mathbf{a},p}(Q)} \leqslant C\|\nabla_{\mathbf{a}} u_j^{x_0,r}(y)\|_{\mathrm{L}^p(Q)}$ with an absolute constant C, which will ensure $\mathrm{W}^{\mathbf{a},p}$ boundedness of our sequence. Coming back to the previous equation, let us change variables in the integral letting $y := \tau^{x_0,r}(x)$. This yields

$$\begin{aligned}\lim_{j\to\infty} \int_Q \varphi\psi(\nabla_{\mathbf{a}} u_j^{x_0,r}) r^{|\mathbf{a}^{-1}|}\,\mathrm{d}y &= \int_Q \varphi\langle \nu_{x_0+r\odot y}, \psi\rangle r^{|\mathbf{a}^{-1}|}\,\mathrm{d}y \\ &\quad + \int_Q \varphi\langle \nu^\infty_{x_0+r\odot y}, \psi_p^\infty\rangle\,\mathrm{d}(\tau^{x_0,r}{}_{\#}\lambda).\end{aligned}$$

Here $\tau^{x_0,r}{}_{\#}\lambda$ is the push-forward of λ by $\tau^{x_0,r}$ defined through

$$\tau^{x_0,r}{}_{\#}\lambda(A) := \lambda((\tau^{x_0,r})^{-1}(A)),$$

where $(\tau^{x_0,r})^{-1}(A)$ is the preimage of A under $\tau^{x_0,r}$. Denote

$$\begin{aligned} f(x) &:= \langle \nu_x, \psi\rangle, \\ g^a(x) &:= \langle \nu_x^\infty, \psi_p^\infty\rangle \frac{\mathrm{d}\lambda}{\mathrm{d}x}(x), \\ g^s(x) &:= \langle \nu_x^\infty, \psi_p^\infty\rangle. \end{aligned}$$

We have $f, g^a \in \mathrm{L}^1(\Omega, \mathcal{L}|\Omega)$ and $g^s \in \mathrm{L}^1(\Omega, \lambda^s)$, so the set of (box) Lebesgue points of $f + g^a$ has full Lebesgue measure in Ω. Similarly, the set of box λ^s-continuity points (analogous to Lebesgue points but with the Lebesgue measure replaced by λ^s) has full λ^s measure in Ω, thus we have the following:
For Lebesgue almost every $x_0 \in \Omega$:

$$\begin{aligned} 0 &= \lim_{r\to 0^+} \fint_{Q_r(x_0)} |f + g^a - (f+g^a)(x_0)|\,\mathrm{d}x \\ &= \lim_{r\to 0^+} \fint_Q |f(x_0 + r\odot y) + g^a(x_0 + r\odot y) - f(x_0) - g^a(x_0)|\,\mathrm{d}y. \end{aligned} \tag{3.3}$$

For λ^s-a.e. $x_0 \in \Omega$:

$$0 = \lim_{r\to 0^+} \fint_{Q_r(x_0)} |g^s - g^s(x_0)|\,\mathrm{d}\lambda^s = \lim_{r\to 0^+} \fint_Q |g^s(x_0 + r\odot y) - g^s(x_0)|\,\mathrm{d}(\lambda^s)^{x_0,r}. \tag{3.4}$$

Since $\mathcal{L}^N \perp \lambda^s$ we have, for $\mathcal{L}^N$-a.e. $x_0 \in \Omega$,

$$\lim_{r\to 0^+} \frac{1}{r^{|\mathbf{a}^{-1}|}} \lambda^s(Q_r(x_0)) = 0. \tag{3.5}$$

Similarly, for λ^s-a.e. $x_0 \in \Omega$ we have

$$\lim_{r\to 0^+} \frac{1}{\lambda^s(Q_r(x_0))} \int_{Q_r(x_0)} (f + g^a + 1)\,\mathrm{d}x = 0. \tag{3.6}$$

Decomposing $\lambda = \frac{\mathrm{d}\lambda}{\mathrm{d}x}\,\mathrm{d}x + \lambda^s$ and using (3.3) and (3.5) we find a Lebesgue-null set $N(\varphi,\psi) \subset \Omega$ such that for all $x_0 \in \Omega \setminus N(\varphi,\psi)$ we have

$$\lim_{r\to 0^+}\lim_{j\to\infty} \int_Q \varphi\psi(\nabla_{\mathbf{a}} u_j^{x_0,r})\,\mathrm{d}y = \int_Q \varphi\,\mathrm{d}y \left(\langle \nu_{x_0}, \psi\rangle + \langle \nu_{x_0}^\infty, \psi_p^\infty\rangle \frac{\mathrm{d}\lambda}{\mathrm{d}x}(x_0)\right). \tag{3.7}$$

Let us fix two countable families $\mathcal{D}_1 = \{\varphi \in \mathcal{D}_1\}$ and $\mathcal{D}_2 = \{\psi \in \mathcal{D}_2\}$ that are dense in $C(\overline{Q})$ and $C(\mathbb{R}^{n\times m})$ respectively and such that their tensor products $\varphi\otimes\psi$ belong to E_p and are dense in that space — this is indeed possible, see for example Lemma 4.7 in [148]. Following the argument above, we may construct a Lebesgue-null set $N := \bigcup_{\mathcal{D}_1,\mathcal{D}_2} N(\varphi,\psi)$ such that for all $x_0 \in \Omega\setminus N$ the convergence in (3.7) holds for all $\varphi\otimes\psi \in \mathcal{D}_1\otimes\mathcal{D}_2$. Using a standard diagonal extraction argument we may now, for any fixed $x_0 \in \Omega\setminus N$, obtain sequences $j_k \to \infty$ and $r_k \to 0^+$ as $k\to\infty$ such that

$$\lim_{k\to\infty} \int_Q \varphi\psi(\nabla_{\mathbf{a}} u_{j_k}^{x_0,r_k})\,\mathrm{d}y = \int_Q \varphi\,\mathrm{d}y \left(\langle \nu_{x_0}, \psi\rangle + \langle \nu_{x_0}^\infty, \psi_p^\infty\rangle \frac{\mathrm{d}\lambda}{\mathrm{d}x}(x_0)\right),$$

for all $\varphi\otimes\psi \in \mathcal{D}_1\otimes\mathcal{D}_2$. Since $\mathcal{D}_1\otimes\mathcal{D}_2$ is dense in E_p, this shows that the sequence $\nabla_{\mathbf{a}} u_{j_k}^{x_0,r_k}$ generates the homogeneous Young measure $((\nu_{x_0}), \frac{\mathrm{d}\lambda}{\mathrm{d}x}(x_0), \nu_{x_0}^\infty)$ on Q. To prove that this measure is indeed a homogeneous $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure we note that, since $(x,z)\mapsto |z|^p \in \mathrm{E}_p$, we know that the sequence $\|\nabla_{\mathbf{a}} u_{j_k}^{x_0,r_k}\|_{\mathrm{L}^p(Q)}$ converges, so it is bounded. Since we have introduced the polynomial correction terms to $u_{j_k}^{x_0,r_k}$ (see equation (3.2)) we have $\|u_{j_k}^{x_0,r_k}\|_{\mathrm{W}^{\mathbf{a},p}(Q)} \leqslant C\|\nabla_{\mathbf{a}} u_{j_k}^{x_0,r_k}\|_{\mathrm{L}^p(Q)}$, thus the sequence is bounded in $\mathrm{W}^{\mathbf{a},p}(Q)$, and so our measure is indeed a homogeneous $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure.

Now, for the singular part. Observe that

$$\begin{aligned}\lim_{j\to\infty} \int_Q \varphi\psi(\nabla_{\mathbf{a}} u_j^{x_0,r}) \frac{r^{|\mathbf{a}^{-1}|}}{\lambda^s(Q_r(x_0))}\,\mathrm{d}y = &\\ \frac{r^{|\mathbf{a}^{-1}|}}{\lambda^s(Q_r(x_0))} &\left(\int_Q \varphi(y)(f(x_0 + r\odot y) + g^a(x_0 + r\odot y))\,\mathrm{d}y\right) + \\ &\int_Q \varphi(y) g^s(x_0 + r\odot y)\,\mathrm{d}(\lambda^s)^{x_0,r}.\end{aligned}$$

Denote $\theta(r) := (\lambda^s(Q_r(x_0))r^{-|\mathbf{a}^{-1}|})^{1/p}$ and define

$$u_j^{x_0,r_k}(y) := \theta(r)^{-1} r^{-1} u_j(x_0 + r\odot y) - P_j^{x_0,r_k},$$

where again, the polynomial $P_j^{x_0,r_k}$ is the one given by Proposition 2.5.2 applied to $\theta(r)^{-1}r^{-1}u_j(x_0+r\odot y)$. Then, similarly to before,

$$\lim_{j\to\infty}\int_Q \varphi\psi(\nabla_{\mathbf{a}}u_j^{x_0,r_k})\,\mathrm{d}y = \int_Q \varphi(y)\left(\langle \nu_{x_0+r\odot y},\psi(\theta(r)^{-1}(\cdot))\rangle + \theta(r)^{-p}\langle \nu^\infty_{x_0+r\odot y},\psi_p^\infty\rangle \frac{\mathrm{d}\lambda}{\mathrm{d}x}(x_0+r\odot y)\right)\mathrm{d}y + \int_Q \varphi(y)g^s(x_0+r\odot y)\,\mathrm{d}(\lambda^s)^{x_0,r}. \tag{3.8}$$

As before we find a λ^s-null set $N^s := \bigcup_{\mathcal{D}_1,\mathcal{D}_2} N^s(\varphi,\psi)\subset\Omega$ such that (3.4) and (3.6) hold for all $x_0\in\Omega\setminus N^s$ and for all $\varphi\in\mathcal{D}_1$, $\psi\in\mathcal{D}_2$. Fix such a point x_0 and $\mu\in\mathrm{Tan}(\lambda^s,x_0)$. Pick a sequence $r_k\searrow 0$ for which $(\lambda^s)^{x_0,r_k}\overset{*}{\rightharpoonup}\mu$ in $C(\overline{Q})^*$. We have

$$\int_Q \varphi(y)g^s(x_0+r\odot y)\,\mathrm{d}(\lambda^s)^{x_0,r_k} = \int_Q \varphi\,\mathrm{d}(\lambda^s)^{x_0,r_k}g^s(x_0) + \int_Q \varphi(y)\left(g^s(x_0+r\odot y)-g^s(x_0)\right)\mathrm{d}(\lambda^s)^{x_0,r_k},$$

and from (3.8) we deduce that

$$\lim_{k\to\infty}\lim_{j\to\infty}\int_Q \varphi\psi(\nabla_{\mathbf{a}}u_j^{x_0,r_k})\,\mathrm{d}y = \int_Q\varphi\,\mathrm{d}y\psi(0) + \int_Q \varphi\,\mathrm{d}\mu\langle\nu^\infty_{x_0},\psi_p^\infty\rangle,$$

for all $\varphi\in\mathcal{D}_1$, $\psi\in\mathcal{D}_2$. Thus, another diagonal extraction argument paired up with the fact that $\mathcal{D}_1\otimes\mathcal{D}_2$ is dense in E_p ends the proof. □

The following is a simple corollary of our last Proposition. It isolates the information we have for oscillation $\mathrm{W}^{\mathbf{a},p}$-gradient Young measures.

Corollary 3.4.3. *Let $\Omega\subset\mathbb{R}^N$ be an open and bounded domain. Fix $1<p<\infty$ and let $\nu=\{\nu_x\}$ be an oscillation $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure on Ω. Then, for almost every $x\in\Omega$, the measure ν_x is a homogeneous oscillation $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure.*

3.5 a-quasiconvexity

This section aims to develop the mixed smoothness equivalent of quasiconvexity introduced first by Morrey in [129]. This notion, in the classical setting of first order gradients, has been extensively studied in the literature, some of which we have recalled in the introduction. Without attempting to give a comprehensive list of references we instead direct the reader to the book by Dacorogna (see [40]) and the bibliography therein. We will revisit **a**-quasiconvexity in more detail in the next chapter, when we study lower semicontinuity properties of functionals. For now, we content ourselves with presenting only a few background results that we will need to establish the dual characterisation of $W^{\mathbf{a},p}$-gradient Young measures.

The central notion here is the following adaptation to the mixed smoothness framework of the classical notion of quasiconvexity (recall that m is the cardinality of the set $\{\alpha \in \mathbb{Z}_+^N : \langle \alpha, \mathbf{a}^{-1}\rangle = 1\}$, so that $\mathbb{R}^{n\times m}$ is the space where $\nabla_{\mathbf{a}} u$ takes its values):

Definition 3.5.1. *We say that a function $f\colon \mathbb{R}^{n\times m} \to [-\infty, \infty)$ is* **a**-*quasiconvex if for every $V \in \mathbb{R}^{n\times m}$ one has*

$$f(V) \leqslant \inf_{u\in C_c^\infty(Q;\mathbb{R}^n)} \fint_Q f(V + \nabla_{\mathbf{a}} u(x))\,\mathrm{d}x.$$

Let us note that examples of **a**-quasiconvex functions may be readily constructed using, as building blocks, functions that are quasiconvex in the usual (isotropic) sense. First of all note that any convex function is **a**-quasiconvex. As a non-convex and anisotropic example one may consider, for instance, $\mathbf{a} := (2, 1, 1)$ with f acting on **a**-gradients of functions $u\colon [-1, 1]^3 \to \mathbb{R}^2$ as

$$f(\partial_t^2 u, \nabla_x u) := |\partial_t^2 u|^2 + \det(\nabla_x u), \tag{3.9}$$

i.e.,

$$f(V_t, V_x) := |V_t|^2 + \det(V_x),$$

for $V_t \in \mathbb{R}^{2\times 1}$ and $V_x \in \mathbb{R}^{2\times 2}$ — in our example $\partial_t^2 u(t, x)$ is a two-dimensional vector, and $\nabla_x u(t, x)$ denotes the first derivative of u with respect to the x variable, thus is a two-by-two matrix. The fact that f is **a**-quasiconvex results immediately from convexity of $|\cdot|^2$ and quasiconvexity of $\det(\cdot)$. For any function $u \in C_c^\infty(Q; \mathbb{R}^2$ with $Q = [-1, 1]^3$ and $V = (V_t, V_x) \in \mathbb{R}^{2\times 3}$ we may write

$$\fint_Q f(V + \nabla_{\mathbf{a}} u(t, x))\,\mathrm{d}(t, x) = \fint_Q |V_t + \partial_t^2 u(t, x)|^2 + \det(V_x + \nabla_x u(t, x))\,\mathrm{d}t\,\mathrm{d}x.$$

Since $u \in C_c^\infty(Q; \mathbb{R}^2)$ we have, for any fixed x, $u(\cdot, x) \in C_c^\infty([-1, 1]; \mathbb{R}^2)$ and similarly, for any fixed t, $u(t, \cdot) \in C_c^\infty([-1, 1]^2; \mathbb{R}^2)$. Thus

$$\fint_{[-1,1]} |V_t + \partial_t^2 u(t, x)|^2\,\mathrm{d}t \geqslant |V_t|^2 \text{ for every } x,$$

and

$$\fint_{[-1,1]^2} \det(V_x + \nabla_x u(t, x))\,\mathrm{d}x \geqslant \det(V_x) \text{ for every } t,$$

by convexity of $|\cdot|^2$ and quasiconvexity of det, respectively. Putting the two together (and using Fubini's theorem) we immediately deduce that f is indeed **a**-quasiconvex.

For functions that are not **a**-quasiconvex we introduce their quasiconvexifications in the following natural way:

Definition 3.5.2. *For a measurable function $g\colon \mathbb{R}^{n\times m} \to [-\infty, \infty)$ we define the function $\mathcal{Q}g\colon \mathbb{R}^{n\times m} \to [-\infty; \infty)$ by*

$$\mathcal{Q}g(V) := \inf \left\{ \fint_Q g(V + \nabla_{\mathbf{a}} u(x))\, \mathrm{d}x \colon u \in C_c^\infty(Q) \right\}. \tag{3.10}$$

Remark 5. The expression on the right-hand side of the above definition is often called Dacorogna's formula in the standard first order gradient case (see [40]). For us, its principal use is to obtain (in the next section) a theoretical characterisation of the oscillation Young measures generated by **a**-gradients in the spirit of Kinderlehrer and Pedregal, who have done the same for first order gradients in [95].

In what follows we will often refer to $\mathcal{Q}g$ as the **a**-quasiconvex envelope of g, and the next lemma justifies this terminology, by showing that $\mathcal{Q}g$ is the largest **a**-quasiconvex function that is smaller than or equal to g. By definition $\mathcal{Q}g \leqslant g$, simply by testing the definition with $u \equiv 0$ and, again from the definition, one sees immediately that any **a**-quasiconvex function that is no bigger than g must be no bigger than $\mathcal{Q}g$, thus it only remains to show that $\mathcal{Q}g$ is **a**-quasiconvex. The proof that we will present here follows the approach of Fonseca and Müller from [67], which was developed in the $\mathcal{A}$-free framework.

Lemma 3.5.3. *For a continuous function $g\colon \mathbb{R}^{n\times m} \to \mathbb{R}$ we have*

$$\mathcal{Q}(\mathcal{Q}g) = \mathcal{Q}g,$$

that is, $\mathcal{Q}g$ is indeed **a**-*quasiconvex.*

Proof. For $R > 0$ we let

$$\mathcal{Q}^R g(V) := \inf \left\{ \fint_Q g(V + \nabla_{\mathbf{a}} w(x))\, \mathrm{d}x \colon w \in C_c^\infty(Q), \|\nabla_{\mathbf{a}} w\|_{\mathrm{L}^\infty} \leqslant R \right\}.$$

First we show that $\mathcal{Q}^R g$ is continuous. Fix any $\rho > 0$ and let $\omega(\cdot)$ be the modulus of continuity of g on $B(0, \rho + R)$, i.e.,

$$\omega(r) := \sup\{|g(V) - g(V')| \,:\, V, V' \in \overline{B}(0, \rho + R), |V - V'| \leqslant r\}.$$

Since g is uniformly continuous on $\overline{B}(0, \rho + R)$ we know that ω is finite for all r and $\lim_{r\to 0^+} \omega(r) = 0$.

Now, for all $V, V' \in B(0, \rho)$ and every admissible w, i.e., $w \in C_c^\infty(Q)$ with $\|\nabla_{\mathbf{a}} w\|_{\mathrm{L}^\infty} \leqslant R$ we have

$$\begin{aligned} \fint_Q g(V + \nabla_{\mathbf{a}} w(x))\,\mathrm{d}x &\geqslant \fint_Q g(V' + \nabla_{\mathbf{a}} w(x))\,\mathrm{d}x - \omega(|V - V'|) \\ &\geqslant \mathcal{Q}^R g(V') - \omega(|V - V'|), \end{aligned}$$

by definition of $\mathcal{Q}^R g$. Taking infimum over admissible w yields

$$\mathcal{Q}^R g(V) - \mathcal{Q}^R g(V') \geqslant \omega(|V - V'|).$$

Inverting the roles of V and V' yields a symmetric inequality, so that

$$\left|\mathcal{Q}^R g(V) - \mathcal{Q}^R g(V')\right| \leqslant \omega(|V - V'|),$$

which in turn means that $\mathcal{Q}^R g$ is uniformly continuous in $B(0, \rho)$.

We wish to show that for any $w \in C_c^\infty(Q)$ and $V \in \mathbb{R}^{n\times m}$ we have

$$\fint_Q (\mathcal{Q}^R g(V + \nabla_{\mathbf{a}} w(x))\,\mathrm{d}x \geqslant \mathcal{Q} g(V).$$

To this end fix an arbitrary $w \in C_c^\infty(Q)$ and $V \in \mathbb{R}^{n\times m}$. For a fixed $j \in \mathbb{N}$ decompose Q into I (a finite constant depending on j) boxes of the form

$$Q = \bigcup_{i\in I} Q_i^j = \bigcup_{i\in I} \left(\frac{1}{j} \odot Q + a_{j,i}\right).$$

Note that an exact decomposition of this form is possible only for certain natural numbers j, but we can simply restrict our attention to j's of the form $j = k^{a_1\cdot\ldots\cdot a_N}$ for $k \in \mathbb{N}$, for which the decomposition is exact. Define a piecewise constant approximation of $\nabla_{\mathbf{a}} w$ by

$$W^j := \sum_i \mathbb{1}_{Q_i^j} W_i^j,$$

where $W_i^j := \fint_{Q_i^j} \nabla_{\mathbf{a}} w(x)\,\mathrm{d}x$. Since $\nabla_{\mathbf{a}} w$ is continuous we have $W^j \to \nabla_{\mathbf{a}} w$ in L^∞ as $j \to \infty$. Fix an $\varepsilon > 0$. Continuity of $\mathcal{Q}^R g$ yields, for j large enough, that

$$\int_Q \mathcal{Q}^R g(V + \nabla_{\mathbf{a}} w(x))\,\mathrm{d}x \geqslant \int_Q \mathcal{Q}^R g(V + W^j(x))\,\mathrm{d}x - \varepsilon = \sum_i |Q_i^j| \mathcal{Q}^R g(V + W_i^j) - \varepsilon,$$

because $\mathcal{Q}^R g$ is uniformly continuous on compact sets and the L^∞ norms of W^j are bounded by $\|\nabla_{\mathbf{a}} w\|_{\mathrm{L}^\infty} < \infty$.

By definition of $\mathcal{Q}^R g$ there exist $u_i^j \in C_c^\infty(Q)$ with $\|\nabla_{\mathbf{a}} u_i^j\|_{\mathrm{L}^\infty} \leqslant R$ such that

$$\mathcal{Q}^R g(V + W_i^j) \geqslant \fint_Q g(V + W_i^j + \nabla_{\mathbf{a}} u_i^j(x))\,\mathrm{d}x - \varepsilon.$$

Set

$$U^J := \sum_i \mathbb{1}_{Q_i^j} W_i^j + \nabla_{\mathbf{a}} \left[j^{-1} u_i^j (j \odot (y - a_i^j)) \right].$$

With this scaling we may write

$$\begin{aligned}\fint_Q g(V + U^j)\,\mathrm{d}x = &\sum_i \int_{Q_i^j} g(V + W_i^j + \nabla_{\mathbf{a}} \left[j^{-1} u_i^j (j \odot (y - a_i^j)) \right])\,\mathrm{d}y = \\ &\sum_i |Q_i^j| \fint_Q g(V + W_i^j + \nabla_{\mathbf{a}} u_i^j(y))\,\mathrm{d}y \leqslant \\ &\sum_i |Q_i^j| \mathcal{Q}^R g(V + W_i^j) + \varepsilon \leqslant \fint_Q \mathcal{Q}^R g(V + \nabla_{\mathbf{a}} w(x))\,\mathrm{d}x + 2\varepsilon.\end{aligned}$$

Finally, observe that if we let

$$\widetilde{U^j} := w + \sum_i j^{-1} u_i^j (j \odot (y - a_i^j))$$

then $\widetilde{U^j} \in C_c^\infty(Q)$ with $\|\nabla_{\mathbf{a}} \widetilde{U^j}\|_{\mathrm{L}^\infty} \leqslant \|\nabla_{\mathbf{a}} w\|_{\mathrm{L}^\infty} + R$ for all j. Since U^j also satisfies $\|U^j\|_{\mathrm{L}^\infty} \leqslant \|\nabla_{\mathbf{a}} w\|_{\mathrm{L}^\infty} + R$ and

$$\lim_{j \to \infty} \|U^j - \nabla_{\mathbf{a}} \widetilde{U^j}\|_{\mathrm{L}^\infty} = 0$$

we have, by uniform continuity of g on compacts, that for j large enough

$$\fint_Q g(V + U^j(x)) \geqslant \fint_Q g(V + \nabla_{\mathbf{a}} \widetilde{U^j}(x))\,\mathrm{d}x - \varepsilon.$$

Thus, putting all of this together yields

$$\mathcal{Q} g(V) \leqslant \fint_Q g(V + \nabla_{\mathbf{a}} \widetilde{U^j}(x))\,\mathrm{d}x \leqslant \fint_Q \mathcal{Q}^R g(V + \nabla_{\mathbf{a}} w(x))\,\mathrm{d}x + 3\varepsilon.$$

Since ε was arbitrary we deduce that for any $w \in C_c^\infty$ we have

$$\mathcal{Q} g(V) \leqslant \fint_Q \mathcal{Q}^R g(V + \nabla_{\mathbf{a}} w(x))\,\mathrm{d}x$$

Finally, since $\mathcal{Q}^R g \searrow \mathcal{Q} g$ pointwise with $R \to \infty$ we deduce, by the Monotone Convergence Theorem, that

$$\mathcal{Q} g(V) \leqslant \inf_{w \in C_c^\infty} \fint_Q \mathcal{Q} g(V + \nabla_{\mathbf{a}} w(x))\,\mathrm{d}x = \mathcal{Q}(\mathcal{Q} g)(V).$$

Since the inverse inequality $\mathcal{Q} g \geqslant \mathcal{Q}(\mathcal{Q} g)$ is obvious we have $\mathcal{Q} g = \mathcal{Q}(\mathcal{Q} g)$ and the proof is complete. $\square$

3.6 Topological structure of the space of oscillation $\mathrm{W}^{\mathbf{a},p}$-gradient Young measures

Our next aim is to obtain a dual characterisation of oscillation $\mathrm{W}^{\mathbf{a},p}$-gradient Young measures in terms of $\mathrm{W}^{\mathbf{a},p}$-quasiconvex functions. This, and the following, section is based on [67] and the proofs we give are adaptations of the techniques presented therein. However, the root of the principal results (Theorems 3.7.1 and 3.7.2) goes back to the seminal work of Kinderlehrer and Pedregal (see [95]). Before we proceed with the dual characterisation we need to look more closely at the structure of the space of oscillation Young measures generated by **a**-gradients of sequences of functions in $\mathrm{W}^{\mathbf{a},p}$, which we will do now.

In the following we denote by $\mathcal{E}_p$ an analogue of the space E_p defined earlier for functions that do not depend on the spatial variable x. We set

$$\mathcal{E}_p := \left\{ g \in C(\mathbb{R}^{n\times m}) \colon \lim_{|W|\to\infty} \frac{g(W)}{1+|W|^p} \text{ exists in } \mathbb{R} \right\},$$

and we equip $\mathcal{E}_p$ with the norm

$$\|g\|_{\mathcal{E}_p} := \sup_{W\in\mathbb{R}^{n\times m}} \frac{|g(W)|}{1+|W|^p}.$$

Similarly to before, we see that $\mathcal{E}_p$ is a separable Banach space. Furthermore, the space of probability measures with finite p-th moment is clearly a subset of $\mathcal{E}_p^*$ through the pairing

$$\langle \nu, g\rangle := \int_{\mathbb{R}^{n\times m}} g(W)\,\mathrm{d}\nu(W),$$

for $\nu \in \{\mu \in \mathcal{P}(\mathbb{R}^{n\times m}) \colon \int_{\mathbb{R}^{n\times m}} |W|^p \,\mathrm{d}\nu(W) < \infty\}$ and $g \in \mathcal{E}_p$. In particular, the space of homogeneous oscillation $\mathrm{W}^{\mathbf{a},p}$-gradient Young measures is a subset of $\mathcal{E}_p^*$. For future use, we note the following technical result.

Lemma 3.6.1. *Fix a bounded open set $E \subset \mathbb{R}^N$ and a function $\varphi \in \mathrm{W}_0^{\mathbf{a},p}(Q)$. Suppose that for every $j \in \mathbb{N}$ we have fixed a countable family $\{Q_i^j\}_i = \{Q_{x_i^j}(r_i^j)\}_i$ of pairwise disjoint anisotropic boxes contained in E and with $\lim_{j\to\infty}\sup_i r_i^j = 0$. Assume furthermore that $\lim_{j\to\infty} |E \setminus \bigcup_i Q_i^j| = 0$. Define*

$$\varphi_j := \sum_i r_i^j \varphi((r_i^j)^{-1} \odot (x - x_i^j)).$$

Then $\varphi_j \rightharpoonup 0$ weakly in $\mathrm{W}_0^{\mathbf{a},p}(E)$. The sequence $\nabla_{\mathbf{a}}\varphi_j$ is p-equiintegrable and generates the homogeneous oscillation Young measure $\nabla_{\mathbf{a}}\varphi_{\#}\left(\frac{\mathcal{L}^N \llcorner Q}{|Q|}\right)$.

Proof. First of all note that the function φ_j is well-defined as an element of $\mathrm{W}_0^{\mathbf{a},p}(E)$. To see this observe that for every $\langle \beta, \mathbf{a}^{-1} \rangle \leqslant 1$ and every i we have

$$\|\partial^\beta [r_i^j \varphi((r_i^j)^{-1} \odot x)]\|_{\mathrm{L}^p(Q_i^j)} = (r_i^j)^{1-\langle \beta, \mathbf{a}^{-1}\rangle} \frac{|Q_i^j|}{|Q|} \|\partial^\beta \varphi\|_{\mathrm{L}^p(Q)}.$$

Furthermore, this shows that the sequence φ_j converges strongly to 0 in $\mathrm{L}^p(E)$ and that it is bounded in $\mathrm{W}_0^{\mathbf{a},p}(E)$, thus converges weakly to 0 in that space, since the only possible limit of any weakly convergent subsequence is 0. The p-equiintegrability of $\{\nabla_{\mathbf{a}} \varphi_j\}$ follows from the fact that for any $M > 0$ we have

$$\sup_j \int_E |\nabla_{\mathbf{a}} \varphi_j|^p \mathbb{1}_{|\nabla_{\mathbf{a}} \varphi_j| > M} \, \mathrm{d}x = \frac{|E|}{|Q|} \int_Q |\nabla_{\mathbf{a}} \varphi|^p \mathbb{1}_{|\nabla_{\mathbf{a}} \varphi| > M} \, \mathrm{d}x,$$

which we get by a simple change of variables on each Q_i^j.

To show that $\nabla_{\mathbf{a}} \varphi_j$ generates the desired Young measure let us fix arbitrary functions $f \in C_c^\infty(E)$ and $g \in \mathcal{E}_p$. Then

$$\int_E f(x) g(\nabla_{\mathbf{a}} \varphi_j(x)) \, \mathrm{d}x = \sum_i \int_{Q_i^j} f(x) g(\nabla_{\mathbf{a}} \varphi_j(x)) \, \mathrm{d}x.$$

Since f is Lipschitz we may write

$$\left| \sum_i \int_{Q_i^j} (f(x) - f(x_i^j)) g(\nabla_{\mathbf{a}} \varphi_j(x)) \, \mathrm{d}x \right| \leqslant \varepsilon_j \left| \sum_i \int_{Q_i^j} g(\nabla_{\mathbf{a}} \varphi_j(x)) \, \mathrm{d}x \right| \leqslant \varepsilon_j \frac{|E|}{|Q|} \int_Q |g(\nabla_{\mathbf{a}} \varphi)| \, \mathrm{d}x,$$

where $\varepsilon_j = \varepsilon_j(f, \sup_i r_i^j) \to 0$ as $j \to \infty$. Changing variables on each Q_i^j we may write that

$$\sum_i \int_{Q_i^j} f(x_i^j) g(\nabla_{\mathbf{a}} \varphi_j(x)) \, \mathrm{d}x = \sum_i \frac{|Q_i^j|}{|Q|} f(x_i^j) \int_Q g(\nabla_{\mathbf{a}} \varphi(y)) \, \mathrm{d}y.$$

Again, using the fact that f is Lipschitz we note that, since the term $\sum_i |Q_i^j| f(x_i^j)$ is a Riemann sum of f on the anisotropic boxes Q_i^j we deduce that it converges, as $j \to \infty$ to the integral $\int_E f(x) \, \mathrm{d}x$, so that finally

$$\lim_{j \to \infty} \int_E f(x) g(\nabla_{\mathbf{a}} \varphi_j(x)) \, \mathrm{d}x = \left(\int_E f(x) \, \mathrm{d}x \right) \left(\fint_Q g(\nabla_{\mathbf{a}} \varphi(y)) \, \mathrm{d}y \right),$$

and a standard density argument on f and g ends the proof. □

Since every bounded open set may be covered, up to a subset of measure 0, with arbitrarily small anisotropic boxes we get the following:

Corollary 3.6.2. *For any* $\varphi \in \mathrm{W}_0^{\mathbf{a},p}(Q)$ *and any bounded open set* $E \in \mathbb{R}^N$ *the measure given by* $\nabla_{\mathbf{a}}\varphi_{\#}\left(\frac{\mathcal{L}^N \llcorner Q}{|Q|}\right)$ *is a homogeneous (oscillation)* $\mathrm{W}^{\mathbf{a},p}$*-gradient Young measure on* E*.*

We denote by $\mathbb{H}_0^p(Q)$ the space of homogeneous (oscillation) $\mathrm{W}^{\mathbf{a},p}$-gradient Young measures on Q with barycentre 0.

Lemma 3.6.3. *The set* $\mathbb{H}_0^p$ *is convex.*

Proof. Fix $\nu, \mu \in \mathbb{H}_0^p$ and $\theta \in (0,1)$. Let $v_j, u_j \subset C_c^\infty(Q)$ satisfy $v_j \rightharpoonup 0$, $u_j \rightharpoonup 0$ in $\mathrm{W}_0^{\mathbf{a},p}(Q)$ (because ν and μ have their barycenters at 0) with $\nabla_{\mathbf{a}}v_j$, $\nabla_{\mathbf{a}}u_j$ generating ν and μ respectively.

Pick a sequence of smooth cut-off functions $\varphi_i \in C_c^\infty(Q;[0,1])$ with $\varphi_i \nearrow \mathbb{1}_{(0,\theta)\times Q^{N-1})}$ and such that $|\{\varphi_i \neq \mathbb{1}_{(0,\theta)\times Q^{N-1})}\}| \to 0$ with $i \to \infty$. Here Q^{N-1} is the $(N-1)$ dimensional cube. Define $w_j^i := \varphi_i v_j + (1-\varphi_i)u_j$. Then

$$\nabla_{\mathbf{a}}w_j^i = \varphi_i \nabla_{\mathbf{a}}v_j + (1-\varphi_i)\nabla_{\mathbf{a}}u_j + R(i,j),$$

where $R(i,j)$ is the remainder term that includes the terms where we put some of the derivatives on φ or $(1-\varphi)$. Due to Lemma 2.2.7 the lower order derivatives of v_j and u_j converge stronly to 0 in L^p, thus

$$\|R(i,j)\|_{\mathrm{L}^p} = \|\nabla_{\mathbf{a}}w_j^i - (\varphi_i\nabla_{\mathbf{a}}v_j + (1-\varphi_i)\nabla_{\mathbf{a}}u_j\|_{\mathrm{L}^p} \leqslant C(\varphi_i)(\|\widehat{\nabla_{\mathbf{a}}}v_j\|_{\mathrm{L}^p} + \|\widehat{\nabla_{\mathbf{a}}}u_j\|_{\mathrm{L}^p}) \to 0$$

with $j \to \infty$ for any fixed i (here $C(\varphi_i)$ is a finite constant that depends on the function φ_i). Given this we may select a subsequence $j(i) \to \infty$ as $i \to \infty$ such that

$$\|\nabla_{\mathbf{a}}w_{j(i)}^i - (\varphi_i\nabla_{\mathbf{a}}v_{j(i)} + (1-\varphi_i)\nabla_{\mathbf{a}}u_{j(i)})\|_{\mathrm{L}^p} \to 0$$

as $i \to \infty$. Let $w_i := w_{j(i)}^i \in C_c^\infty(Q)$. It is straightforward to see that $\nabla_{\mathbf{a}}w_i$ generates $\mathbb{1}_{(0,\theta)\times Q^{N-1}}\nu + \mathbb{1}_{(\theta,1)\times Q^{N-1}}\mu$ as its oscillation Young measure. That is because this clearly holds for $\mathbb{1}_{(0,\theta)\times Q^{N-1}}\nabla_{\mathbf{a}}v_{j(i)} + \mathbb{1}_{(\theta,1)\times Q^{N-1}}\nabla_{\mathbf{a}}u_{j(i)}$ and, by construction, the difference of this sequence and $(\varphi_i\nabla_{\mathbf{a}}v_{j(i)} + (1-\varphi_i)\nabla_{\mathbf{a}}u_{j(i)})$ converges to 0 in measure, whilst the difference of $\nabla_{\mathbf{a}}w_i$ and the latter sequence converges to 0 in L^p. Finally, w_i is compactly supported in Q, so we may extend it periodically and treat it as a function on the whole of $\mathbb{R}^N$ and define

$$w_i^k(x) := R_k^{-1}w_i(R_k \odot x),$$

where $R_k := k^{a_1\cdot\dots\cdot a_N}$. Then, by Lemma 3.6.1, for all $\varphi \in C_0^\infty$ and $\psi \in \mathcal{E}_p$ we have

$$\lim_{i\to\infty}\lim_{k\to\infty} \fint_Q \varphi(x)\psi(\nabla_{\mathbf{a}}w_i^k(x))\,\mathrm{d}x = \lim_{j\to\infty}\fint_Q \varphi(x)\left(\fint_Q \psi(\nabla_{\mathbf{a}}w_i(y))\,\mathrm{d}y\right)\mathrm{d}x. \tag{3.11}$$

Finally, since we have already identified the measure generated by w_i, we may write

$$\lim_{i\to\infty}\lim_{k\to\infty}\fint_Q \varphi(x)\psi(\nabla_{\mathbf{a}} w_i^k(x))\,\mathrm{d}x = \fint_Q \varphi(x)\,\mathrm{d}x(\theta\langle\nu,\psi\rangle + (1-\theta)\langle\mu,\psi\rangle).$$

A standard density argument and a diagonal extraction in the separable spaces $\mathrm{L}^1(Q)$ and $C_0(\mathbb{R}^{n\times m})$ let us obtain a subsequence $\widetilde{w}_l \subset \{w_i^k\} \subset C^\infty$ with **a**-gradients generating the measure $\theta\nu + (1-\theta)\mu$, which ends the proof.

□

Lemma 3.6.4. *The set $\mathbb{H}_0^p$ is relatively closed in $\mathcal{P}(\mathbb{R}^{n\times m}) \cap \{\mu\colon \int_{\mathbb{R}^N} |W|^p\,\mathrm{d}\mu < \infty\}$ with respect to the weak* topology on $\mathcal{E}_p^*$.*

Proof. Fix an arbitrary $\nu \in \overline{\mathbb{H}_0^p}^{E^*} \cap \mathcal{P}(\mathbb{R}^{n\times m})$ and let $\{f_i\} \subset C^\infty(Q)$, $\{g_j\} \subset C_c^\infty(\mathbb{R}^{n\times m})$ be countable dense subsets of $L^1(Q)$ and $C_0(\mathbb{R}^{n\times m})$ respectively. Take also $f_0(x) \equiv 1$ and $g_0(W) := |W|^p$. By definition of the weak* topology, for any fixed $g \in C_0(\mathbb{R}^{n\times m})$ there exists a $\nu_k \in \mathbb{H}_0^p$ with

$$|\langle \nu - \nu_k, g\rangle| < \frac{1}{2k}.$$

Through a diagonal argument we may ensure that this is satisfied simultaneously for any finite set of g's, i.e., for any $k \in \mathbb{N}$ there exists a $\nu_k \in \mathbb{H}_0^p$ such that

$$|\langle \nu - \nu_k, g_j\rangle| < \frac{1}{2k}, \text{ for all } j \in \{0, 1, \dots, k\}.$$

Since $\nu_k \in \mathbb{H}_0^p$ we may find a sequence $\{w_j^k\} \subset \mathrm{W}_0^{\mathbf{a},p}(Q)$ with **a**-gradients generating ν_k.

Theorem 3.1.1 implies that for any $g \in C_c^\infty(\mathbb{R}^{n\times m})$ we have $g(\nabla_{\mathbf{a}} w_j^k) \rightharpoonup \langle g, \nu_k\rangle$ in $\mathrm{L}^1(Q)$. Another diagonal extraction and the triangle inequality let us establish existence of a sequence $\{w_k\} \subset \mathrm{W}_0^{\mathbf{a},p}(Q)$ such that

$$\left|\langle \nu, g_j\rangle \int_Q f_i\,\mathrm{d}x - \int_Q f_i g_j(\nabla_{\mathbf{a}} w_k)\,\mathrm{d}x\right| < \frac{1}{k}, \text{ for all } 0 \leqslant i, j \leqslant k, \tag{3.12}$$

as all f_i's are smooth and therefore bounded, so that they are admissible test functions for weak convergence in L^1.

Setting $i = j = 0$ shows that $\{\nabla_{\mathbf{a}} w_k\}$ is bounded in $\mathrm{L}^p(Q)$, therefore we may find a subsequence generating some $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure μ. For notational simplicity we assume that the entire sequence does. From (3.12) and the fact that $g_j(\nabla_{\mathbf{a}} w_k) \rightharpoonup \langle \mu, g_j\rangle$ in L^1 as $k \to \infty$ we infer that

$$\langle \nu, g_j\rangle \int_Q f_i\,dx = \int_Q f_i \langle \mu, g_j\rangle\,dx, \text{ for all } i, j.$$

By density of f_i in $\mathrm{L}^p(Q)$ we deduce that $\langle \mu, g_j \rangle = \langle \nu, g_j \rangle$ in $(\mathrm{L}^p)^*$. In particular, they are equal almost everywhere, so that for almost every $x \in Q$ we have $\langle \mu_x, g_j \rangle = \langle \nu, g_j \rangle$ for all j. By density of $\{g_j\}$ in $C_0(\mathbb{R}^{n\times m})$ we deduce that $\mu_x = \nu$ for almost all x, which shows that μ is in fact homogeneous so $\nu = \mu \in \mathbb{H}_0^p$, which ends the proof. □

Lemma 3.6.5. *If a sequence $\{\nu_j\} \subset \mathcal{P}(\mathbb{R}^{n\times m}) \cap \mathcal{E}_p^*$ converges to some $\nu \in \mathcal{P}(\mathbb{R}^{n\times m}) \cap \mathcal{E}_p^*$ in the space $\mathcal{E}_p^*$ then it also converges in the sense of weak convergence of probability measures (or weak*). In particular, by the portmanteau theorem, we have*

$$\lim_{j\to\infty} \int_{\mathbb{R}^{n\times m}} g \,\mathrm{d}\nu_j = \int_{\mathbb{R}^{n\times m}} g \,\mathrm{d}\nu$$

for all bounded and continuous functions g, and

$$\liminf_{j\to\infty} \int_{\mathbb{R}^{n\times m}} g \,\mathrm{d}\nu_j \geqslant \int_{\mathbb{R}^{n\times m}} g \,\mathrm{d}\nu$$

for all lower semicontinuous functions g bounded from below.

Proof. Immediately follows from bounded continuous functions being a subset of $\mathcal{E}_p$. □

3.7 Dual characterisation of oscillation $\mathrm{W}^{\mathbf{a},p}$-gradient Young measures

With the groundwork established in the previous section we are now ready to prove the main results of this chapter. Let us note that our proofs closely follow the approach of Fonseca and Müller in [67], however they need to be adapted to the mixed smoothness setting, which is why we present them in full.

Theorem 3.7.1. *A probability measure $\mu \in \mathcal{P}(\mathbb{R}^{n\times m})$ is a homogeneous oscillation $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure with mean W_0 if and only if μ satisfies $\int_{\mathbb{R}^{n\times m}} W \,\mathrm{d}\mu(W) = W_0$, $\int_{\mathbb{R}^{n\times m}} |W|^p \,\mathrm{d}\mu(W) < \infty$ and*

$$\int_{\mathbb{R}^{n\times m}} g(W) \,\mathrm{d}\mu(W) \geqslant \mathcal{Q}g(W_0)$$

for all $g \in \mathcal{E}_p$.

Proof. First observe that it is enough to consider the case $W_0 = 0$, as we may always take a translation. For the proof of this case we argue by contradiction. Suppose that $\nu \in \mathcal{P}(\mathbb{R}^{n\times m})$ satisfies

$$\begin{cases} \int_{\mathbb{R}^{n\times m}} W \,\mathrm{d}\nu = 0, \\ \int_{\mathbb{R}^{n\times m}} |W|^p \,\mathrm{d}\nu(W) < \infty, \\ \int_{\mathbb{R}^{n\times m}} g(W) \,\mathrm{d}\nu(W) \geqslant \mathcal{Q}g(0) \text{ for all } g \in \mathcal{E}_p, \end{cases}$$

but $\mu \notin \mathbb{H}_0^p$. The p-th moment assumption on ν implies that $\nu \in \mathcal{E}_p^*$. By Lemmas 3.6.3 and 3.6.4 and the Hahn-Banach separation theorem there exist $g \in \mathcal{E}_p$ and $\alpha \in \mathbb{R}$ such that

$$\mathcal{Q}g(0) \leqslant \langle \nu, g \rangle < \alpha \leqslant \langle \mu, g \rangle \text{ for all } \mu \in \mathbb{H}_0^p.$$

For any $\varphi \in C_c^\infty(Q)$ the measure $(\nabla_{\mathbf{a}}\varphi)_\# \frac{\mathcal{L}^N \llcorner Q}{|Q|}$ is an element of $\mathbb{H}_0^p$ (see Corollary 3.6.2) so that

$$\fint_Q g(\nabla_{\mathbf{a}}\varphi(x))\,\mathrm{d}x \geqslant \alpha.$$

Taking infimum over all such φ's yields

$$\mathcal{Q}g(0) = \inf_{\varphi \in C_c^\infty(Q)} \fint_Q g(\nabla_{\mathbf{a}}\varphi(x))\,\mathrm{d}x \geqslant \alpha,$$

which is a contradiction with $\mathcal{Q}g(0) \leqslant \langle \nu, g \rangle < \alpha$ and shows that ν is indeed in the set $\mathbb{H}_0^p$.

For the reverse implication let $\varphi_k \in C_c^\infty$ be such that $\{\nabla_{\mathbf{a}}\varphi_k\}$ is a p-equiintegrable sequence generating ν (which is possible thanks to Corollary 3.2.4). Then for all k we have

$$\int_{\mathbb{R}^{n\times m}} g(\nabla_{\mathbf{a}}\varphi_k(x))\,\mathrm{d}x \geqslant \mathcal{Q}g(0)$$

by definition of $\mathcal{Q}g$. On the other hand p-equiintegrability of $\{\nabla_{\mathbf{a}}\varphi_k\}$ and the growth bound on g imply that we may use point iii) of Theorem 3.1.1 to deduce that

$$\lim_{k\to\infty} \int_{\mathbb{R}^{n\times m}} g(\nabla_{\mathbf{a}}\varphi_k(x))\,\mathrm{d}x = \int_{\mathbb{R}^{n\times m}} g\,\mathrm{d}\nu,$$

as for a continuous function the boundedness from below assumption is irrelevant, since we may simply consider the positive and negative parts of g separately. This and the previous estimate finish the proof. □

A similar result holds for non-homogeneous Young measures. We state it in the following:

Theorem 3.7.2. *Fix a bounded open domain Ω satisfying the weak $\mathbf{a}$-horn condition. Let $\{\nu_x\}_{x\in\Omega}$ be a weak* measurable family of probability measures on $\mathbb{R}^{n\times m}$. Then there exists a $\mathrm{W}^{\mathbf{a},p}(\Omega)$-bounded sequence $\{v_n\} \subset \mathrm{W}^{\mathbf{a},p}(\Omega)$ with $\{\nabla_{\mathbf{a}}v_n\}$ generating the oscillation Young measure ν if and only if the following conditions hold:*

i) there exists $v \in \mathrm{W}^{\mathbf{a},p}(\Omega)$ such that

$$\nabla_{\mathbf{a}}v(x) = \langle \nu_x, \mathrm{Id} \rangle \textit{ for a.e. } x \in \Omega;$$

ii)

$$\int_\Omega \int_{\mathbb{R}^{n\times m}} |W|^p \, \mathrm{d}\nu_x(W) \, \mathrm{d}x < \infty;$$

iii) for a.e. $x \in \Omega$ and all $g \in \mathcal{E}_p$ we have

$$\langle \nu_x, g \rangle \geqslant \mathcal{Q}g(\langle \nu_x, \mathrm{Id} \rangle).$$

Proof. We define the space

$$\mathsf{E} := C(\overline{\Omega}; \mathcal{E}_p),$$

equipped with the natural supremum norm:

$$||f||_{\mathsf{E}} := \sup_{x \in \overline{\Omega}} ||f(x, \cdot)||_{\mathcal{E}_p}.$$

Furthermore, we introduce the following sets:

$$\mathbb{X} := \left\{ \nu \colon \Omega \to \mathcal{P}(\mathbb{R}^{n\times m}) \colon \nu \text{ is weak* measurable, } \langle \nu_x, |\cdot|^p \rangle \in \mathrm{L}^1,\ \langle \nu_x, \mathrm{Id} \rangle = 0 \text{ a.e. } x \in \Omega \right\},$$

$$\mathbb{Y} := \{ \nu \in \mathbb{X} \colon \nu \text{ is generated by a } p\text{- equiintegrable sequence } \nabla_{\mathbf{a}} u_n \text{ for } u_n \in \mathrm{W}^{\mathbf{a},p}(\Omega) \},$$

$$\mathbb{W} := \{ \nu \in \mathbb{X} : \langle \nu_x, g \rangle \geqslant \mathcal{Q}g(0) \text{ for a.e. } x \in \Omega \text{ and all } g \in \mathcal{E}_p \}.$$

Assuming that a family of measures $\nu = \{\nu_x\}_x$ satisfies the assumptions of our Theorem, we may, just as in the case of homogeneous Young measures, restrict ourselves to considering families satisfying

$$\langle \nu_x, \mathrm{Id} \rangle = 0 \text{ for a.e. } x \in \Omega.$$

Then our assertion is equivalent to showing that $\mathbb{W} \subset \mathbb{Y}$. In order to do so, we first show that $\mathbb{Y}$ is relatively closed in $\mathbb{X}$ with respect to the weak star topology induced by E. Then, we show that piecewise-homogeneous Young measures form a dense subset of $\mathbb{W}$ and, finally, show that these measures all belong to $\mathbb{Y}$ as well, thus ending the proof.

For the first part observe that, due to the p-th moment restriction on ν in condition ii), it is immediate that $\mathbb{W}, \mathbb{Y} \subset \mathbb{X} \subset \mathsf{E}^*$. Showing that

$$\overline{\mathbb{Y}}^{\mathsf{E}^*} \cap \mathbb{X} = \mathbb{Y},$$

is entirely analogous to the proof of Lemma 3.6.4, hence we simply admit this result without copying the aforementioned argument.

For the second part, for any $k \in \mathbb{N}$, we define

$$\mathcal{G}_k := \left\{ \frac{1}{k} \odot (y + Q) \colon y \in \mathbb{Z}^N,\ \frac{1}{k} \odot (y + Q) \subset \Omega \right\},$$

where Q is the unit cube, and we set

$$G_k := \bigcup_{\mathcal{U}\in\mathcal{G}_k} \mathcal{U}$$

to be the part of Ω covered by boxes in $\mathcal{G}_k$. Observe that clearly $\lim_{k\to\infty} |\Omega\setminus G_k| = 0$. We now define

$$\mathbb{W}_k := \left\{\nu \in \mathbb{W}\colon \nu\big|_{\mathcal{U}} \text{ is homogeneous for any } \mathcal{U}\in\mathcal{G}_k \text{ and } \nu\big|_{(\Omega\setminus G_k)} = \delta_0\right\}.$$

We then set

$$D := \bigcup_{k\in\natural} \mathbb{W}_k,$$

and we aim to show that

$$\mathbb{W} \subset \overline{D}^{\mathsf{E}^*}.$$

Fix an arbitrary $\nu \in \mathbb{W}$ and define

$$\nu_x^k := \begin{cases} \frac{1}{|\mathcal{U}|}\int_{\mathcal{U}} \nu_y \,\mathrm{d}y & \text{if } x\in\mathcal{U} \text{ with } \mathcal{U}\in\mathcal{G}_k,\\ \delta_0 \quad \text{otherwise.}\end{cases}$$

Clearly $\nu^k \in \mathbb{W}_k$, so all we need to show is that

$$\langle \nu^k, f\rangle \to \langle \nu, f\rangle \text{ for any } f\in\mathsf{E}. \tag{3.13}$$

Fix an arbitrary $f \in \mathsf{E}$. Since f is continuous on $\overline{\Omega}$ it is uniformly continuous, so we may let ω be its modulus of uniform continuity, so that

$$\omega(\delta) := \sup\{\|f(x,\cdot) - f(y,\cdot)\|_{\mathcal{E}_p}\colon x,y\in\overline{\Omega}, |x-y|\leqslant\delta\}.$$

For every $\mathcal{U}\in\mathcal{G}_k$ pick a point $x_{\mathcal{U}}\in\mathcal{U}$, for example the center of $\mathcal{U}$, so that

$$|x_{\mathcal{U}} - x| \leqslant C(k) \text{ for all } x\in\mathcal{U}\in\mathcal{G}_k,$$

where $C(k)$ is a constant depending on k and $\mathbf{a}$ and satisfying $C(k)\to 0$ as $k\to\infty$. Explicitly $C(k) = \sqrt{\sum_{i=1}^N k^{-2/a_i}}$. We may now estimate

$$\left|\int_{\mathcal{U}}\int_{\mathbb{R}^{n\times m}} f(x,W)\,\mathrm{d}\nu_x(W)\,\mathrm{d}x - \int_{\mathcal{U}}\int_{\mathbb{R}^{n\times m}} f(x,W)\,\mathrm{d}\nu_x^k(W)\,\mathrm{d}x\right| \leqslant$$

$$\left|\int_{\mathcal{U}}\int_{\mathbb{R}^{n\times m}} f(x_{\mathcal{U}},W)\,\mathrm{d}\nu_x(W)\,\mathrm{d}x - \int_{\mathcal{U}}\int_{\mathbb{R}^{n\times m}} f(x_{\mathcal{U}},W)\,\mathrm{d}\nu_x^k(W)\,\mathrm{d}x\right| +$$

$$+\omega(C(k))||f||_{\mathsf{E}}\left(\int_{\mathcal{U}}\int_{\mathbb{R}^{n\times m}}(1+|W|^p)\,\mathrm{d}\nu_x(W)\,\mathrm{d}x + \int_{\mathcal{U}}\int_{\mathbb{R}^{n\times m}}(1+|W|^p)\,\mathrm{d}\nu_x^k(W)\,\mathrm{d}x\right).$$

Now, since we integrate over $\mathcal{U}$ neither of the integrands depend on x. Therefore, due to the definition of ν_x^k the first term vanishes and we may replace the second one by twice the integral with respect to ν_x, so that finally

$$\left|\int_{\mathcal{U}}\int_{\mathbb{R}^{n\times m}} f(x,W)\,\mathrm{d}\nu_x(W)\,\mathrm{d}x - \int_{\mathcal{U}}\int_{\mathbb{R}^{n\times m}} f(x,W)\,\mathrm{d}\nu_x^k(W)\,\mathrm{d}x\right| \leqslant$$
$$\leqslant 2\omega(C(k))||f||_{\mathsf{E}}\int_{\mathcal{U}}\int_{\mathbb{R}^{n\times m}}(1+|W|^p)\,\mathrm{d}\nu_x(W)\,\mathrm{d}x.$$

Summing over all $\mathcal{U}\in\mathcal{G}_k$ and adding the remainder we have

$$\begin{aligned}\left|\langle\nu^k,f\rangle-\langle\nu,f\rangle\right| \leqslant \quad & 2\omega(C(k))\|f\|_{\mathsf{E}}\int_{G_k}\int_{\mathbb{R}^{n\times m}}(1+|W|^p)\,\mathrm{d}\nu_x(W)\,\mathrm{d}x + \\ & 2\|f\|_{\mathsf{E}}\int_{\Omega\setminus G_k}\int_{\mathbb{R}^{n\times m}}(1+|W|^p)\,\mathrm{d}\nu_x(W)\,\mathrm{d}x.\end{aligned}$$

Since

$$\int_{G_k}\int_{\mathbb{R}^{n\times m}}(1+|W|^p)\,\mathrm{d}\nu_x(W)\,\mathrm{d}x \leqslant \int_{\Omega}\int_{\mathbb{R}^{n\times m}}(1+|W|^p)\,\mathrm{d}\nu_x(W)\,\mathrm{d}x < \infty$$

and

$$\lim_{k\to\infty}\omega(C(k))=0,$$

the first term in our inequality tends to 0 as $k\to\infty$. The second one does the same, as the Lebesgue measure of the set $\Omega\setminus G_k$ tends to 0 with $k\to\infty$ and the function given by $x\mapsto\int_{\mathbb{R}^{n\times m}}(1+|W|^p)\,\mathrm{d}\nu_x(W)$ is Lebesgue integrable. This establishes (3.13) and so proves that

$$\mathbb{W}\subset\overline{D}^{\mathsf{E}^*}.$$

What is left to show is that $D=\bigcup_k\mathbb{W}_k\subset Y$. Fix $k\in\mathbb{N}$ and let $G_k=\{Q_i\}_{i=1}^I$ for some $I<\infty$. Fix $\nu\in\mathbb{W}_k$ with $\nu\big|_{Q_i}=\nu^i$. By Theorem 3.7.1 we know that ν^i is a homogeneous $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure with barycentre 0, so there exists a sequence $\{u_j^i\}_j\in\mathrm{W}_0^{\mathbf{a},p}(Q)$ with $\{\nabla_{\mathbf{a}}u_j^i\}_j$ generating ν^i. Denote by c_i the centre of Q_i so that $Q_i=c_i+\frac{1}{k}\odot Q$. Extend u_j^i by 0 to the whole of $\mathbb{R}^N$ and set $\widetilde{u_j^i}(x):=u_j^i(k\odot(x-c_i))$. Then $\{\nabla_{\mathbf{a}}\widetilde{u_j^i}(x)\}_j$ are supported in Q_i and generate ν^i. Thus, if we set $u_j:=\sum_{i=1}^I\widetilde{u_j^i}\in\mathrm{W}_0^{\mathbf{a},p}(\Omega)$, we immediately see that $\{\nabla_{\mathbf{a}}u_j\}$ generate ν, thus $\nu\in\mathbb{Y}$.

This shows that $D\subset Y$, so that

$$W\subset\overline{D}^{\mathsf{E}^*}\subset\overline{\mathbb{Y}}^{\mathsf{E}^*}=\mathbb{Y},$$

which ends the proof. □

Chapter 4

Existence

We aim to study existence of solutions to variational problems in the mixed smoothness setting via the direct method mentioned in the introduction. In this chapter we study the two necessary ingredients: coercivity, in Section 4.1, and lower semicontinuity, in Section 4.2. We elaborate on lower semicontinuity further in Section 4.3, where we establish formulas for sequentially weakly lower semicontinuous envelopes of relaxation for functionals in our framework. The principal results that we obtain are as follows: in Theorem 4.1.3 we characterise coercivity as an intrinsic property of the integrand that may be phrased in terms of its $\mathbf{a}$-quasiconvexity; in Theorem 4.2.8 we identify $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity as equivalent to lower semicontinuity of functionals given by continuous integrands of p-growth; in Theorem 4.3.2 we obtain a relaxation formula for integrands of p-growth under certain additional assumptions; finally, in Theorem 4.3.13 we prove a relaxation formula for integrands that may take the value $+\infty$, at the cost of relaxing the notion of convergence considered.

4.1 Coercivity

The purpose here is to prove Theorem 4.1.3 characterising coercivity of functionals in the mixed smoothness setting. This, combined with the lower semicontinuity results to follow, identifies a number of situations where the minimisation problem under a fixed Dirichlet boundary condition is well-posed and allows one to talk about its solutions and, later on, their regularity. Our result is a generalisation to the mixed smoothness framework of a recent result due to Chen and Kristensen from [37], that was further improved by Gmeineder and Kristensen in [82]. Certain parts of Chen and Kristensen's argument carry through with only cosmetic adjustments, thus we try to focus on the parts that require more substantial changes.

Here $F\colon \mathbb{R}^{n\times m}\to\mathbb{R}$ is a continuous integrand satisfying the growth condition

$$|F(X)| \leqslant C(|X|^p + 1) \tag{4.1}$$

for some $C \in (0,\infty)$ and all $X \in \mathbb{R}^{n\times m}$. For a non-empty bounded open set $\Omega \subset \mathbb{R}^N$ we denote by $\mathrm{I}(\cdot,\Omega)\colon \mathrm{W}^{\mathbf{a},p}(\Omega)\to\mathbb{R}$ the mapping defined by $\mathrm{I}(u,\Omega) := \int_\Omega F(\nabla_{\mathbf{a}} u)\,\mathrm{d}x$. To be able to properly introduce a Dirichlet problem for this functional, we define the Dirichlet class

$$\mathrm{W}^{\mathbf{a},p}_g(\Omega) := \{g+\varphi\colon \varphi \in \mathrm{W}^{\mathbf{a},p}_0(\Omega)\},$$

for a fixed $g \in \mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N)$. Observe that we require the boundary datum g to be defined on the whole of $\mathbb{R}^N$, rather than just $\Omega \subset \mathbb{R}^N$, thus bypassing possible trace issues.

Fix an exponent $q \in [1,p]$. Following [37] we introduce the following notions of coercivity:

Definition 4.1.1. *We say that $\mathrm{I}(\cdot,\Omega)$ is L^q coercive on $\mathrm{W}^{\mathbf{a},p}_g(\Omega)$ if for any sequence $u_j \in \mathrm{W}^{\mathbf{a},p}_g(\Omega)$ with $\|\nabla_{\mathbf{a}} u_j\|_{\mathrm{L}^q}\to\infty$ one has $\mathrm{I}(u_j,\Omega)\to\infty$.*

Definition 4.1.2. *We say that $\mathrm{I}(\cdot,\Omega)$ is L^q mean coercive on $\mathrm{W}^{\mathbf{a},p}_g(\Omega)$ if for any $u \in \mathrm{W}^{\mathbf{a},p}_g(\Omega)$ we have*

$$\mathrm{I}(u,\Omega) \geqslant C_1\|\nabla_{\mathbf{a}} u\|^q_{\mathrm{L}^q} - C_2$$

for some strictly positive constants C_1, C_2 independent of u, but not necessarily of g.

As in [37] (see also Proposition 3.1 in [82]), our main goal here is to establish the following:

Theorem 4.1.3. *Let $F\colon \mathbb{R}^{n\times m}\to\mathbb{R}$ be a continuous integrand satisfying the growth condition (4.1). Then, for any $q\in[1,p]$ the following are equivalent:*

a) For any bounded open set $\Omega \subset \mathbb{R}^N$ and any boundary datum $g \in \mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N)$ the functional $\mathrm{I}(\cdot,\Omega)$ is L^q coercive on $\mathrm{W}^{\mathbf{a},p}_g(\Omega)$.

b) There exist a non-empty bounded open set $\Omega \subset \mathbb{R}^N$ and a boundary datum $g \in \mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N)$ such that the functional $\mathrm{I}(\cdot,\Omega)$ is L^q coercive on $\mathrm{W}^{\mathbf{a},p}_g(\Omega)$.

c) For any bounded open set $\Omega \subset \mathbb{R}^N$ and any boundary datum $g \in \mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N)$ the functional $\mathrm{I}(\cdot,\Omega)$ is L^q mean coercive on $\mathrm{W}^{\mathbf{a},p}_g(\Omega)$.

d) There exist a non-empty bounded open set $\Omega \subset \mathbb{R}^N$ and a boundary datum $g \in \mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N)$ such that the functional $\mathrm{I}(\cdot,\Omega)$ is L^q mean coercive on $\mathrm{W}^{\mathbf{a},p}_g(\Omega)$.

e) For any bounded open set $\Omega \subset \mathbb{R}^N$ *and any boundary datum* $g \in \mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N)$ *all* $\mathrm{W}^{\mathbf{a},p}_g(\Omega)$ *minimising sequences for the functional* $\mathrm{I}(\cdot,\Omega)$ *are bounded in* $\mathrm{W}^{\mathbf{a},q}_g(\Omega)$*.*

f) There exist a non-empty bounded open set $\Omega \subset \mathbb{R}^N$ *and a boundary datum* $g \in \mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N)$ *such that all* $\mathrm{W}^{\mathbf{a},p}_g(\Omega)$ *minimising sequences for the functional* $\mathrm{I}(\cdot,\Omega)$ *are bounded in* $\mathrm{W}^{\mathbf{a},q}_g(\Omega)$*.*

g) There exist a constant $c > 0$ *and a point* $X_0 \in \mathbb{R}^{n\times m}$ *such that the integrand* $X \mapsto F(X) - c|X|^q$ *is* $\mathbf{a}$*-quasiconvex at* X_0*.*

Simply put, the above asserts that L^q coercivity, L^q mean coercivity, and boundedness of all minimising sequences are mutually equivalent, and it is enough to check either of these on a particular choice of domain and boundary datum. Furthermore, coercivity may be characterized in terms of $\mathbf{a}$-quasiconvexity of $F(\cdot) - c|\cdot|^q$ at a single point. That last statement is particularly interesting, as it means that understanding $\mathbf{a}$-quasiconvexity yields not only lower semicontinuity, but also coercivity results. Thus, both ingredients needed for the direct method may be phrased in terms of $\mathbf{a}$-quasiconvexity.

Lemma 4.1.4 (see Proposition 3.1 in [37])**.** *Under the hypotheses of Theorem 4.1.3 its point a) is equivalent to b), and point c) is equivalent to d).*

Proof. The fact that a) is equivalent to b) in the first order gradient case is the content of Proposition 3.1 in [37]. The proof of c) being equivalent to d) is very similar — we will do this here omitting the first part, as all that needs to be changed in the argument of [37] is the scaling factor.

It is obvious that c) implies d), thus we only need to prove the other implication. Let Ω_0 and g_0 be the domain and the boundary datum for which $\mathrm{I}(\cdot,\Omega_0)$ is L^q mean coercive on $\mathrm{W}^{\mathbf{a},p}_{g_0}(\Omega_0)$. Fix an arbitrary non-empty bounded open set $\Omega \subset \mathbb{R}^N$ and a boundary datum $g \in \mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N)$. As in Chen and Kristensen's proof, we take boxes $Q_{2r}(x_0) \Subset \Omega_0$ (for some $x_0 \in \Omega_0$) and $Q_R(0) \Supset \Omega$, and we fix cut-off functions $\varphi \in C^\infty_c(Q_{2r}(x_0))$ and $\rho \in C^\infty_c(Q_R(0))$ satisfying

$$\mathbf{1}_{Q_r(x_0)} \leqslant \varphi \leqslant \mathbf{1}_{Q_{2r}(x_0)},$$

$$\mathbf{1}_{\Omega} \leqslant \rho \leqslant \mathbf{1}_{Q_R(0)}.$$

Any $u \in \mathrm{W}^{\mathbf{a},p}_g(\Omega)$ may be seen as a function in $\mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N)$ if we extend it by $u = g$ outside Ω, simply from our definition of the Dirichlet classes. We may then cut-off outside $Q_R(0)$ by setting $\widetilde{w} := \rho u \in \mathrm{W}_0^{\mathbf{a},p}(Q_R(0))$. Define $w(x) := R^{-1}\widetilde{w}(R \odot x) \in \mathrm{W}_0^{\mathbf{a},p}(Q_1(0))$ and set

$$v(x) := (1-\varphi(x))g_0(x) + rw(r^{-1} \odot (x - x_0)) \text{ for } x \in \Omega_0.$$

Clearly $v \in \mathrm{W}_{g_0}^{\mathbf{a},p}(\Omega_0)$ and we may calculate, using a simple change of variables, that

$$\begin{aligned}\|\nabla_{\mathbf{a}} v\|_{\mathrm{L}^q(\Omega_0)}^q = &\int_{\Omega_0 \setminus Q_r(x_0)} |\nabla_{\mathbf{a}}(1-\varphi(x))g_0|^q \,\mathrm{d}x + \\ &(\frac{r}{R})^{|\mathbf{a}^{-1}|} \int_{Q_R(0)\setminus\Omega} |\nabla_{\mathbf{a}}(\rho g)|^q \,\mathrm{d}x + (\frac{r}{R})^{|\mathbf{a}^{-1}|} \|\nabla_{\mathbf{a}} u\|_{\mathrm{L}^q(\Omega)}^q,\end{aligned}$$

so that $\|\nabla_{\mathbf{a}} v\|_{\mathrm{L}^q(\Omega_0)}^q = D_1 + D_2\|\nabla_{\mathbf{a}} u\|_{\mathrm{L}^q(\Omega)}^q$, where D_i do not depend on u and $D_2 > 0$. In the same way we get

$$\begin{aligned}\int_{\Omega_0} F(\nabla_{\mathbf{a}} v)\,\mathrm{d}x = &\int_{\Omega_0 \setminus Q_r(x_0)} F(\nabla_{\mathbf{a}}(1-\varphi(x))g_0)\,\mathrm{d}x + \\ &(\frac{r}{R})^{|\mathbf{a}^{-1}|}\left(\int_{Q_R(0)\setminus\Omega} F(\nabla_{\mathbf{a}}(\rho g))\,\mathrm{d}x + \int_\Omega F(\nabla_{\mathbf{a}} u)\,\mathrm{d}x\right).\end{aligned}$$

Thus, $\mathrm{I}(u,\Omega)$ differs from $\mathrm{I}(v,\Omega_0)$ by a constant and a scaling. Since the same is true for the L^q norms of u and v we easily conclude that $\mathrm{I}(u,\Omega) \geqslant -c_1 + c_2\|u\|_{\mathrm{L}^q(\Omega)}^q$ for some $c_i > 0$, which ends the proof. □

Definition 4.1.5. *Let $F\colon \mathbb{R}^{n\times m} \to \mathbb{R}$ be a continuous integrand satisfying the growth condition* (4.1)*. For $q \in [1,p]$ and a non-empty bounded and open subset $\Omega \subset \mathbb{R}^N$ we define, for $t \geqslant 0$,*

$$\theta(t) = \theta_q^\Omega(t) := \inf\left\{ \fint_\Omega F(\nabla_{\mathbf{a}}\varphi)\,\mathrm{d}x \colon \varphi \in \mathrm{W}_0^{\mathbf{a},p}(\Omega,\mathbb{R}^n), \fint_\Omega |\nabla_{\mathbf{a}}\varphi|^q\,\mathrm{d}x \geqslant t \right\}.$$

Lemma 4.1.6 (see Lemma 2.1 in [37])**.** *For any open, bounded, and non-empty domains $\omega, \Omega \subset \mathbb{R}^N$ and any function $\varphi \in C_c^\infty(\Omega)$ there exists a function $\psi \in C_c^\infty(\omega)$ such that the pushforward measures $[\nabla_{\mathbf{a}}\varphi]_\# \left(\frac{\mathcal{L}^N\lfloor\Omega}{\mathcal{L}^N(\Omega)}\right)$ and $[\nabla_{\mathbf{a}}\psi]_\# \left(\frac{\mathcal{L}^N\lfloor\omega}{\mathcal{L}^N(\omega)}\right)$ are equal.*

Proof. This is proven using a standard exhaustion argument, as in Lemma 2.1 in [37]. We omit the proof here, as the only difference with the aforementioned paper is that one needs to change the scaling to our anisotropic variant. □

As an immediate corollary of the above we get the following:

Corollary 4.1.7. *The unit cube Q in the definition of the* **a***-quasiconvex envelope may be replaced by any other domain without changing the resulting function. That is, for any open bounded domain $\Omega \subset \mathbb{R}^N$ we have*

$$\mathcal{Q}F(\cdot) = \inf_{\varphi \in C_c^\infty(\Omega)} \fint_\Omega F(\cdot + \nabla_{\mathbf{a}}\varphi(x))\,\mathrm{d}x.$$

Lemma 4.1.8. *Suppose that F is a continuous integrand satisfying the growth condition* (4.1)*. Then $\mathcal{Q}F$ is either identically equal to $-\infty$ or it is real-valued everywhere and satisfies the growth condition* (4.1)*, albeit possibly with a larger constant.*

Proof. Assume, for a contradiction, that $\mathcal{Q}F$ is finite at some point $X_0 \in \mathbb{R}^{n\times m}$, but it does not satisfy the p growth condition. Let D be the constant with which F satisfies the p growth assumption (4.1). Since $\mathcal{Q}F \leqslant F$, it follows that the p growth bound must fail for the negative part of $\mathcal{Q}F$. Thus, there exists a sequence of points $X_j \in \mathbb{R}^{n\times m}$ such that $\mathcal{Q}F(X_j) < -j(|X_j|^p + 1)$. From this and Corollary 4.1.7 we infer existence of a sequence of functions $\varphi_j \in C_c^\infty(Q_{1/2})$ such that

$$\fint_{Q_{1/2}} F(X_j + \nabla_{\mathbf{a}}\varphi_j)\,\mathrm{d}x < -j(|X_j|^p + 1).$$

Fix a cut-off function $\rho \in C_c^\infty(Q)$ with $\rho \equiv 1$ on $Q_{1/2}$. For each j let P_j be the **a**-polynomial given by $P_j(x) := \sum_{\langle\alpha,\mathbf{a}^{-1}\rangle=1}(X_j - X_0)x^\alpha$ and set $\psi_j := \rho P_j + \varphi_j \in C_c^\infty(Q)$. Then

$$\begin{aligned}\mathcal{Q}F(X_0) \leqslant{}& \liminf_{j\to\infty} |Q|^{-1}\int_Q F(X_0 + \nabla_{\mathbf{a}}\psi_j)\,\mathrm{d}x = \\ & \liminf_{j\to\infty} |Q|^{-1}\int_{Q_{1/2}} F(X_j + \nabla_{\mathbf{a}}\varphi_j)\,\mathrm{d}x + |Q|^{-1}\int_{Q\setminus Q_{1/2}} F(X_0 + \nabla_{\mathbf{a}}(\rho P_j))\,\mathrm{d}x \leqslant \\ & \liminf_{j\to\infty}\left(\frac{-1}{2}\right) j(|X_j|^p + 1) + \fint_Q D(|X_0 + \nabla_{\mathbf{a}}(\rho P_j)|^p + 1)\,\mathrm{d}x,\end{aligned}$$

where we have used the definition of ψ_j, the definition of X_j, and the p-growth bound on F respectively. We note that $|X_0 + \nabla_{\mathbf{a}}(\rho P_J)|^p \leqslant 2^{p-1}|X_0|^p + 2^{p-1}|\nabla_{\mathbf{a}}(\rho P_j)|^p$, and by construction $|\partial^\alpha P_j(x)| \leqslant C|X_j - X_0|$ with a uniform constant C for all $x \in Q$ and all α with $\langle\alpha,\mathbf{a}^{-1}\rangle \leqslant 1$, thus $|\nabla_{\mathbf{a}}(\rho P_j)|^p \leqslant C|X_j - X_0|^p$, since ρ is just a fixed C_c^∞ function. Plugging this into our inequality we get

$$\mathcal{Q}F(X_0) \leqslant \liminf_{j\to\infty}\left(\frac{-1}{2}\right) j(|X_j|^p + 1) + C(|X_0|^p + |X_j - X_0|^p + 1) = -\infty,$$

which yields a contradiction with the assumption that $\mathcal{Q}F(X_0)$ is finite. □

The above asserts that the quasiconvex envelope of an integrand of p-growth is either degenerate and equal to $-\infty$ everywhere, or it inherits the original integrand's growth bounds. Let us remark that in the context of classical quasiconvexity this is typically shown using rank-one convexity. Therefore, the usual proofs do not generalise to the mixed smoothness setting, hence we present a new argument that we believe is, in a sense, more straightforward.

Lemma 4.1.9. *For continuous integrands F satisfying the growth condition* (4.1) *the class $C_c^\infty(\Omega)$ in the formula for the quasiconvex envelope may be replaced by $\mathrm{W}_0^{\mathbf{a},p}(\Omega)$, i.e., for any open bounded domain $\Omega \subset \mathbb{R}^N$ we have*

$$\mathcal{Q}F(\cdot) = \inf_{\varphi \in \mathrm{W}_0^{\mathbf{a},p}(\Omega)} \fint_\Omega F(\cdot + \nabla_{\mathbf{a}}\varphi(x)) \,\mathrm{d}x.$$

Proof. Results immediately from the density of C_c^∞ in $\mathrm{W}_0^{\mathbf{a},p}$ and the fact that continuous integrands of p growth induce functionals continuous in the L^p norm. □

Lemma 4.1.10. *If F is a continuous integrand satisfying the p-growth condition* (4.1) *and such that its* **a***-quasiconvex envelope is not identically equal to $-\infty$, then the functional*

$$u \mapsto \fint_\Omega \mathcal{Q}F(\nabla_{\mathbf{a}}u) \,\mathrm{d}x$$

is sequentially upper semicontinuous along sequences u_j converging to a given u in the $\mathrm{W}^{\mathbf{a},p}$ norm and with $\nabla_{\mathbf{a}}u_j \to \nabla_{\mathbf{a}}u$ almost everywhere.

Proof. By Lemma 4.1.8 $\mathcal{Q}F$ satisfies the p-growth condition (4.1) as well (perhaps with a different constant C). Furthermore, since $X \mapsto \fint_\Omega F(X + \nabla_{\mathbf{a}}\varphi(x)) \,\mathrm{d}x$ is continuous for any $\varphi \in \mathrm{W}_0^{\mathbf{a},p}$ we see that $\mathcal{Q}F$ is a pointwise infimum of a family of continuous functions, thus $\mathcal{Q}F$ is upper semicontinuous. Since the functions $C(1 + |\nabla_{\mathbf{a}}u_j|^p) - \mathcal{Q}F(\nabla_{\mathbf{a}}u_j)$ are all non-negative, we get, by Fatou's lemma, that

$$\liminf_{j\to\infty} \fint_\Omega C(1 + |\nabla_{\mathbf{a}}u_j|^p) - \mathcal{Q}F(\nabla_{\mathbf{a}}u_j) \,\mathrm{d}x \geqslant \fint_\Omega \liminf_{j\to\infty} C(1 + |\nabla_{\mathbf{a}}u_j|^p) - \mathcal{Q}F(\nabla_{\mathbf{a}}u_j) \,\mathrm{d}x.$$

Rearranging and using strong L^p convergence of $\nabla_{\mathbf{a}}u_j$ yields

$$\limsup_{j\to\infty} \fint_\Omega \mathcal{Q}F(\nabla_{\mathbf{a}}u_j) \,\mathrm{d}x \leqslant \fint_\Omega \limsup_{j\to\infty} \mathcal{Q}F(\nabla_{\mathbf{a}}u_j) \,\mathrm{d}x.$$

Finally, $\nabla_{\mathbf{a}}u_j \to \nabla_{\mathbf{a}}u$ almost everywhere by assumption, and $\mathcal{Q}F$ is upper semicontinuous, thus $\limsup_{j\to\infty} \mathcal{Q}F(\nabla_{\mathbf{a}}u_j) \leqslant \mathcal{Q}F(\nabla_{\mathbf{a}}u)$ almost everywhere, which ends the proof. □

The following result justifies the omission of the underlying set Ω in our notation $\theta(t)$ by showing that this auxiliary function θ does not depend on the choice of Ω. Moreover, it shows that θ only depends on $\mathcal{Q}F$ rather than F itself, thus showing that F is L^q (mean) coercive if and only if $\mathcal{Q}F$ is.

Lemma 4.1.11 (see Lemma 3.1 in [37])**.** *Let $\omega, \Omega \subset \mathbb{R}^{n\times m}$ be non-empty bounded open subsets of $\mathbb{R}^N$, and define $\theta_q^\Omega(t)$ as above. Define also*

$$\theta^{qc}(t) := \inf\left\{ \fint_\omega \mathcal{Q}F(\nabla_{\mathbf{a}}\varphi)\,\mathrm{d}x \colon \varphi \in \mathrm{W}_0^{\mathbf{a},p}(\omega, \mathbb{R}^n), \fint_\omega |\nabla_{\mathbf{a}}\varphi|^q\,\mathrm{d}x \geqslant t \right\}.$$

Then $\theta_q^\Omega(t) = \theta^{qc}(t)$ for all $t \geqslant 0$.

Proof. The argument follows that of [37], but we have decided to present it in full, as there are technical details that need to be changed, for example the piecewise polynomial approximation that we use requires a countable partition of the domain, whereas the original paper uses a finite one. Moreover, due to the lack of rank one directions we cannot, in general, guarantee continuity of the quasiconvex evenlope $\mathcal{Q}F$ (see however Lemma 4.3.1), which introduces some technical difficulties.

The case $t = 0$ is just the classical fact that the quasiconvex envelope does not depend on the choice of the underlying domain, and is the content of Corollary 4.1.7. Another easy case is when $\mathcal{Q}F$ is identically equal to $-\infty$. Then one can, for example, choose two disjoint open subsets $\Omega_1, \Omega_2 \Subset \Omega$ and construct functions $\varphi_i \in C_c^\infty(\Omega_i)$ and use φ_1 to satisfy the restriction $\fint_{\Omega_i} |\nabla_{\mathbf{a}}\varphi_i|^q\,\mathrm{d}x \geqslant t$, whilst using φ_2 to make the integral $\int_{\Omega_2} F(\nabla_{\mathbf{a}}\varphi_2)\,\mathrm{d}x$ as small as one wishes, with the last point being possible thanks to the fact that $\mathcal{Q}F(0) = -\infty$. Then setting $\varphi := \varphi_1 + \varphi_2$ and testing on that we easily see that $\theta_q^\Omega(t) = -\infty$ for all t, and clearly the same holds for θ^{qc}. Thus, in the following we assume that $t > 0$ and that $\mathcal{Q}F > -\infty$.

Let us fix $t > 0$. For any $\varepsilon > 0$ we may find a function $\varphi \in C_c^\infty(\Omega)$ satisfying

$$\fint_\Omega |\nabla_{\mathbf{a}}\varphi|^q\,\mathrm{d}x \geqslant t \quad \text{and} \quad \theta^{\mathrm{qc}}(t) + \varepsilon > \fint_\Omega \mathcal{Q}F(\nabla_{\mathbf{a}}\varphi)\,\mathrm{d}x.$$

To see this, it is enough to observe that the existence of $\widetilde{\varphi} \in \mathrm{W}_0^{\mathbf{a},p}(\Omega)$ satisfying these two inequalities, the second with a smaller ε', is guaranteed by the definition of θ^{qc}. To get the same with $\varphi \in C_c^\infty(\Omega)$, it is enough to observe that $\widetilde{\varphi}$ may be approximated in $\mathrm{W}_0^{\mathbf{a},p}(\Omega)$ by a sequence $\varphi_j \in C_c^\infty(\Omega)$ with $\|\nabla_{\mathbf{a}}\varphi_j\|_{\mathrm{L}^q} \geqslant \|\nabla_{\mathbf{a}}\widetilde{\varphi}\|_{\mathrm{L}^q}$ for all j and $\nabla_{\mathbf{a}}\varphi_j \to \nabla_{\mathbf{a}}\widetilde{\varphi}$ almost everywhere. Then we simply use Lemma 4.1.10 and conclude.

From here it is easy to pass to a function $\psi \in \mathrm{W}_0^{\mathbf{a},p}(\Omega)$ satisfying the above with 2ε instead of ε in the second inequality and for which $\nabla_{\mathbf{a}}\psi$ is piecewise constant on a large part of Ω. We simply apply Proposition 2.5.4 to $\varphi \in C_c^\infty(\Omega)$ (preserving the boundary values) and obtain a sequence φ_j approximating φ in $\mathrm{W}^{\mathbf{a},p}$ and such that the measure of the complement of the set of boxes T, denoted τ_j, on which $\nabla_{\mathbf{a}}\varphi_j$ is constant goes to 0 as j goes to infinity. We may then rescale the sequence to satisfy the condition $\fint_\Omega |\nabla_{\mathbf{a}}\varphi_j|^q \,\mathrm{d}x \geqslant t$, pass to a subsequence for which the gradients converge almost everywhere, and finally use Lemma 4.1.10 again to deduce that elements φ_j for large enough j's satisfy the desired inequalities, still with an ε. Without loss of generality assume that this holds for all j. Since $\nabla_{\mathbf{a}}\varphi_j$ is strongly L^p convergent it is also p-equiintegrable, thus (as $\mathcal{Q}F$ satisfies the p-growth bound) $\mathcal{Q}F(\nabla_{\mathbf{a}}\varphi_j)$ and $F(\nabla_{\mathbf{a}}\varphi_j)$ are equiintegrable as well. Therefore, we may find a j_0 such that the measure of $\Omega \setminus \bigcup_{\tau_j} T$ is small enough so that

$$\left|\int_{\Omega\setminus\bigcup_{\tau_{j_0}} T} \mathcal{Q}F(\nabla_{\mathbf{a}}\varphi_j)\,\mathrm{d}x\right| < \varepsilon|\Omega| \quad \text{and} \quad \left|\int_{\Omega\setminus\bigcup_{\tau_{j_0}} T} F(\nabla_{\mathbf{a}}\varphi_j)\,\mathrm{d}x\right| < \varepsilon|\Omega|.$$

We set $\psi := \varphi_{j_0}$ and $\tau = \tau_{j_0}$.

By Corollary 4.1.7 we may, for each $T \in \tau$, find a function $\varphi_T \in C_c^\infty(T)$ satisfying

$$\fint_T F(\nabla_{\mathbf{a}}\psi + \nabla_{\mathbf{a}}\varphi_T)\,\mathrm{d}x < \mathcal{Q}F(\nabla_{\mathbf{a}}\psi) + \varepsilon. \tag{4.2}$$

We set $\varphi := \sum_{T\in\tau} \varphi_T$. Since the sum is finite this function is well defined and belongs to the class $C_c^\infty(\Omega)$. We then have

$$\begin{aligned}\theta^{\mathrm{qc}}(t) + 2\varepsilon &> \fint_\Omega \mathcal{Q}F(\nabla_{\mathbf{a}}\psi)\,\mathrm{d}x = |\Omega|^{-1}\left(\int_{\bigcup_\tau} \mathcal{Q}F(\nabla_{\mathbf{a}}\psi)\,\mathrm{d}x + \int_{\Omega\setminus\bigcup_\tau} \mathcal{Q}F(\nabla_{\mathbf{a}}\psi)\,\mathrm{d}x\right)\\ &> |\Omega|^{-1}\int_{\bigcup_\tau} F(\nabla_{\mathbf{a}}\psi + \nabla_{\mathbf{a}}\varphi)\,\mathrm{d}x - 2\varepsilon > \fint_\Omega F(\nabla_{\mathbf{a}}\psi + \nabla_{\mathbf{a}}\varphi)\,\mathrm{d}x - 3\varepsilon.\end{aligned}$$

Here the first inequality results from the construction of ψ above, the second one results from estimating the remainder integral $\left|\int_{\Omega\setminus\bigcup_\tau} \mathcal{Q}F(\nabla_{\mathbf{a}}\psi)\,\mathrm{d}x\right|$ and using (4.2) on $T \in \tau$, and the last one is just an estimate on $\left|\int_{\Omega\setminus\bigcup_\tau T} F(\nabla_{\mathbf{a}}\psi)\,\mathrm{d}x\right|$ paired with the fact that $\varphi \equiv 0$ on $\Omega \setminus \bigcup_\tau T$. Finally, since $\varphi \in C_c^\infty(\Omega)$, Jensen's inequality gives

$$\fint_\Omega |\nabla_{\mathbf{a}}\psi + \nabla_{\mathbf{a}}\varphi|^q\,\mathrm{d}x \geqslant \fint_\Omega |\nabla_{\mathbf{a}}\psi|^q\,\mathrm{d}x \geqslant t.$$

Since $\varepsilon > 0$ was arbitrary this already shows, that if $\omega = \Omega$ then $\theta^{\mathrm{qc}} = \theta_q^\Omega$. To pass to an arbitrary ω we need to realise the distribution of $\nabla_{\mathbf{a}}(\psi + \varphi)$ on ω, which follows easily from Lemma 4.1.6 — observe that this also shows that we preserve the moment restrictions. □

Lemma 4.1.12 (see Proposition 3.2 in [37])**.** *Let $F\colon \mathbb{R}^{n\times m} \to \mathbb{R}$ be a continuous integrand satisfying the growth condition* (4.1)*. Then its associated auxiliary function $\theta := \theta_q^\Omega\colon [0,\infty) \to \mathbb{R}\cup\{-\infty\}$ is convex.*

Proof. We omit the proof as our case only differs from the one in [37] by adding a scaling exponent. □

We are now ready to finish the proof of the main result of this section.

Proof of Theorem 4.1.3, part 2. We have already seen that a) and b) are equivalent, and so are c) and d). That c) implies a) is obvious. We will now show that a) implies d). If a) holds, i.e., if F is L^q coercive, then its auxiliary function θ satisfies $\lim_{t\to\infty}\varphi(t) = \infty$. This, paired with the fact that θ is convex (by Lemma 4.1.12), implies that there exist $c_1 > 0$ and $c_2 \in \mathbb{R}$ such that $\theta(t) \geqslant c_1 t + c_2$ for all $t \geqslant 0$. Thus, taking Ω to be any admissible domain and imposing zero boundary conditions, we may take $t = \|\nabla_{\mathbf{a}}\varphi\|^q_{\mathrm{L}^q(\Omega)}$ for any $\varphi \in \mathrm{W}^{\mathbf{a},p}_0(\Omega)$ and conclude that F is L^q mean coercive. Thus, we conclude that the conditions a) – d) are all equivalent.

The fact that c) implies e) and f) is immediate. For the other direction let us start with a simple case of $\Omega = Q$ — the unit cube, and boundary datum g such that $\nabla_{\mathbf{a}}g$ is some constant $X_0 \in \mathbb{R}^{n\times m}$. Let us assume that in this case all minimising sequences are bounded, we now wish to prove that F is L^q mean coercive at X_0 on Q. Consider the auxiliary function $\widetilde{\theta}$ of the translated integrand $\widetilde{F}(\cdot) := F(X_0 + \cdot)$, which clearly still satisfies the L^p growth bound. The assumption that all minimising sequences are bounded easily implies that the infimum of our variational problem, equal to $\widetilde{\theta}(0)$, cannot be $-\infty$. If it was, one could take an arbitrary (bounded) minimising sequence φ_j, rescale to be supported on a smaller cube, and then use the remaining room to make the $\mathrm{W}^{\mathbf{a},q}$ norm of the resulting sequence arbitrarily large without spoiling the minimising property, thanks to the p growth assumption on F.

Clearly $\widetilde{\theta}$ is non-decreasing, and convex as shown in Lemma 4.1.12. If $\widetilde{\theta}$ was bounded from above we would deduce existence of minimising sequences of arbitrarily large $\mathrm{W}^{\mathbf{a},q}$ norms, which would contradict our assumption. Thus, $\widetilde{\theta}$ is real-valued, unbounded, and convex, and so we deduce, as previously, that F is L^q mean coercive at z_0 on Q, which implies point d) of our theorem, and thus all the points a) - d).

Now let us take an arbitrary Ω and g and assume, for contradiction, that F is not mean coercive, but all minimising sequences are bounded. As previously, the infimum of the variational problem has to be a real number. Take any minimising sequence u_j and note that, due to Proposition 2.5.4, we may approximate any u_j with a sequence $u_j^\varepsilon \in \mathrm{W}^{\mathbf{a},p}_{u_j} = \mathrm{W}^{\mathbf{a},p}_{u}$, elements of which are piecewise polynomial on a finite family of disjoint open boxes, covering Ω up to a set of measure ε, and converge to u_j in the $\mathrm{W}^{\mathbf{a},p}$ norm. Since F is a continuous integrand of p growth, the map $u \mapsto \int_\Omega F(\nabla_{\mathbf{a}}u)$ is continuous in $\mathrm{W}^{\mathbf{a},p}$, thus extracting a diagonal subsequence we may assume that our (not renamed) minimising sequence u_j is such that for each j there exists a finite family $\{Q_i^j\}_i$ of pairwise disjoint boxes with $\left|\Omega \setminus \bigcup_i Q_i^j\right| < 1/j$ and with $\nabla_{\mathbf{a}} u_j$ constant on each Q_i^j.

By the previous step, since we assume that F is not L^q mean coercive we may, for any open box and any $\mathbf{a}$-polynomial boundary datum, construct an unbounded minimising sequence. Doing so on each Q_i^j with boundary datum u_j (which is indeed an $\mathbf{a}$-polynomial on Q_i^j) we may construct functions $\varphi_i^j \in \mathrm{W}_0^{\mathbf{a},p}(Q_i^j)$ such that

$$j\left|Q_i^j\right| < \int_{Q_i^j} \left|\nabla_{\mathbf{a}}(u_j + \varphi_i^j)\right|^q \,\mathrm{d}x,$$

and

$$\int_{Q_i^j} F(\nabla_{\mathbf{a}}(u_j + \varphi_i^j))\,\mathrm{d}x \leqslant \int_{Q_i^j} F(\nabla_{\mathbf{a}} u_j) + 1/j\,\mathrm{d}x.$$

Extending each φ_i^j by zero outside Q_i^j and letting $v_j := u_j + \sum_i \varphi_i^j$ yields a sequence $v_j \in \mathrm{W}_u^{\mathbf{a},p}(\Omega)$ that clearly still minimizes the functional, but is unbounded in $\mathrm{W}^{\mathbf{a},q}$, and thus we have a contradiction, which proves that f) implies c), thus the points a) through f) are equivalent.

That g) implies d) is easy. Assuming that $F - \delta|\cdot|^q$ is quasiconvex at some $X_0 \in \mathbb{R}^{n\times m}$ we may write

$$F(X_0) - \delta|X_0|^q \leqslant \fint_\Omega F(X_0 + \nabla_{\mathbf{a}}\varphi(x)) - \delta|X_0 + \nabla_{\mathbf{a}}\varphi(x)|\,\mathrm{d}x$$

for all $\varphi \in C_c^\infty(\Omega)$. This is exactly L^q mean coercivity on $\mathrm{W}_g^{\mathbf{a},p}(\Omega)$ with g a polynomial such that $\nabla_{\mathbf{a}} g = X_0$, $c_1 = \delta > 0$ and $c_2 = F(X_0) - \delta|X_0|^q$.

As the last step we prove that c) implies g). Assume that F is L^q mean coercive, take Ω to be any admissible domain, and let the boundary condition be $g \equiv 0$. If the corresponding coercivity constants are c_1 and c_2 then take any $\delta \in (0, c_1)$ and put $G(X) := F(X) - \delta|X|^q$ for $X \in \mathbb{R}^{n\times m}$. Clearly, the auxiliary function of G is bounded from below by $(c_1 - \delta)t + c_2$ and, as always, $G \geqslant \mathcal{Q}G$. From Lemma 4.1.11 we know that the auxiliary functions of G and $\mathcal{Q}G$ are the same, thus we may write

$$\fint_\Omega G(\nabla_{\mathbf{a}}\varphi)\,\mathrm{d}x \geqslant \fint_\Omega \mathcal{Q}G(\nabla_{\mathbf{a}}\varphi)\,\mathrm{d}x \geqslant (c_1 - \delta)\fint_\Omega |\nabla_{\mathbf{a}}\varphi|^q\,\mathrm{d}x + c_2$$

for all $\varphi \in \mathrm{W}_0^{\mathbf{a},p}(\Omega)$. Thanks to Corollary 4.1.7 we may find a sequence $\varphi_j \subset C_c^\infty(\Omega)$ for which

$$\mathcal{Q}G(0) \leqslant \fint_\Omega \mathcal{Q}G(\nabla_{\mathbf{a}}\varphi_j)\,\mathrm{d}x \leqslant \fint_\Omega G(\nabla_{\mathbf{a}}\varphi_j)\,\mathrm{d}x \searrow \mathcal{Q}G(0).$$

Since we have picked $\delta < c_1$ it is easy to conclude that G is still L^q mean coercive. Since $\fint_\Omega G(\nabla_{\mathbf{a}}\varphi_j)\,\mathrm{d}x$ is bounded we infer that the sequence $\nabla_{\mathbf{a}}\varphi_j$ is bounded in L^q. As in [37] we consider the following probability measures on $\mathbb{R}^{n\times m}$

$$\nu_j := (\nabla_{\mathbf{a}}\varphi_j)_\# \left(\frac{\mathcal{L}^N \restriction \Omega}{|\Omega|}\right)$$

and observe that they have uniformly bounded q-th moments. Thus, passing to a subsequence if necessary, we may assume that $\nu_j \overset{*}{\rightharpoonup} \nu$ in $C_0(\mathbb{R}^{n\times m})^*$, where ν is some probability measure on $\mathbb{R}^{n\times m}$ with finite qth moment. Setting $H(X) := G(X) - \mathcal{Q}G(X)$, which is a non-negative and lower semicontinuous function, we may use the portmanteau theorem to write

$$0 \leqslant \int_{\mathbb{R}^{n\times m}} H\,\mathrm{d}\nu \leqslant \liminf_{j\to\infty} \int_{\mathbb{R}^{n\times m}} H\,\mathrm{d}\nu_j = 0,$$

thus $H = 0$ on the support of ν. However, then $G = \mathcal{Q}G$ on the support of ν, i.e., $F - \delta|\cdot|^q$ is $\mathbf{a}$-quasiconvex on the support of ν, and since ν is a probability measure its support is non-empty, which ends the proof. □

4.2 Lower semicontinuity

We are now ready to move on to the central part of our existence theory, i.e., sequential weak lower semicontinuity of integral functionals acting on Sobolev spaces of mixed smoothness. The main results here are contained in Lemmas 4.2.6 and 4.2.7, which show that **a**-quasiconvexity is a sufficient and necessary condition for lower semicontinuity. To be precise, the exact notion of **a**-quasiconvexity that is relevant depends on the assumptions on the integrand and on the space on which one wishes to establish lower semicontinuity. Thus, we begin with the definitions of (closed) $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity and a short discussion of the relationship between the three notions. The proofs in this part are somehow standard — necessity is proven through a simple application of the Riemann-Lebesgue lemma, whereas for sufficiency we employ the Fundamental Theorem of Young Measures and the characterisation of Young measures generated by sequences of **a**-gradients. Thus, the bulk of the work is hidden in the technical results of Chapter 3. Let us also remark here that **a**-quasiconvexity being equivalent to lower semicontinuity of functionals does not say much about relaxation formulas. Indeed, these are much more delicate and significantly more difficult to prove, but we postpone further discussion until Section 4.3.

For the reader's convenience let us recall the quasiconvexity notion that we have defined in the previous section. We say that F is $\mathbf{a}$-quasiconvex if

$$F(X) \leqslant \inf_{\varphi \in C_c^\infty(Q)} \fint_Q F(X + \nabla_{\mathbf{a}}\varphi)\,\mathrm{d}x \text{ for all } X \in \mathbb{R}^{n\times m}.$$

Similarly to Ball and Murat in [13] we define the following:

Definition 4.2.1. *We say that a function* $F\colon \mathbb{R}^{n\times m} \to \mathbb{R}$ *is* $\mathrm{W}^{\mathbf{a},p}$*-quasiconvex if for every* $X \in \mathbb{R}^{n\times m}$ *one has*

$$F(X) \leqslant \inf_{u \in \mathrm{W}_0^{\mathbf{a},p}(Q;\mathbb{R}^n)} \fint_Q F(X + \nabla_{\mathbf{a}} u(x))\,\mathrm{d}x.$$

Observe that this definition differs from that of $\mathbf{a}$-quasiconvexity only by replacing the test space C_c^∞ by $\mathrm{W}_0^{\mathbf{a},p}$. In fact, for continuous integrands satisfying a p-growth bound the two are equivalent. In the classical case of first order gradients, this has been shown by Ball and Murat in [13], and the same proof works in our case. We present it below for the reader's convenience.

Lemma 4.2.2 (see [13])**.** *Suppose that* $F\colon \mathbb{R}^{n\times m} \to \mathbb{R}$ *is continuous and satisfies* $|F(X)| \leqslant C(1 + |X|^p)$ *for every* $X \in \mathbb{R}^{n\times m}$*. Then* F *is* $\mathrm{W}^{\mathbf{a},p}$*-quasiconvex if and only if it is* $\mathbf{a}$*-quasiconvex.*

Proof. The fact that $C_c^\infty \subset \mathrm{W}_0^{\mathbf{a},p}$ shows that $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity implies $\mathbf{a}$-quasiconvexity (and there are no extra assumptions needed).

For the other direction, fix an arbitrary $u \in \mathrm{W}_0^{\mathbf{a},p}(Q)$ and $X \in \mathbb{R}^{n\times m}$. Pick a sequence $u_j \in C_c^\infty(Q)$ with $u_j \to u$ in $\mathrm{W}_0^{\mathbf{a},p}(Q)$. In particular, this implies that $\nabla_{\mathbf{a}} u_j \to \nabla_{\mathbf{a}} u$ in L^p. By passing to a subsequence if necessary we may assume that $\nabla_{\mathbf{a}} u_j \to \nabla_{\mathbf{a}} u$ almost everywhere. Then, in particular, $F(X + \nabla_{\mathbf{a}} u_j) \to F(X + \nabla_{\mathbf{a}} u)$ almost everywhere. Since $C(1 + |X + \nabla_{\mathbf{a}} u_j|^p) - F(X + \nabla_{\mathbf{a}} u_j) \geqslant 0$ for all j we have, by Fatou's lemma,

$$\liminf_{j\to\infty} \fint_Q C(1+|X+\nabla_{\mathbf{a}} u_j|^p) - F(X+\nabla_{\mathbf{a}} u_j)\,\mathrm{d}x \geqslant \fint_Q C(1+|X+\nabla_{\mathbf{a}} u|^p) - F(X+\nabla_{\mathbf{a}} u)\,\mathrm{d}x.$$

Since $\nabla_{\mathbf{a}} u_j \to \nabla_{\mathbf{a}} u$ strongly in L^p we deduce that

$$\fint_Q F(X + \nabla_{\mathbf{a}} u)\,\mathrm{d}x \geqslant \liminf_{j\to\infty} \fint_Q F(X + \nabla_{\mathbf{a}} u_j)\,\mathrm{d}x \geqslant F(X),$$

where the last inequality follows from the $\mathbf{a}$-quasiconvexity assumption on F. Since $u \in \mathrm{W}_0^{\mathbf{a},p}(Q)$ was arbitrary this ends the proof. □

This and Lemma 3.5.3 immediately imply the following.

Corollary 4.2.3. *Suppose that* $F\colon \mathbb{R}^{n\times m} \to \mathbb{R}$ *is continuous and satisfies* $|F(X)| \leqslant C(1+|X|^p)$ *for every* $X \in \mathbb{R}^{n\times m}$*. Then* QF *is* $\mathrm{W}^{\mathbf{a},p}$*-quasiconvex.*

Remark 6. Let us note here that the growth assumption on F cannot, in general, be removed from the statement of Lemma 4.2.2. In other words, $\mathrm{W}^{\mathbf{a},p}$ and $\mathrm{W}^{\mathbf{a},q}$ gradient quasiconvexity are different notions in general. The previously mentioned paper [13] contains an example of a function that is $\mathrm{W}^{1,p}$ quasiconvex only for sufficiently large exponents p, thus showing the importance of the relevant test space.

Finally, for future use, we introduce a third notion (analogous to the one introduced by Pedregal in [140] and later discussed by Kristensen in [102]) that will be particularly useful for dealing with extended-real valued integrands.

Definition 4.2.4. *We say that a function* $F\colon \mathbb{R}^{n\times m} \to (-\infty, \infty]$ *is closed* $\mathrm{W}^{\mathbf{a},p}$*-quasiconvex if* F *is lower semicontinuous and Jensen's inequality holds for* F *and every homogeneous oscillation* $\mathrm{W}^{\mathbf{a},p}$*-gradient Young measure, i.e.,*

$$F(X) \leqslant \inf_{\nu\in\mathbb{H}^p_X} \int_{\mathbb{R}^{n\times m}} F(W)\,\mathrm{d}\nu(W),$$

where $\mathbb{H}^p_X$ *is the set of all homogenous* $\mathrm{W}^{\mathbf{a},p}$*-gradient Young measures with mean* X*.*

Note that, as in Theorem 3.7.1, we may note that replacing $\nu \in \mathbb{H}^p_X$ with $\delta_X * \mu$ for $\mu \in \mathbb{H}^p_0$ allows us to rewrite the inequality in the definition of closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity as

$$F(X) \leqslant \inf_{\mu\in\mathbb{H}^p_0} \int_{\mathbb{R}^{n\times m}} F(X+W)\,\mathrm{d}\mu(W),$$

for all $X \in \mathbb{R}^{n\times m}$.

Before we proceed, we note the following lemma which, in the case of classical gradients, was first proven by Ball and Zhang in [14]:

Lemma 4.2.5. *Suppose that* $F\colon \mathbb{R}^{n\times m} \to \mathbb{R}$ *is a continuous integrand satisfying* $|F(X)| \leqslant C(|X|^p+1)$ *for some constant* C *and for all* $X \in \mathbb{R}^{n\times m}$*. Then* F *is* $\mathbf{a}$*-quasiconvex if and only if it is* $\mathrm{W}^{\mathbf{a},p}$*-quasiconvex and if and only if it is closed* $\mathrm{W}^{\mathbf{a},p}$*-quasiconvex.*

Proof. In the case of continuous integrands of p growth the equivalence between $\mathbf{a}$-quasiconvexity and $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity has already been established in Lemma 4.2.2. To prove the equivalence of closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity with these notions let us first observe that clearly closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity implies $\mathbf{a}$-quasiconvexity, as for every $\varphi \in C_c^\infty(Q)$ the measure given by $(\nabla_{\mathbf{a}}\varphi)_\# \frac{\mathcal{L}^N \llcorner Q}{|Q|}$ is a homogeneous $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure (see Corollary 3.6.2) — note that this part does not use any assumptions on F — indeed, closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity is the strongest of the three notions.

For the other direction fix a point $X \in \mathbb{R}^{n\times m}$ and an arbitrary $\mu \in \mathbb{H}_0^p$ and let $\varphi_j \in \mathrm{W}_0^{\mathbf{a},p}(Q)$ be such that the sequence $\{\nabla_{\mathbf{a}}\varphi_j\}$ is p-equintegrable and generates μ, which is possible thanks to Corollary 3.2.4. Due to the growth conditions on F we know that the family $\{F(X+\nabla_{\mathbf{a}}\varphi_j)\}$ is equiintegrable as well, so that Theorem 3.1.1 shows that

$$\int_{\mathbb{R}^{n\times m}} F(X+W)\,\mathrm{d}\mu(W) = \lim_{j\to\infty} \fint_Q F(X+\nabla_{\mathbf{a}}\varphi_j(x))\,\mathrm{d}x \geqslant F(X),$$

where the last inequality is due to $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity of F. Since X and ν were arbitrary, this ends the proof. Observe that we do not need to assume that F is bounded from below to apply Theorem 3.1.1, as we may simply consider its positive and negative parts separately. □

In the following we establish an initial result that identifies closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity as a sufficient condition for lower semicontinuity of integral functionals in the mixed smoothness setting.

Lemma 4.2.6. *Let Ω be an open and bounded domain. Suppose that $F\colon \mathbb{R}^{n\times m} \to [0,\infty]$ is a closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvex integrand. Then the functional*

$$\mathrm{I}(u) := \int_\Omega F(\nabla_{\mathbf{a}}u)\,\mathrm{d}x$$

is sequentially weakly lower semicontinuous on $\mathrm{W}^{\mathbf{a},p}(\Omega)$.

Proof. Fix an arbitrary $u \in \mathrm{W}^{\mathbf{a},p}(\Omega)$ and a sequence $u_j \rightharpoonup u$ in $\mathrm{W}^{\mathbf{a},p}(\Omega)$. We need to show that $\liminf \mathrm{I}(u_j) \geqslant \mathrm{I}(u)$. Since u_j is weakly convergent in $\mathrm{W}^{\mathbf{a},p}(\Omega)$ it is also bounded in that space. Passing to a subsequence if necessary, we may assume that the lim inf is a true limit and, passing to a further subsequence, that $\nabla_{\mathbf{a}}u_j$ generates some $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure $\nu = \{\nu_x\}_{x\in\Omega}$. Observe that, since $p \in (1,\infty)$, the barycentre of ν_x is $\nabla_{\mathbf{a}}u(x)$ for almost every x in Ω. By Theorem 3.1.1 we know that

$$\lim_{j\to\infty} \mathrm{I}(u_j) = \lim_{j\to\infty} \int_\Omega F(\nabla_{\mathbf{a}}u_j)\,\mathrm{d}x \geqslant \int_\Omega \int_{\mathbb{R}^{n\times m}} F\,\mathrm{d}\nu_x\,\mathrm{d}x.$$

Corollary 3.4.3 tells us that, for Lebesgue almost every x, the measure ν_x is a homogeneous $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure. Hence F, as a closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvex function, satisfies Jensen's inequality when tested with ν_x for almost every x, which ends the proof of this part, as we have

$$\lim_{j\to\infty} \mathrm{I}(u_j) \geqslant \int_\Omega \int_{\mathbb{R}^{n\times m}} F \,\mathrm{d}\nu_x \,\mathrm{d}x \geqslant \int_\Omega F(\overline{\nu_x}) \,\mathrm{d}x = \int_\Omega F(\nabla_{\mathbf{a}} u(x)) \,\mathrm{d}x = \mathrm{I}(u).$$

□

Remark 7. Observe that, since the proof above is based on localisation of Young measures, as an immediate corollary we get that, under the assumptions of the above lemma, the functional

$$\mathrm{W}^{\mathbf{a},p}(\Omega) \ni u \mapsto \int_A F(\nabla_{\mathbf{a}} u) \,\mathrm{d}x$$

is sequentially weakly lower semicontinuous on $\mathrm{W}^{\mathbf{a},p}(\Omega)$ for any measurable subset $A \subset \Omega$.

Lemma 4.2.7. *Let Ω be a bounded open domain satisfying the weak* **a***-horn condition. Suppose that $F\colon \mathbb{R}^{n\times m} \to (-\infty, \infty]$ is a measurable integrand and that its associated functional $\mathrm{I}(u) := \int_\Omega F(\nabla_{\mathbf{a}} u) \,\mathrm{d}x$ is sequentially weakly lower semicontinuous on $\mathrm{W}^{\mathbf{a},p}(\Omega)$. Then F is $\mathrm{W}^{\mathbf{a},p}$-quasiconvex.*

Proof. The proof is somehow standard in the theory of quasiconvexity, but we provide an outline for the sake of completeness. Fix an arbitrary $\varphi \in \mathrm{W}^{\mathbf{a},p}_0(Q)$ and a $X \in \mathbb{R}^{n\times m}$. Let u be an **a**-polynomial with $\nabla_{\mathbf{a}} u \equiv X$. For an arbitrary $k \in \mathbb{N}$ we may cover Ω, up to a set of measure zero, with a countable family of anisotropic boxes of the form $\{x_i^j + r_i^j \odot Q\}_i$, with $r_i^j < 1/j$ for all j, i. Let $\varphi_i^j(x) := (r_i^j)\varphi((r_i^j)^{-1} \odot (x - x^j))$ and set $\varphi_j := \sum_{i=1}^\infty \varphi_i^j$. Then φ_j converges weakly to 0 in $\mathrm{W}^{\mathbf{a},p}(\Omega)$ (see Lemma 3.6.1), thus $u + \varphi_j \rightharpoonup u$ in $\mathrm{W}^{\mathbf{a},p}(\Omega)$, so that by our assumption

$$\liminf_{j\to\infty} \int_\Omega F(\nabla_{\mathbf{a}}(u + \varphi_j) \,\mathrm{d}x = \liminf_{j\to\infty} \int_\Omega F(X + \nabla_{\mathbf{a}} \varphi_j) \,\mathrm{d}x \geqslant \int_\Omega F(\nabla_{\mathbf{a}} u) \,\mathrm{d}x = |\Omega| F(X).$$

However, for every j we have, by a change variables,

$$\int_\Omega F(X + \nabla_{\mathbf{a}} \varphi_j) \,\mathrm{d}x = |\Omega| \fint_Q F(X + \nabla_{\mathbf{a}} \varphi) \,\mathrm{d}x.$$

which ends the proof. □

As an immediate corollary of Lemmas 4.2.5, 4.2.6, and 4.2.7 we obtain the following result:

Theorem 4.2.8. *Let Ω be a bounded open domain satisfying the weak* $\mathbf{a}$*-horn condition. Suppose that $F\colon \mathbb{R}^{n\times m} \to [0,\infty)$ is a continuous integrand satisfying the p-growth condition $|F(X)| \leqslant C(|X|^p + 1)$ for some constant C and all $X \in \mathbb{R}^{n\times m}$. Then the functional*

$$\mathrm{I}(u) := \int_\Omega F(\nabla_{\mathbf{a}} u)\,\mathrm{d}x$$

is sequentially weakly lower semicontinuous on $\mathrm{W}^{\mathbf{a},p}$ *if and only if F is* $\mathrm{W}^{\mathbf{a},p}$*-quasiconvex.*

4.3 Relaxation

The results of the previous subsection identify $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity as an equivalent condition for lower semicontinuity of integral functionals in the mixed smoothness framework. The question that remains to be answered is what happens when the functional lacks lower semicontinuity? As discussed in the introduction, the general idea in calculus of variations is to approximate the initial functional by some sort of a relaxed version of it.

The relaxed functional of interest is typically defined as the sequentially weakly lower semicontinuous envelope of the original functional, i.e., through

$$\bar{\mathrm{I}}(u) := \inf_{u_j \rightharpoonup u} \left\{ \liminf_{j\to\infty} \mathrm{I}(u_j) \right\},$$

where the infimum is taken over all sequences u_j converging to u in the appropriate sense — here we require that $u_j \rightharpoonup u$ in $\mathrm{W}^{\mathbf{a},p}(\Omega)$. With this definition the relaxed functional $\bar{\mathrm{I}}$ is sequentially lower semicontinuous with respect to weak convergence in $\mathrm{W}^{\mathbf{a},p}(\Omega)$. The aim of this section is to show that, under appropriate assumptions, the relaxation is again an integral functional with integrand given by $\mathbf{a}$-quasiconvexification of the original one which, in the classical setting of first order gradients and under (p,p) growth conditions on the integrand, was first done by Dacorogna in [38]. The main results of this section are contained in Theorems 4.3.2 and 4.3.13, which deal with the p-growth case and the extended-real valued case, respectively.

4.3.1 The p-growth case

We begin with the case of integrands of p-growth. Before we move on to relaxation, let us prove the following important fact:

Lemma 4.3.1. *Suppose that* $F\colon \mathbb{R}^{n\times m} \to [0,\infty)$ *is a continuous and* L^p *coercive integrand with* $F(X) \leqslant C(|X|^p + 1)$*. Assume further that* F *is locally Lipschitz with*

$$|F(X) - F(W)| \leqslant D(1 + |X|^{p-1} + |W|^{p-1})|X - W|, \tag{4.3}$$

for all $X, W \in \mathbb{R}^{n\times m}$*. Then* $\mathcal{Q}F$ *is continuous.*

Proof. For a continuous integrand of p-growth the mapping $X \mapsto \fint_Q F(X + \nabla_{\mathbf{a}}\varphi)\,\mathrm{d}x$ is continuous for any fixed $\varphi \in C_c^\infty(Q)$. Thus, by the Dacorogna formula (3.10), we immediately see that the quasiconvex envelope is upper semicontinuous, as it is a pointwise infimum of a family of continuous functions.

Before we proceed further, we need to take a closer look at the coercivity assumption. By Theorem 4.1.3 we know that the functional induced by F is L^p mean coercive with any boundary datum. However, the constants describing the coercivity may depend on the datum — if this was not the case then F would have to satisfy pointwise coercivity bounds of the form $F(X) \geqslant C|X|^p - C^{-1}$, which we do not assume. Nevertheless, in the presence of the additional assumption on local Lipschitz continuity of F, it may be shown that the constants may be chosen uniformly on compact sets. To prove that, let us take X_0 to be a point at which $F - c|\cdot|^p$, with some positive constant c, is $\mathbf{a}$-quasiconvex — such a point exists by Theorem 4.1.3. For an arbitrary $X \in \mathbb{R}^{n\times m}$ and any $\varphi \in C_c^\infty(Q)$ we may, using (4.3), write

$$\fint_Q |F(X + \nabla_{\mathbf{a}}\varphi) - F(X_0 + \nabla_{\mathbf{a}}\varphi)|\,\mathrm{d}x \leqslant$$

$$D\fint_Q \left(1 + |X + \nabla_{\mathbf{a}}\varphi|^{p-1} + |X_0 + \nabla_{\mathbf{a}}\varphi|^{p-1}\right)|X - X_0|\,\mathrm{d}x.$$

Rearranging we get

$$\fint_Q |F(X + \nabla_{\mathbf{a}}\varphi) - F(X_0 + \nabla_{\mathbf{a}}\varphi)|\,\mathrm{d}x \leqslant c_1 + c_2\fint_Q |\nabla_{\mathbf{a}}\varphi|^{p-1}\,\mathrm{d}x,$$

where the constants c_i depend on X but may be chosen uniformly on compact sets. These constants also depend on D and X_0, but, since the integrand is fixed, this dependence is not important here. Rearranging and using the strict $\mathbf{a}$-quasiconvexity at X_0 given by Theorem 4.1.3 we may now write

$$\fint_Q F(X + \nabla_{\mathbf{a}}\varphi)\,\mathrm{d}x \geqslant c_0\fint_Q |\nabla_{\mathbf{a}}\varphi|^p\,\mathrm{d}x - c_1 - c_2\fint_Q |\nabla_{\mathbf{a}}\varphi|^{p-1}\,\mathrm{d}x,$$

with a different constant c_1. Using weighted Young's inequality we may finally write

$$\fint_Q F(X + \nabla_{\mathbf{a}}\varphi)\,\mathrm{d}x \geqslant \frac{c_0}{2}\fint_Q |\nabla_{\mathbf{a}}\varphi|^p\,\mathrm{d}x - c_1,$$

where, again, the constant c_1 has changed, but it is still independent of φ and may be chosen locally uniformly in X. Thus, we have shown that for any compact set K there exists a constant $c > 0$ such that

$$\fint_Q F(X + \nabla_{\mathbf{a}}\varphi)\,\mathrm{d}x \geqslant c\fint_Q |\nabla_{\mathbf{a}}\varphi|^p\,\mathrm{d}x - c^{-1}, \tag{4.4}$$

for all $X \in K$ and all $\varphi \in C_c^\infty(Q)$.

It remains to show that $\mathcal{Q}F$ is lower semicontinuous. To that end, take an arbitrary point $X \in \mathbb{R}^{n\times m}$ and a sequence $X_j \to X$. Assume, without loss of generality, that $\liminf_{j\to\infty} \mathcal{Q}F(X_j) = \lim_{j\to\infty} \mathcal{Q}F(X_j) < \infty$. For every j we may, by the Dacorogna formula (3.10) find a function $\varphi_j \in C_c^\infty(Q)$ such that

$$\mathcal{Q}F(X_j) > \fint_Q F(X_j + \nabla_{\mathbf{a}}\varphi_j)\,\mathrm{d}x.$$

Since X_j is a convergent sequence, it is contained in a compact set. Similarly, the sequence $\mathcal{Q}F(X_j)$ is bounded, thus, based on what we have just proven in (4.4), we deduce that the family $\{\nabla_{\mathbf{a}}\varphi_j\}$ is bounded in L^p. Similarly to before, using the Lipschitz bound (4.3), we may write

$$\begin{aligned}\fint_Q F(X + \nabla_{\mathbf{a}}\varphi_j)\,\mathrm{d}x \leqslant &\fint_Q F(X_j + \nabla_{\mathbf{a}}\varphi_j)\,\mathrm{d}x \\ &+ D\fint_Q \left(1 + |X|^{p-1} + |X_j|^{p-1} + |\nabla_{\mathbf{a}}\varphi_j|^{p-1}\right)|X - X_j|\,\mathrm{d}x.\end{aligned} \tag{4.5}$$

Now, $|X - X_j|$ converges to 0 as $j \to \infty$, and the integral

$$\fint_Q \left(1 + |X|^{p-1} + |X_j|^{p-1} + |\nabla_{\mathbf{a}}\varphi_j|^{p-1}\right)\,\mathrm{d}x$$

is bounded, hence the last term in (4.5) goes to 0 with j. Thus

$$\begin{aligned}\mathcal{Q}F(X) &\leqslant \liminf_{j\to\infty}\fint_Q F(X + \nabla_{\mathbf{a}}\varphi_j)\,\mathrm{d}x \\ &\leqslant \liminf_{j\to\infty}\fint_Q F(X_j + \nabla_{\mathbf{a}}\varphi_j)\,\mathrm{d}x = \liminf_{j\to\infty}\mathcal{Q}F(X_j),\end{aligned}$$

which ends the proof.

□

Let us remark that the above could be easily generalised to the $\mathcal{A}$-free setting of compensated compactness due to Murat and Tartar (see [135], [136] by Murat and [164], [165], [168] by Tartar). It is known that in general, when the characteristic cone of $\mathcal{A}$ does not span the entire space, $\mathcal{A}$-quasiconvex envelopes of smooth functions need not be continuous (see, for example, Remark 3.5 in [67] by Fonseca and Müller). Our proof shows that this is not a problem if one restricts to coercive integrands satisfying the bound (4.3). However, in the context of $\mathcal{A}$-quasiconvexity, this may also be resolved in a different way. In a recent collaboration with Raiţă (see [146]), we argue that it is natural to restrict the space of definition of the integrand to the span of the characteristic cone of the operator in question. However, this approach cannot be applied in the mixed smoothness case, and thus we do not elaborate on it further.

Our main result is the following:

Theorem 4.3.2. *Let Ω be a bounded open domain satisfying the weak* **a***-horn condition. Supppose that $F\colon \mathbb{R}^{n\times m} \to [0,\infty)$ is a continuous and L^p coercive integrand with $F(X) \leqslant C(|X|^p+1)$. Assume furthermore that F satisfies*

$$F(X) \geqslant D|X|^p - D^{-1} \tag{4.6}$$

or

$$|F(X) - F(W)| \leqslant D(1+|X|^{p-1}+|W|^{p-1})|X-W|, \tag{4.7}$$

for all $X, W \in \mathbb{R}^{n\times m}$. Then the sequentially weakly lower semicontinuous envelope of the functional I_F is given by

$$\bar{\mathrm{I}}_F(u) := \inf_{u_j \rightharpoonup u} \left\{ \liminf_{j\to\infty} \mathrm{I}_F(u_j) \right\} = \int_\Omega \mathcal{Q}F(\nabla_{\mathbf{a}} u(x))\,\mathrm{d}x = \mathrm{I}_{\mathcal{Q}F}(u),$$

where the infimum is taken over all sequences u_j converging to u weakly in $\mathrm{W}^{\mathbf{a},p}(\Omega)$.

Proof. When F satisfies the first of the two alternative conditions we have provided, i.e., when F is of p-growth from below as well as from above, this result is a corollary of a more general relaxation result (for extended real-valued integrands) that we will demonstrate next, thus we postpone this part of the proof.

Under the assumption that F is locally Lipschitz, we have shown in Lemma 4.3.1, that $\mathcal{Q}F$ is continuous. We also know, from Lemma 4.1.8, that $\mathcal{Q}F$ satisfies the same p-growth assumption as F. Lemma 3.5.3 tells us that $\mathcal{Q}F$ is **a**-quasiconvex, which in view of the continuity and growth bounds implies, by Lemma 4.2.5, that it is $\mathrm{W}^{\mathbf{a},p}$-quasiconvex. Finally, Lemma 4.2.6 shows that the functional induced by $\mathcal{Q}F$ is indeed sequentially weakly lower semicontinuous on $\mathrm{W}^{\mathbf{a},p}(\Omega)$. This translates to

$$\bar{\mathrm{I}}_F(u) \geqslant \mathrm{I}_{\mathcal{Q}F}(u),$$

and so it remains to prove the reverse inequality.

Proceeding similarly to a proof in Dacorogna's book [40] (see Theorem 9.1 therein) let us start with the simple case of $\Omega = Q$ and u with $\nabla_{\mathbf{a}} u = X$ for some constant $X \in \mathbb{R}^{n\times m}$. By the Dacorogna formula (3.10) there exists a sequence $\varphi_j \in C_c^\infty(Q)$ with

$$\int_Q F(\nabla_{\mathbf{a}} u + \nabla_{\mathbf{a}} \varphi_j)\,\mathrm{d}x \to \int_Q \mathcal{Q}F(\nabla_{\mathbf{a}} u)\,\mathrm{d}x.$$

It only remains to show that the sequence φ_j may be chosen in such a way as to satisfy $\varphi_j \rightharpoonup 0$ in $\mathrm{W}^{\mathbf{a},p}(Q)$. The argument here is essentially a simpler version of the one in the proof of Lemma 3.6.1. For a fixed j extend φ_j periodically and consider the sequence

$$\varphi_j^k(x) := r_k^{-1} \varphi_j(r_k \odot x),$$

with $r_k := k^{a_1 \cdot \ldots \cdot a_N}$. Then $\nabla_{\mathbf{a}} \varphi_j^k$ preserves the integral, i.e.,

$$\int_Q F(\nabla_{\mathbf{a}} u + \nabla_{\mathbf{a}} \varphi_j^k)\,\mathrm{d}x = \int_Q F(\nabla_{\mathbf{a}} u + \nabla_{\mathbf{a}} \varphi_j)\,\mathrm{d}x,$$

for every k, and $\varphi_j^k \rightharpoonup 0$ in $\mathrm{W}^{\mathbf{a},p}(Q)$. Thus, a diagonal extraction argument ends the proof in this basic case.

For the general case let us fix an arbitrary function $u \in \mathrm{W}^{\mathbf{a},p}(\Omega)$. Using Proposition 2.5.4 we may find a sequence $v_j \in \mathrm{W}^{\mathbf{a},p}_u(\Omega)$ with $v_j \to u$ in $\mathrm{W}^{\mathbf{a},p}(\Omega)$ and such that for each j there exists a finite family of anisotropic boxes $\{Q_i^j\}_i$ such that $\nabla_{\mathbf{a}} v_j$ is constant on each Q_i^j and $\left|\Omega \setminus \bigcup_i Q_i^j\right| \to 0$ as $j \to \infty$. We know, from Lemma 4.3.1, that $\mathcal{Q}F$ is continuous, and from Lemma 4.1.8 that it satisfies the p-growth bound. Thus

$$\int_\Omega \mathcal{Q}F(\nabla_{\mathbf{a}} v_j)\,\mathrm{d}x \to \int_\Omega \mathcal{Q}F(\nabla_{\mathbf{a}} u)\,\mathrm{d}x.$$

Since v_j converges to u strongly in $\mathrm{W}^{\mathbf{a},p}$ the sequence $\nabla_{\mathbf{a}} v_j$ is p-equiintegrable. By assumption and Lemma 4.1.8 we know that both F and $\mathcal{Q}F$ satisfy the p-growth bound from above, so that

$$\int_{\Omega\setminus\bigcup_i Q_i^j} F(\nabla_{\mathbf{a}} v_j)\,\mathrm{d}x \to 0 \text{ and } \int_{\Omega\setminus\bigcup_i Q_i^j} \mathcal{Q}F(\nabla_{\mathbf{a}} v_j)\,\mathrm{d}x \to 0.$$

In particular,

$$\int_{\bigcup_i Q_i^j} \mathcal{Q}F(\nabla_{\mathbf{a}} v_j)\,\mathrm{d}x + \int_{\Omega\setminus\bigcup_i Q_i^j} F(\nabla_{\mathbf{a}} v_j)\,\mathrm{d}x \to \int_\Omega \mathcal{Q}F(\nabla_{\mathbf{a}} u)\,\mathrm{d}x.$$

Using the previous step we may, for any fixed Q_i^j, find a sequence $\varphi^j_{i,k} \in C_c^\infty(Q_i^j)$ with $\varphi^j_{i,k} \rightharpoonup 0$ in $\mathrm{W}^{\mathbf{a},p}(\Omega)$ as $k \to \infty$ and such that

$$\int_{Q_i^j} F(\nabla_{\mathbf{a}} v_j + \nabla_{\mathbf{a}} \varphi^j_{i,k})\,\mathrm{d}x \to \int_{Q_i^j} \mathcal{Q}F(\nabla_{\mathbf{a}} v_j)\,\mathrm{d}x \text{ as } k \to \infty.$$

Letting $\varphi^j_k := \sum_i \varphi^j i, k \in C_c^\infty(\Omega)$, we have $\varphi^j_k \rightharpoonup 0$ in $\mathrm{W}^{\mathbf{a},p}(\Omega)$ as $k \to \infty$ and

$$\int_{\bigcup_i Q_i^j} F(\nabla_{\mathbf{a}} v_j + \nabla_{\mathbf{a}} \varphi^j_k)\,\mathrm{d}x \to \int_{\bigcup_i Q_i^j} \mathcal{Q}F(\nabla_{\mathbf{a}} v_j)\,\mathrm{d}x.$$

A standard diagonal extraction argument allows us to construct a, non-relabelled, diagonal sequence $\varphi_j \in C_c^\infty(\Omega)$ with $\varphi_j \rightharpoonup 0$ in $\mathrm{W}^{\mathbf{a},p}(\Omega)$ and such that

$$\int_\Omega F(\nabla_{\mathbf{a}} v_j + \nabla_{\mathbf{a}} \varphi_j)\,\mathrm{d}x \to \int_\Omega \mathcal{Q}F(\nabla_{\mathbf{a}} u)\,\mathrm{d}x.$$

Therefore, setting $u_j := v_j + \varphi_j$ shows that

$$\liminf_{j\to\infty} \mathrm{I}_F(u_j) \leqslant \mathrm{I}_{\mathcal{Q}F}(u),$$

and thus ends the proof. □

Observe that we only use the locally Lipschitz assumption on F in order to deduce continuity of $\mathcal{Q}F$ from Lemma 4.3.1. If continuity can be ensured in a different way then we can dispense with the assumption (4.7). In the classical setting of first order gradients it is known that quasiconvex functions are convex along rank-one directions, and these span the entire space, so that one may deduce continuity from directional convexity. We have already mentioned that, in general, there is no good analogue of rank-one convexity in the mixed smoothness setting. However, if $\mathbf{a}$ is such that all multiindices α with $\langle \alpha, \mathbf{a}^{-1}\rangle = 1$ are of the same parity, this has been resolved by Kazaniecki, Stolyarov, and Wojciechowski in [93]. Under this assumption they have shown that $\mathbf{a}$-quasiconvex functions are convex along directions of the form

$$\sum_{\langle \alpha, \mathbf{a}^{-1}\rangle = 1} i^{|\alpha|+|\alpha_0|} x^\alpha b^i e_{\alpha,i}. \tag{4.8}$$

Here $x \in \mathbb{R}^N$ and $b \in \mathbb{R}^n$ are arbitrary vectors, $i^2 = -1$, $e_{\alpha,i}$ is the canonical basis of $\mathbb{R}^{n\times m}$, and α_0 is an arbitrary multiindex on the hyperplane of homogeneity, i.e., with $\langle \alpha_0, \mathbf{a}^{-1}\rangle = 1$. When all the multiindices have the same parity, the coefficients in (4.8) are real and vectors of this form span $\mathbb{R}^{n\times m}$, thus continuity follows from directional convexity. Thus, we obtain the following:

Corollary 4.3.3. *If $\mathbf{a}^{-1}$ is such that all the multiindices α with $\langle \alpha, \mathbf{a}^{-1} \rangle = 1$ have orders of the same parity then the conclusion of Theorem 4.3.13 holds without the additional assumptions* (4.6) *or* (4.7)*.*

If, on the other hand, the parities of $|\alpha|$ do not match, then the coefficients in (4.8) complexify, and we do not get any directional convexity. It is still possible to use the calculation leading to determining the form of the vectors in (4.8) to show that **a**-quasiconvex functions are (pluri)subharmonic in a certain sense (see the discussion in [93]), but we have not yet been able to use that to strengthen our relaxation results. We believe, however, that this should be studied further and intend to do so in future work.

4.3.2 Closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvex envelope

The p-growth assumption on the integrand was crucial in the results of the previous subsections. Indeed, when F is not of p-growth the notions of $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity and closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity need not coincide, and this is precisely the reason for introducing this stricter notion of closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvexity. An explicit example, in the isotropic setting of first order gradients, of an integrand that is quasiconvex but not closed quasiconvex is given in Example 1.3 in [102].

The first issue we run into is that, for an extended real-valued integrand, the formula (3.10) need not yield a closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvex function. The purpose of this subsection is to establish a formula that does. We start with the natural definition of the closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvex envelope.

Definition 4.3.4. *For a measurable function $F \colon \mathbb{R}^{n\times m} \to (-\infty, \infty]$ we define its closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvex envelope by*

$$\overline{F}(X) := \sup\{G(X) \colon G \leqslant F,\ G \textit{ is closed } \mathbf{a}\text{-}p \textit{ quasiconvex}\}.$$

Our goal is the following:

Proposition 4.3.5. *For any $p \in (1, \infty)$, the closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvex envelope of a lower semicontinuous function $F \colon \mathbb{R}^{n\times m} \to [0, \infty]$ satisfying the growth condition $F(X) \geqslant c|X|^p$ for some constant $c > 0$ is given by*

$$\overline{F}(X) = \inf_{\nu \in \mathbb{H}^p_0} \langle F(\cdot + X), \nu \rangle = \inf_{\nu \in \mathbb{H}^p_X} \langle F, \nu \rangle.$$

Moreover, the function $\overline{F}$ is indeed closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvex.

The reason we put the lower growth assumption in this result is that in a number of places in the proof we will use an argument that selects, at each $X \in \mathbb{R}^{n\times m}$ a measure $\nu_X \in \mathbb{H}^p_X$ which (nearly) achieves the infimum in question, i.e., $\langle F, \nu_X\rangle = \overline{F}(X)$. The main idea of the proof is that then, with another fixed $\nu \in \mathbb{H}^p_0$ the measure μ defined by $\mathrm{d}\mu := \mathrm{d}\nu_X\,\mathrm{d}\nu$ is again a homogeneous $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure. To prove that we need to ensure that it has a finite p-th moment, which is where coercivity comes into the picture. For now we do not know whether it is possible to relax this assumption, nevertheless it is in line with the relaxation result we prove next. As in the case of integrands of p-growth, to prove a relaxation formula we need some sort of a coercivity assumption on the integrand. If we were to relax the pointwise coercivity to simply L^p or L^p-mean coercivity we would need a Lipschitz assumption on our integrand of the form appearing in Theorem 4.3.13, which we cannot have if we wish to allow F to take the value $+\infty$. Thus, for now at least, we content ourselves with including the lower-growth bound also in our formula for the closed quasiconvex envelope. We note that the lower growth assumption is also present in the relaxation result in Kristensen's paper [102], on which we base our relaxation proof.

Before we proceed to the proof let us recall the following classical result due to Kuratowski and Ryll-Nardzewski:

Theorem 4.3.6 (see [111])**.** *Let X be a metric space and Y be a separable and complete metric space. Fix a multi-valued function $\mathcal{G}\colon X \to 2^Y$. If for any closed set $K \subset Y$ the set $\{x \in X\colon \mathcal{G}(x) \cap K \neq \emptyset\}$ is Borel measurable then $\mathcal{G}$ admits a measurable selector, i.e., there exists a Borel measurable function $g\colon X \to Y$ such that for all $x \in X$ we have $g(x) \in \mathcal{G}(x)$.*

Proof of Proposition 4.3.5. The argument that we present closely follows the one in a recent paper by the author (see [145]), where it was employed in the context of $\mathcal{A}$-quasiconvexity.
Denote

$$R(X) := \inf_{\nu\in\mathbb{H}^p_0} \langle F(\cdot + X), \nu\rangle.$$

Clearly for any $\nu \in \mathbb{H}^p_0$ and $X \in \mathbb{R}^{n\times m}$ we have $\overline{F}(X) \leqslant \langle F(\cdot + X), \nu\rangle$, therefore taking the infimum over $\nu \in \mathbb{H}^p_0$ yields

$$\overline{F}(X) \leqslant R(X),$$

hence showing that R is closed $\mathrm{W}^{\mathbf{a},p}$-quasiconvex will give the reverse inequality and end the proof, as one immediately gets $R \leqslant F$ by testing with $\nu := \delta_0 \in \mathbb{H}^p_0$.

To show that R is lower semicontinuous fix $X_0 \in \mathbb{R}^{n\times m}$, a sequence $X_j \to X_0$, and an $\varepsilon > 0$. We will show that

$$\varepsilon + \liminf_{j\to\infty} R(X_j) \geqslant R(X_0).$$

Without loss of generality assume that $\lim_{j\to\infty} R(X_j) = \liminf_{j\to\infty} R(X_j) < \infty$, and let M be such that $R(X_j) + \varepsilon \leqslant M$ for all j. By definition of R, for each X_j there exists $\nu_j \in \mathbb{H}_0^p$ with

$$M \geqslant R(X_j) + \varepsilon \geqslant \langle F(\cdot + X_j), \nu_j\rangle.$$

Our growth assumption on F and boundedness of $|X_j|$ (as a convergent sequence) then imply

$$M \geqslant \int_{\mathbb{R}^{n\times m}} c\,|X + X_j|^p \,\mathrm{d}\nu_j \geqslant C\left(\int_{\mathbb{R}^{n\times m}} |X|^p \,\mathrm{d}\nu_j - 1\right),$$

which yields $\sup_j \int_{\mathbb{R}^{n\times m}} |X|^p \,\mathrm{d}\nu_j < \infty$. We see that the family $\{\nu_j\}$ is bounded in $\mathcal{E}_p^*$, therefore we may extract a weakly*-convergent subsequence from it — without loss of generality assume that the whole sequence converges, i.e., $\nu_j \overset{*}{\rightharpoonup} \nu_0$ in $\mathcal{E}_p^*$. By Lemma 3.6.4 we have $\nu_0 \in \mathbb{H}_0^p$. Moreover $\delta_{X_j} * \nu_j \overset{*}{\rightharpoonup} \delta_{X_0} * \nu_0$. Since F is lower semicontinuous and bounded from below we have

$$\varepsilon + \liminf_{j\to\infty} R(X_j) \geqslant \liminf_{j\to\infty} \langle F, \delta_{X_j} * \nu_j\rangle \geqslant \langle F, \delta_{X_0} * \nu_0\rangle = \int_{\mathbb{R}^{n\times m}} F(\cdot + X_0)\,\mathrm{d}\nu_0 \geqslant R(X_0),$$

where the last inequality comes from the definition of R and the fact that $\nu \in \mathbb{H}_0^p$. Since $\varepsilon > 0$ was arbitrary we conclude that R is in fact lower semicontinuous.

It now remains to show that R satisfies the Jensen's inequality with respect to homogeneous oscillation $\mathrm{W}^{\mathbf{a},p}$-gradient Young measures. To that end fix $X_0 \in \mathbb{R}^{n\times m}$ and $\nu \in \mathbb{H}_{X_0}^p$. We wish to show that $R(X_0) \leqslant \int_{\mathbb{R}^{n\times m}} R\,\mathrm{d}\nu$. Observe that we may assume without loss of generality that $\int_{\mathbb{R}^{n\times m}} R\,\mathrm{d}\nu < \infty$, as the case where this integral is infinite is trivial. Let us fix an $\varepsilon > 0$ and observe that, by definition of R, for all $X \in \mathbb{R}^{n\times m}$ there exists $\nu_X \in \mathbb{H}_0^p$ satisfying

$$\langle F(\cdot + X), \nu_X\rangle \leqslant \varepsilon + R(X),$$

so that, for now only formally,

$$\int_{\mathbb{R}^{n\times m}} \left(\int_{\mathbb{R}^{n\times m}} F(\cdot + X)\,\mathrm{d}\nu_X\right) \mathrm{d}\nu(X) \leqslant \varepsilon + \int_{\mathbb{R}^{n\times m}} R\,\mathrm{d}\nu.$$

Now — if we manage to show that ν_X may be chosen in such a way that $X \mapsto \nu_X$ is weak* measurable and that the measure μ defined by duality as

$$\langle g, \mu \rangle := \int_{\mathbb{R}^{n\times m}} \left(\int_{\mathbb{R}^{n\times m}} g(\cdot + X)\,\mathrm{d}\nu_X \right) \mathrm{d}\nu(X) \tag{4.9}$$

is a homogeneous $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure with mean X_0 then the claim will follow, as by definition $\langle F, \mu \rangle \geqslant R(X_0)$.

Remark 8. There is a delicate point to be emphasized here. A careful reader may notice that weak* measurability of $X \mapsto \nu_X$ only means Lebesgue measurability of $X \mapsto \int_{\mathbb{R}^{n\times m}} g(\cdot + X)\,\mathrm{d}\nu_X$, which is not necessarily enough to integrate this function with respect to ν. However, if we manage to get Borel measurability of the function in question then the construction is justified, as ν is a Radon (hence Borel) measure — we will call such a map Borel weak* measurable. It is clear that if one makes sense of the integration on the right-hand side of (4.9) then it defines a linear functional on $C_0(\mathbb{R}^{n\times m})$. Its boundedness follows from the fact that all ν_X, ν are probability measures, thus showing that the functional is given by some finite Radon measure μ.

For the measurable selection part we define a multifunction $\mathcal{G}$ given by

$$\mathcal{G}(X) := \left\{ \mu \in \mathbb{H}_0^p \colon \int_{\mathbb{R}^{n\times m}} F(\cdot + X)\,\mathrm{d}\mu \leqslant \varepsilon + R(X) \right\}.$$

For the measurable selection result we intend to use (see Theorem 4.3.6) we need $\mathcal{G}$ to take values in 2^Y for some complete metric space Y. For that we define, for a given $M > 0$,

$$\Omega_M := \{X \in \mathbb{R}^{n\times m} \colon |X| < M, R(X) \leqslant M\}.$$

Observe that since we assumed R to be integrable with respect to ν, we have that $X \in \bigcup_{M=1}^{\infty} \Omega_M$ for ν-a.e. $X \in \mathbb{R}^{n\times m}$. Let us fix $M \in \mathbb{N}$. Then, for any $X \in \Omega_M$ and any $\mu \in \mathcal{G}(X)$, we have

$$\int_{\mathbb{R}^{n\times m}} F(W + X)\,\mathrm{d}\mu(W) \leqslant \varepsilon + R(X) \leqslant 2\varepsilon + R(X) \leqslant M + 2\varepsilon.$$

The factor 2 in front of ε is not important here, we only put it there to allow for some room in the later part of the argument. Due to the growth assumption on F we have

$$\begin{aligned}\int_{\mathbb{R}^{n\times m}} F(W + X)\,\mathrm{d}\mu(W) \geqslant & C \int_{\mathbb{R}^{n\times m}} |W + X|^p \,\mathrm{d}\mu(W) \geqslant \\ & C \int_{\mathbb{R}^{n\times m}} |W|^p \,\mathrm{d}\mu(W) - C^{-1} |X|^p .\end{aligned}$$

Finally $\int_{\mathbb{R}^{n\times m}} |W|^p \,\mathrm{d}\mu(W) \leqslant C_M$, holds for all $\mu \in \mathcal{G}(X)$, with the constant C_M depending only on M (and ε). Therefore, we may consider our operator $\mathcal{G}$ as a map $\Omega_M \to 2^{Y_M}$, where

$$Y_M := \left\{ \mu \in \mathbb{H}_0^p \colon \int_{\mathbb{R}^{n\times m}} |W|^p \,\mathrm{d}\mu \leqslant C_M \right\}.$$

The set Y_M may be equipped with the weak* topology inherited from $\mathcal{E}_p^*$. Since we put a uniform bound on the p-th moments (so also on the norm in $\mathcal{E}_p^*$), this topology is metrisable in a complete and separable manner when restricted to Y_M. To prove that first recall that due to Lemma 3.6.4 $\mathbb{H}_0^p$ is weak* closed in $\mathcal{E}_p^*$. Since $|\cdot|^p \in \mathcal{E}_p$ we know that the map $\mu \mapsto \int_{\mathbb{R}^{n\times m}} |W|^p \,\mathrm{d}\mu$ is weak* continuous, thus Y_M is weak* closed and bounded. The Banach-Alaoglu Theorem (see for example Theorem 3.16 in [27]) then implies that Y_M is weak* compact. Since $\mathcal{E}_p$ is clearly separable we deduce that the weak* topology on Y_M is metrisable (see Theorem 3.28 in [27]). Finally, compact metric spaces are complete and separable, thus proving our claim.

Lemma 4.3.7. *For any $X \in \Omega_M$ the set $\mathcal{G}(X)$ is non-empty and closed.*

Proof. The fact that $\mathcal{G}(X) \neq \emptyset$ comes straight from the definition of R. To show that it is closed it is enough to show that it is sequentially closed. Let us then fix a sequence $\{\mu_j\} \subset \mathcal{G}(X)$ and assume that it converges weak* in $\mathcal{E}_p^*$ to some $\mu \in Y_M$. Since the function F is lower semicontinuous and bounded from below we get by Lemma 3.6.5 that

$$R(X) + \varepsilon \geqslant \liminf_{j\to\infty} \int_{\mathbb{R}^{n\times m}} F(\cdot + X) \,\mathrm{d}\mu_j \geqslant \int_{\mathbb{R}^{n\times m}} F(\cdot + X) \,\mathrm{d}\mu,$$

so $\mu \in \mathcal{G}(X)$, which ends the proof. □

Lemma 4.3.8. *For any non-empty closed set $O \subset Y_M$ the set*

$$\{X \in \Omega_M \colon \mathcal{G}(X) \cap O \neq \emptyset\}$$

is Borel measurable.

Proof. First note that we may rewrite the set in question as

$$\bigcap_{k=1}^{\infty} \left\{ X \in \Omega_M \colon \inf_{\mu\in O} \int_{\mathbb{R}^{n\times m}} F(\cdot + X) \,\mathrm{d}\mu \leqslant R(X) + \varepsilon(1 + 2^{-k}) \right\}.$$

Hence, it is enough to show that the sets

$$\left\{ X \in \Omega_M \colon \inf_{\mu\in O} \int_{\mathbb{R}^{n\times m}} F(\cdot + X) \,\mathrm{d}\mu \leqslant R(X) + \varepsilon(1 + 2^{-k}) \right\}$$

are all Borel measurable. Define

$$U(X) := \inf_{\mu \in O} \int_{\mathbb{R}^{n\times m}} F(\cdot + X)\, \mathrm{d}\mu.$$

We claim that U is lower semicontinuous. Let $X_j \to X$. We need to show that $\liminf_{j\to\infty} U(X_j) \geqslant U(X)$. Without loss of generality assume that the lim inf is a true limit and that it is finite, i.e.,

$$\lim_{j\to\infty} U(X_j) = \liminf_{j\to\infty} U(X_j) < \infty.$$

By definition of U, for each k there exists a measure $\mu_j \in O$ with

$$\int_{\mathbb{R}^{n\times m}} F(\cdot + X_j)\, \mathrm{d}\mu_j \leqslant U(X_j) + 1/k.$$

Therefore

$$\lim_{j\to\infty} \int_{\mathbb{R}^{n\times m}} F(\cdot + X_j)\, \mathrm{d}\mu_j = \lim_{j\to\infty} U(X_j).$$

Since the set O is a closed subset of a compact space Y_M we may extract an $\mathcal{E}_p^*$ weak* convergent subsequence from μ_j. Without loss of generality assume that the entire sequence μ_j converges weak* to some $\mu \in O$. This, combined with $X_j \to X$, implies that we have

$$\delta_{X_j} * \mu_j \overset{*}{\rightharpoonup} \delta_X * \mu$$

in the sense of probability measures. Therefore, since F is lower semicontinuous, the portmanteau theorem yields

$$\liminf_{j\to\infty} \int_{\mathbb{R}^{n\times m}} F(\cdot + X_j)\, \mathrm{d}\mu_j = \liminf_{j\to\infty} \int_{\mathbb{R}^{n\times m}} F\, \mathrm{d}\left(\delta_{X_j} * \mu_j\right) \geqslant \int_{\mathbb{R}^{n\times m}} F\, \mathrm{d}\left(\delta_X * \mu\right) = \int_{\mathbb{R}^{n\times m}} F(\cdot + X)\, \mathrm{d}\mu \geqslant U(X),$$

which shows that U is indeed lower semicontinuous. Since the set

$$\left\{X \in \Omega_M \colon \inf_{\mu\in O} \int_{\mathbb{R}^{n\times m}} F(\cdot + X)\, \mathrm{d}\mu \leqslant R(X) + \varepsilon(1 + 2^{-k})\right\}$$

is the same as

$$\{X \in \Omega_M \colon U(X) \leqslant R(X) + \varepsilon(1 + 2^{-k})\},$$

and both U and R are lower semicontinuous (hence Borel measurable) the set in question is Borel measurable as well, which ends the proof. □

Now, thanks to Lemmas 4.3.7 and 4.3.8 we may use Theorem 4.3.6 to deduce the existence of a weak* measurable map $\nu^M \colon \Omega_M \to \mathbb{H}_0^p$ such that for any $X \in \Omega_M$ the measure ν_X^M satisfies

$$\int_{\mathbb{R}^{n\times m}} F(\cdot + X) d\nu_X^M \leqslant \varepsilon + R(X).$$

Finally let us define the map $\widetilde{\nu} \colon \mathbb{R}^{n\times m} \to \mathbb{H}_0^p$ by

$$\widetilde{\nu}_X := \begin{cases} \nu_X^M \text{ for } X \in \Omega_M \setminus \Omega_{M-1} \\ \widetilde{\mu} \text{ for } X \notin \bigcup_{M=1}^{\infty} \Omega_M, \end{cases}$$

where $\widetilde{\mu}$ is some arbitrary element of the (non-empty) set $\mathbb{H}_0^p$. Observe that the choice of $\widetilde{\mu}$ does not matter, as we have already observed that the set $\mathbb{R}^{n\times m} \setminus \bigcup_{M=1}^{\infty} \Omega_M$ is of ν measure 0. This set is also Borel since we already know that R is Borel measurable, hence each Ω_M is Borel. Clearly the map $\widetilde{\nu}$ is Borel weak* measurable in the sense of Remark 8, i.e., it is a measurable map from $\mathbb{R}^{n\times m}$ equipped with the Borel σ-algebra into $\mathbb{H}_0^p$ equipped with the weak* topology inherited from $\mathcal{E}_p^*$. Therefore, we may define $\mu \in (C_0(\mathbb{R}^{n\times m}))^*$ as in (4.9). It only remains to show that $\mu \in \mathbb{H}_0^p$.

Positivity of μ results immediately from positivity of all ν_X and ν. In the same way we show that μ is a probability measure, as

$$\langle 1, \mu \rangle = \int_{\mathbb{R}^{n\times m}} \left(\int_{\mathbb{R}^{n\times m}} 1 \, \mathrm{d}\nu_X \right) \mathrm{d}\nu(X) = \int_{\mathbb{R}^{n\times m}} 1 \, \mathrm{d}\nu(X) = 1,$$

since all measures considered are probability measures. To prove that μ has a finite p-th moment we write

$$\langle |\cdot|^p, \mu \rangle = \int_{\mathbb{R}^{n\times m}} \left(\int_{\mathbb{R}^{n\times m}} |\cdot + X|^p \, \mathrm{d}\nu_X \right) \mathrm{d}\nu(X).$$

Using the growth assumption on F we get

$$\int_{\mathbb{R}^{n\times m}} |\cdot + X|^p \, \mathrm{d}\nu_X \leqslant C \int_{\mathbb{R}^{n\times m}} F(\cdot + X) \, \mathrm{d}\nu_X \leqslant C(R(X) + \varepsilon),$$

where the last inequality is satisfied for ν-a.e. X. Integrating with respect to ν gives

$$\langle |\cdot|^p, \mu \rangle \leqslant C \left(\varepsilon + \int_{\mathbb{R}^{n\times m}} R(X) \, \mathrm{d}\nu(X) \right) < \infty,$$

since, by assumption, R is integrable with respect to ν. Lastly, it remains to show that μ satisfies the inequality in Theorem 3.7.1. Fix any continuous functions $g \colon \mathbb{R}^{n\times m} \to \mathbb{R}$ with $|g(v)| \leqslant C(1 + |v|^p)$ for some constant C. We have

$$\begin{aligned} \langle \mu, g \rangle &= \int_{\mathbb{R}^{n\times m}} \left(\int_{\mathbb{R}^{n\times m}} g(\cdot + X) \, \mathrm{d}\nu_X \right) \mathrm{d}\nu(X) \\ &\geqslant \int_{\mathbb{R}^{n\times m}} \mathcal{Q}g(X) \, \mathrm{d}\nu(X) \geqslant \mathcal{Q}(\mathcal{Q}g)(X_0) = \mathcal{Q}g(X_0), \end{aligned}$$

where the first inequality comes from the fact that all ν_X's are Young measures with mean 0, the second one from the respective property of ν, and the last equality from Lemma 3.5.3. This shows that we indeed have $\mu \in \mathbb{H}^p_{X_0}$ and ends the proof, as discussed in (4.9). □

4.3.3 Relaxation in the extended real-valued setting

We proceed with the final relaxation result in the setting of extended real-valued integrands. The statements and proofs follow very closely the content of a recent article by the author (see [145]). However, since the proofs had to be altered to fit in the mixed smoothness setting we present them here in full.

We begin by defining a relaxed notion of convergence for vector fields that are nearly (up to a small in L^p error) $\mathbf{a}$-gradients of functions in $\mathrm{W}^{\mathbf{a},p}$. The notion is reminiscent of the one often used in the $\mathcal{A}$-free setting, where instead of working with sequences that satisfy the constraint exactly, i.e., with $\mathcal{A}V_j = 0$, one only requires $\mathcal{A}V_j \to 0$ strongly in $\mathrm{W}^{-1,p}$, see for example [67]. In the case of standard first order gradients it corresponds to the condition $\mathrm{curl} V_j \to 0$ strongly in $\mathrm{W}^{-1,p}(\Omega)$ investigated in [103].

Definition 4.3.9. *We say that a sequence of vector fields $V_j \in \mathrm{L}^p(\Omega;\mathbb{R}^{n\times m})$ is a sequence of approximate $\mathrm{W}^{\mathbf{a},p}$ gradients if there exist sequences $u_j \in \mathrm{W}^{\mathbf{a},p}(\Omega;\mathbb{R}^n)$ and $v_j \in \mathrm{L}^p(\Omega;\mathbb{R}^{n\times m})$ such that*

$$V_j = \nabla_{\mathbf{a}} u_j + v_j$$

and $v_j \to 0$ strongly in L^p.

Following [103] we introduce the following notion of convergence:

Definition 4.3.10. *We say that a sequence of vector fields $V_j \in \mathrm{L}^p(\Omega;\mathbb{R}^{n\times m})$ converges to V in the sense of approximate $\mathrm{W}^{\mathbf{a},p}$ gradients if V_j converges to V weakly in L^p and $(V_j - V)$ is a sequence of approximate $\mathrm{W}^{\mathbf{a},p}$ gradients. In such a case we write $V_j \to_{\mathbf{a}\text{-}p} V$.*

Proposition 3.1.2 immediately implies the following:

Lemma 4.3.11. *Assume that Ω satisfies the weak $\mathbf{a}$-horn condition. Suppose that a sequence $V_j = \nabla_{\mathbf{a}} u_j + v_j \in \mathrm{L}^p(\Omega;\mathbb{R}^{n\times m})$ of approximate $\mathrm{W}^{\mathbf{a},p}$ gradients converges weakly to 0 in L^p and generates an oscillation Young measure ν. Then $\{\nabla_{\mathbf{a}} u_j\}$ generates the same Young measure ν. In particular, any oscillation Young measure generated by a sequence of approximate $\mathrm{W}^{\mathbf{a},p}$ gradients is an oscillation $\mathrm{W}^{\mathbf{a},p}$-gradient Young measure.*

This and Lemma 4.2.6 easily imply the following:

Corollary 4.3.12. *Let Ω be a bounded open domain satisfying the weak* $\mathbf{a}$*-horn condition. Suppose that $F\colon \mathbb{R}^{n\times m} \to (-\infty, \infty]$ is bounded from below and closed* $\mathrm{W}^{\mathbf{a},p}$*-quasiconvex. Then the functional*

$$\mathrm{I}(V) := \int_\Omega F(V)\,\mathrm{d}x$$

is sequentially lower semicontinuous with respect to approximate $\mathrm{W}^{\mathbf{a},p}$ *gradient convergence.*

Proof. Since, by Lemma 4.3.11 the Young measures generated by sequences of approximate $\mathrm{W}^{\mathbf{a},p}$ gradients are $\mathrm{W}^{\mathbf{a},p}$-gradient Young measures, the argument of Lemma 4.2.6 carries through unchanged. □

The main result here is the following:

Theorem 4.3.13. *If $F\colon \mathbb{R}^{n\times m} \to (-\infty, \infty]$ is a continuous integrand satisfying $F(X) \geqslant C|X|^p - C^{-1}$ for some $C > 0$ then the sequentially (with respect to approximate* $\mathrm{W}^{\mathbf{a},p}$ *gradient convergence) weakly lower semicontinuous envelope of the functional* I_F *is given by*

$$\bar{\mathrm{I}}_F[V] := \inf_{V_j \to_{\mathbf{a}\text{-}p} V} \left\{ \liminf_{j\to\infty} \mathrm{I}_F[V_j] \right\} = \int_\Omega \overline{F}(V(x))\,\mathrm{d}x,$$

where the infimum is taken over all sequences V_j converging to V in the sense of approximate $\mathrm{W}^{\mathbf{a},p}$ *gradient convergence. As before, $\overline{F}$ denotes the closed* $\mathrm{W}^{\mathbf{a},p}$*-quasiconvex envelope of F.*

Proof. Corollary 4.3.12 guarantees that $\bar{\mathrm{I}}_F[V] \geqslant \int_\Omega \overline{F}(V(x))\,\mathrm{d}x$, thus we only need to prove the opposite inequality. If F is identically equal $+\infty$ then there is nothing to show, so we may restrict to proper integrands. Using a translation we may, without loss of generality, assume $F(X) \geqslant C|X|^p$. In any case, the fact that F is bounded from below immediately implies the same for $\overline{F}$. Fix any $V \in \mathrm{L}^p$. Without loss of generality we may assume $\int_\Omega \overline{F}(V(x))\,\mathrm{d}x < \infty$, as otherwise there is nothing to prove. Fix an $\varepsilon > 0$ and observe that clearly we must have $\overline{F}(V(x)) < \infty$ a.e. in Ω. Therefore, using Proposition 4.3.5, we may find a family of homogeneous oscillation $\mathrm{W}^{\mathbf{a},p}$-gradient Young measures $\{\nu_x\}_{x\in\Omega}$ with mean 0 and such that, for almost every $x \in \Omega$, we have

$$\overline{F}(V(x)) + \varepsilon \geqslant \int_{\mathbb{R}^{n\times m}} F(\cdot + V(x))\,\mathrm{d}\nu_x. \tag{4.10}$$

Using exactly the same argument as in the proof of Proposition 4.3.5 we may ensure weak* measurability of $x \to \nu_x$. We intend to show that ν is a suitable Young measure using Theorem 3.7.2. Recall that we need to prove the following:

i) there exists $v \in \mathrm{W}^{\mathbf{a},p}(\Omega)$ such that

$$\nabla_{\mathbf{a}} v(x) = \langle \nu_x, \mathrm{Id} \rangle \text{ for a.e. } x \in \Omega;$$

ii)

$$\int_\Omega \int_{\mathbb{R}^{n\times m}} |W|^p \,\mathrm{d}\nu_x(W)\,\mathrm{d}x < \infty;$$

iii) for a.e. $x \in \Omega$ and all continuous functions $g\colon \mathbb{R}^{n\times m} \to \mathbb{R}$ satisfying $|g(W)| \leqslant C(1+|W|^p)$ for some positive constant C one has

$$\langle \nu_x, g \rangle \geqslant \mathcal{Q}g(\langle \nu_x, \mathrm{Id} \rangle).$$

The first point is clearly satisfied, as all our measures are of mean 0. The second one may be checked in the same way as in the already mentioned proof of Proposition 4.3.5, using the growth assumption on F. Finally, the third point results immediately from the fact that all ν_x's are, by definition, elements of $\mathbb{H}^p_0$, so we may use Theorem 3.7.1.

This shows that ν is indeed generated by some p-equiintegrable family $\{\nabla_{\mathbf{a}} w_j\}$ with $w_j \in \mathrm{W}^{\mathbf{a},p}(\Omega)$ and $w_j \rightharpoonup 0$ in $\mathrm{W}^{\mathbf{a},p}$. For a given $M \in \mathbb{N}$ consider $F^M(z) := \min(F(z), M(|z|^p+1))$. Clearly, for each M, the function F^M is continuous and the family $\{F^M(V + \nabla_{\mathbf{a}} w_j)\}_j$ is p-equiintegrable, due to the same property of $\{V + \nabla_{\mathbf{a}} w_j\}$. Theorem 3.1.1 then yields

$$\int_\Omega F^M(V + \nabla_{\mathbf{a}} w_j)\,\mathrm{d}x \to \int_\Omega \left(\int_{\mathbb{R}^{n\times m}} F^M(V(x) + \cdot)\,\mathrm{d}\nu_x \right) \mathrm{d}x.$$

On the other hand, since $F^M \leqslant F$ and ν_x are non-negative and satisfy (4.10), we have

$$\begin{aligned} \int_\Omega \left(\int_{\mathbb{R}^{n\times m}} F^M(V(x) + \cdot)\,\mathrm{d}\nu_x \right) \mathrm{d}x &\leqslant \int_\Omega \left(\int_{\mathbb{R}^{n\times m}} F(V(x) + \cdot)\,\mathrm{d}\nu_x \right) \mathrm{d}x \\ &\leqslant \int_\Omega \overline{F}(V(x))\,\mathrm{d}x + \varepsilon. \end{aligned}$$

From this we deduce, through a diagonal extraction, that there exists a sequence $j(M) \in \mathbb{N}$ with $\lim_{M\to\infty} j(M) = \infty$ such that for all M one has

$$\int_\Omega F^M(V + \nabla_{\mathbf{a}} w_{j(M)})\,\mathrm{d}x \leqslant \int_\Omega \overline{F}(V(x))\,\mathrm{d}x + 2\varepsilon. \tag{4.11}$$

Define the set

$$\mathcal{G}_M := \left\{x \in \Omega\colon F(V(x) + \nabla_{\mathbf{a}} w_{j(M)}(x)) \leqslant M(|V(x) + \nabla_{\mathbf{a}} w_{j(M)}(x)|^p + 1)\right\},$$

and fix some $X_0 \in \mathbb{R}^{n\times m}$ for which $F(X_0) < \infty$ — such a point exists, as F is proper. Next define a vector field W_M in such a way that

$$V(x) + W_M(x) = (V(x) + \nabla_{\mathbf{a}} w_{j(M)}(x))\mathbb{1}_{\mathcal{G}_M} + X_0 \mathbb{1}_{\mathcal{G}_M^c}. \tag{4.12}$$

We claim that $\{V + W_M\}_M$ is an admissible vector field in the $\bar{\mathrm{I}}_F[V]$ problem. For that it is enough to show that $\|V + W_M - (V + \nabla_{\mathbf{a}} w_{j(M)})\|_{\mathrm{L}^p(\Omega)} \to 0$. By definition we have

$$\begin{aligned}\|V + W_M - (V + \nabla_{\mathbf{a}} w_{j(M)})\|_{\mathrm{L}^p(\Omega)} = \|V + W_M - (V + \nabla_{\mathbf{a}} w_{j(M)})\|_{\mathrm{L}^p(\mathcal{G}_M^c)} \leqslant \\ \|X_0\|_{\mathrm{L}^p(\mathcal{G}_M^c)} + M^{-1}\left(\int_\Omega F^M(V + \nabla_{\mathbf{a}} w_{j(M)})\,\mathrm{d}x\right)^{1/p},\end{aligned}$$

where the last inequality comes from the definition of the set $\mathcal{G}_M^c$ (and extending the integral to all of Ω). Now, (4.11) yields

$$M^{-1}\left(\int_\Omega F^M(V + \nabla_{\mathbf{a}} w_{j(M)})\,\mathrm{d}x\right)^{1/p} \leqslant M^{-1}\left(\int_\Omega \overline{F}(V(x))\,\mathrm{d}x + 2\varepsilon\right)^{1/p},$$

thus showing the desired convergence to 0 in L^p, as $\|X_0\|_{\mathrm{L}^p(\mathcal{G}_M^c)} \to 0$ results simply from the fact that clearly the Lebesgue measure of $\mathcal{G}_M^c$ tends to 0, because

$$F(V(x) + \nabla_{\mathbf{a}} w_{j(M)}(x)) > M \text{ on } \mathcal{G}_M^c,$$

and we have a uniform (with respect to M) bound on the integral of the function in question. This implies in particular that $V + W_M$ converges to V in the sense of approximate $\mathrm{W}^{\mathbf{a},p}$ gradients convergence. Therefore, we have

$$\begin{aligned}\bar{\mathrm{I}}_F[V] &\leqslant \liminf_{M\to\infty} \int_\Omega F(V + W_M)\,\mathrm{d}x \\ &= \liminf_{M\to\infty} \int_{\mathcal{G}_M} F^M(V + w_{j(M)})\,\mathrm{d}x + \int_{\mathcal{G}_M^c} F(X_0)\,\mathrm{d}x \\ &\leqslant \liminf_{M\to\infty} \int_\Omega \overline{F}(V(x))\,\mathrm{d}x + 2\varepsilon = \int_\Omega \overline{F}(V(x))\,\mathrm{d}x + 2\varepsilon,\end{aligned}$$

where the last inequality results from (4.11) and the measure of $\mathcal{G}_M^c$ tending to 0. Since $\varepsilon > 0$ was arbitrary the proof is complete. □

Chapter 5

Regularity

Having established the coercivity and lower semicontinuity results describing the existence of minimisers of variational problems in the mixed smoothness setting we now turn to the question of regularity of such minimisers. In the spirit of Evans' partial regularity result (see [61]) we show how a strict **a**-quasiconvexity assumption on the integrand may be used to establish local Hölder continuity of maximal derivatives of minimisers on an open subdomain of full measure. The result we obtain is what is called an ε-regularity result — we characterise the regular points, i.e., the points around which the **a**-gradient of the minimiser is Hölder continuous, through an appropriate smallness condition imposed on the excess of the **a**-gradient. The Lebesgue differentiation theorem then allows us to conclude that the condition is satisfied Lebesgue almost everywhere, although it does not provide any better bounds on the size of the singular set. In fact the question of estimating the Hausdorff dimension of the singular set under a quasiconvexity assumption is still open even in the classical setting of first order gradients, and so far the only result in this direction is due to Kristensen and Mingione (see [105]) and it is conditional on the a priori assumption that the minimiser is Lipschitz, thus for now we content ourselves with a Lebesgue-a.e. type result.

The approach that we take here is, to a large extent, based on the one in [82] by Gmeineder and Kristensen, which, in turn, stems from earlier works by a number of authors, such as De Giorgi (see [46]), Almgren (see [6] and [7]), Giusti and Miranda (see [81]), Morrey (see [130]), Evans (see [61]), and many others who have subsequently contributed to the development of the theory. The general idea of the argument is to use good excess decay estimates (i.e. estimates in terms of Campanato norms) available for linear systems and then carry them over to the minimiser through a local linearisation procedure and a Caccioppoli inequality.

We begin by investigating the regularity properties of linear quasielliptic systems in Section 5.1, largely based on Giusti's work (see [77], [78]). We then introduce the notion of strict **a**-quasiconvexity and, using Widman's (see [179]) hole-filling technique, prove an anisotropic Caccioppoli inequality in Section 5.2. Finally, the core of the argument is contained in Section 5.3 where we explain how the linearisation procedure is carried out and how to choose an appropriate linear equation, solution of which should (in a neighbourhood of a given point) be close to the minimiser in question. Estimates on the deviation of the minimiser from the solution of the linearised equation are then obtained through an explicit construction of an appropriate test function, similarly to [82]. These, in combination with the estimates on the solution of the linearised equation, finally yield estimates on the excess decay of the minimiser. In the last part of this chapter we explain how these estimates may be iterated, thus ending the proof.

Let us remark that, in this thesis, we only study regularity of solutions for variational problems with integrands of controlled quadratic growth, i.e., with bounded second derivative and in the $p = 2$ case. However, in collaboration with my doctoral advisor Professor Jan Kristensen, we are currently working on extending them to cover p-growth in the full reflexive range $p \in (1, \infty)$.

Finally, note that, in all the estimates that follow, the constants may depend on the smoothness vector $\mathbf{a}$, as well as the dimensions N, n, and m. However, these parameters are fixed throughout the chapter, thus we will not track how (and if) they influence the constants.

5.1 Linear quasielliptic equations

We begin with a simpler problem, i.e., the regularity of linear quasielliptic partial differential equations, in the particular case of equations of the form

$$\begin{cases} \nabla_{\mathbf{a}}^* T \nabla_{\mathbf{a}} u = f \text{ in } \Omega, \\ u = u_0 \text{ on } \partial\Omega. \end{cases} \tag{5.1}$$

Here $\nabla_{\mathbf{a}}^*$ denotes the distributional adjoint of $\nabla_{\mathbf{a}}$, Ω is a bounded open domain satisfying the weak **a**-horn condition, and the boundary datum u_0 is assumed to belong to the space $\mathrm{W}^{\mathbf{a},p}(\mathbb{R}^N)$. The condition $u = u_0$ on $\partial\Omega$ is understood as $u \in \mathrm{W}^{\mathbf{a},p}_{u_0}(\Omega)$, i.e., $h = u + \varphi$ for some $\varphi \in \mathrm{W}^{\mathbf{a},p}_0(\Omega)$. The source term f is assumed to belong to the space $\mathrm{W}^{-\mathbf{a},p}(\Omega)$, i.e., the dual of $\mathrm{W}^{\mathbf{a},p}_0(\Omega)$. One may show, as in the standard case, that any $f \in \mathrm{W}^{-\mathbf{a},p}(\Omega)$ may be written as $f = \sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \partial^\alpha f_\alpha$ for some $f_\alpha \in \mathrm{L}^p$. In this case $\|f\|_{\mathrm{W}^{-\mathbf{a},p}(\Omega)} \leqslant \sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \|f_\alpha\|_{\mathrm{L}^p(\Omega)}$, with

$$\|f\|_{\mathrm{W}^{-\mathbf{a},p}(\Omega)} = \inf_{\{f_\alpha\}_\alpha} \sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \|f_\alpha\|_{\mathrm{L}^p(\Omega)},$$

where the infimum is taken over all choices of $\{f_\alpha\}_\alpha$ for which the representation $f = \sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \partial^\alpha f_\alpha$ holds.

Throughout this chapter we assume that the operator $\nabla_{\mathbf{a}}^* T \nabla_{\mathbf{a}}$ is strictly **a**-quasielliptic. For convenience we often say instead that the tensor T is strictly **a**-quasielliptic, by which we understand the following:

Definition 5.1.1. *We say that a constant, symmetric tensor T is strictly* **a***-quasielliptic with* **a***-quasiellipticity constant $\nu > 0$ if*

$$\int_Q \langle T\nabla_{\mathbf{a}} u, \nabla_{\mathbf{a}} u\rangle \,\mathrm{d}x \geqslant \nu \int_Q |\nabla_{\mathbf{a}} u|^2 \,\mathrm{d}x,$$

for all $u \in \mathrm{W}_0^{\mathbf{a},2}(Q)$.

Our goal for this section is the following:

Proposition 5.1.2. *Let T be a constant, strictly* **a***-quasielliptic tensor. For any exponent $p \in [2, \infty)$ the equation*

$$\begin{cases} \nabla_{\mathbf{a}}^* T \nabla_{\mathbf{a}} u = G \text{ in } Q, \\ u \in \mathrm{W}_0^{\mathbf{a},p}(Q) \end{cases} \tag{5.2}$$

with $G \in \mathrm{W}^{-a,p}(Q)$ is uniquely solvable and the solution u satisfies

$$\|u\|_{\mathrm{W}_0^{\mathbf{a},p}} \leqslant C\|G\|_{\mathrm{W}^{-a,p}},$$

with the constant $C = C(p, \nu)$ depending only on p and the strict **a***-quasiellipticity constant of T.*

We defer the proof to the end of this section, as we first need to establish a number of preliminary results. To begin with, let us treat the simple case $p = 2$.

Lemma 5.1.3. *Let T be a constant, strictly* **a***-quasielliptic tensor. Then the equation*

$$\begin{cases} \nabla_{\mathbf{a}}^* T \nabla_{\mathbf{a}} u = G \text{ in } Q, \\ u \in \mathrm{W}_0^{\mathbf{a},2}(Q) \end{cases} \tag{5.3}$$

with $G \in \mathrm{W}^{-a,2}(Q)$ is uniquely solvable and the solution u satisfies

$$\|u\|_{\mathrm{W}_0^{\mathbf{a},2}} \leqslant C\|G\|_{\mathrm{W}^{-a,2}},$$

where the constant $C = C(\nu)$ depends only on the strict **a***-quasiellipticity constant of T.*

Proof. Since T has constant coefficients the bilinear form

$$(u, v) \mapsto \int_Q \langle T\nabla_{\mathbf{a}} u, \nabla_{\mathbf{a}} v\rangle \,\mathrm{d}x,$$

is clearly bounded. Since we assume T to be strictly $\mathbf{a}$-quasielliptic the form is also coercive. Thus, the result follows from the Lax-Milgram theorem. □

It is equally easy to incorporate non-zero boundary conditions:

Lemma 5.1.4. *Let T be a constant, strictly $\mathbf{a}$-quasielliptic tensor. Then the equation*

$$\begin{cases} \nabla_{\mathbf{a}}^* T\nabla_{\mathbf{a}} u = 0 \text{ in } Q, \\ u - u_0 \in \mathrm{W}_0^{\mathbf{a},2}(Q) \end{cases} \tag{5.4}$$

with boundary datum $u_0 \in \mathrm{W}^{\mathbf{a},2}(\mathbb{R}^N)$ is uniquely solvable and the solution u satisfies

$$\|u\|_{\mathrm{W}_0^{\mathbf{a},2}(Q)} \leqslant C\|\nabla_{\mathbf{a}} u_0\|_{\mathrm{L}^2(Q)},$$

with the constant $C = C(\nu)$ depending only on the strict $\mathbf{a}$-quasiellipticity constant of T.

Proof. This is equivalent to minimising

$$u \mapsto \int_Q \langle T\nabla_{\mathbf{a}} u, \nabla_{\mathbf{a}} u\rangle \,\mathrm{d}x$$

over $u \in \mathrm{W}_{u_0}^{\mathbf{a},2}(Q)$. The functional is convex and coercive, thus admits a minimiser which solves the desired equation. It is also clear that the solution is unique under fixed boundary conditions. □

5.1.1 Hypoellipticity

The following is a brief discussion of Hörmander's notion of hypoellipticity (see, for example, [88]), which allows us to deduce interior C^∞ smoothness of solutions of the equation $\nabla_{\mathbf{a}}^* T\nabla_{\mathbf{a}} u = 0$.

Definition 5.1.5 (see Definition 4.1.1 in [88])**.** *We say that a differential operator $P(D)$ (scalar or vectorial) is hypoelliptic if for any open set $A \subset \mathbb{R}^N$ and a distribution $u \in \mathcal{D}'(A)$ the equation $P(D)u = 0$ in $\mathcal{D}'(A)$ implies $u \in C^\infty(A)$.*

Definition 5.1.6 (see Definition 3.1 in [77]). *Suppose that $P(D)$ is a differential operator for which there exists a vector $\gamma = (\gamma_1, \dots, \gamma_N)$ of positive integers such that, with $|\alpha : \gamma| := \sum_1^N \alpha_k/\gamma_k$, we may represent*

$$P(D) = \sum_{|\alpha:\gamma|\leqslant 1} a_\alpha \partial^\alpha$$

with some coefficients $a_\alpha \in \mathbb{R}^{n\times n}$. We set

$$P^0(D) := \sum_{|\alpha:\gamma|=1} a_\alpha \partial^\alpha.$$

If $\det(P^0(-i\xi)) \neq 0$ for all $\xi \in \mathbb{R}^N \setminus \{0\}$ then we say that $P(D)$ is quasielliptic.

Let us remark that Hörmander gives the same definition in Theorem 4.1.8 in [88], where he calls such operators 'semielliptic'. However, in this work we have decided to follow Giusti's terminology, and we will use the word 'quasielliptic'.

Theorem 5.1.7 (see Theorem 4.1.8 in [88]). *Quasielliptic operators are hypoelliptic.*

Proof. For scalar operators this is the content of Theorem 4.1.8 in [88], thus we only give a short explanation on how this extends to systems, following the general approach outlined in Section 3.8 of the aforementioned book. First of all, observe that if $P(D)$ is a quasielliptic system then $\det P(D)$ is a quasielliptic scalar operator. Thus, the theorem mentioned before implies that $\det P(D)$ is hypoelliptic. Having established this, let us assume that, for some $f \in C^\infty$, we have $P(D)u = f$ in some open set Ω. Let us denote by ${}^{\mathrm{co}}P_{i,j}$ the matrix formed by the cofactors in $P_{j,i}$. Then $({}^{\mathrm{co}}P)P = (\det P)\,\mathrm{Id}$, so that applying ${}^{\mathrm{co}}P$ to both sides of our equation yields

$$(\det P)\,\mathrm{Id}\, u = ({}^{\mathrm{co}}P)Pu = ({}^{\mathrm{co}}P)f.$$

Obviously $({}^{\mathrm{co}}P)f$ is still of class C^∞, and since $\det P$ is hypoelliptic this yields that $u \in C^\infty$ as well, which ends the proof. □

Proposition 5.1.8. *If T is a constant, strictly* **a***-quasielliptic tensor then the operator $\nabla_{\mathbf{a}}^* T \nabla_{\mathbf{a}}$ is quasielliptic.*

The proof is a relatively simple calculation, but we defer it to the next section (after Lemma 5.2.3), as we will need a similar argument when dealing with strictly **a**-quasiconvex integrands.

5.1.2 Campanato regularity for linear systems

Before we proceed let us introduce some additional notation. Here we denote by $I(x_0, r)$ the anisotropic box of radius r centred at x_0 and intersected with the underlying domain of reference, i.e., Ω, which remains fixed throughout, and is assumed to be of type (A) with respect to the anisotropic metric $\delta_{\mathbf{a}}$ (see Section 2.6 for definitions). We will usually write simply I_r when there is no risk of confusion with regards to the centre. Finally, $(f)_r$ denotes the average of f over I_r.

First, let us recall two auxiliary lemmas due to Giusti (see [77]).

Lemma 5.1.9 (see Corollary 8.II in [77]). *Let T be a strictly $\mathbf{a}$-quasielliptic tensor. Let $u \in C^\infty(\overline{I_r})$ solve $\nabla_{\mathbf{a}}^* T \nabla_{\mathbf{a}} u = 0$ in I_r. Then, for any α with $\langle \alpha, \mathbf{a}^{-1} \rangle \leqslant 1$, any radius $\rho < r$, any $p \geqslant 2$, and any constant b we have*

$$\|\partial^\alpha u - (\partial^\alpha u)_\rho\|_{\mathrm{L}^p(I_\rho)}^p \leqslant C \left(\frac{\rho}{r}\right)^{|\mathbf{a}^{-1}| + (p/\max_i a_i)} \|\partial^\alpha u - b\|_{\mathrm{L}^p(I_r)}^p,$$

where the constant $C = C(\nu)$ only depends on the $\mathbf{a}$-quasiellipticity constant of T.

Lemma 5.1.10 (see Lemma 9.I in [77]). *Let T be a strictly $\mathbf{a}$-quasielliptic tensor. Let $u \in \mathrm{W}^{\mathbf{a},2}(I_r)$ be a solution of*

$$\begin{cases} \nabla_{\mathbf{a}}^* T \nabla_{\mathbf{a}} u = \sum_{\langle \alpha, \mathbf{a}^{-1} \rangle \leqslant 1} \partial^\alpha f_\alpha \text{ in } I_r, \\ u \in \mathrm{W}_0^{\mathbf{a},2}(I_r), \end{cases} \tag{5.5}$$

with $f_\alpha \in \mathrm{L}^2(I_r)$. Then for any choice of constants c_α we have

$$\|u\|_{\mathrm{W}^{\mathbf{a},2}(I_r)} \leqslant C \left(\sum_{0 < \langle \alpha, \mathbf{a}^{-1} \rangle \leqslant 1} r^{1 - \langle \alpha, \mathbf{a}^{-1} \rangle} \|f_\alpha - c_\alpha\|_{\mathrm{L}^2(I_r)} + r \|f_0\|_{\mathrm{L}^2(I_r)} \right),$$

where the constant $C = C(\nu)$ depends only on the $\mathbf{a}$-quasiellipticity constant of T.

The two lemmas above will be crucial in the following proof, which closely follows Giusti's paper [77]. However, we have decided to include it in full, as it is essential for us that, in the case of zero boundary conditions, the estimates may be carried through all the way to the boundary. Furthermore, since we are only interested in the constant coefficients case, there are several simplifications that may be made with respect to Giusti's original proof. Finally let us also remark that we explicitly include lower order terms on the right-hand side, whereas in [77] the source term is assumed to be of the form $\sum_{\langle \alpha, \mathbf{a}^{-1} \rangle = 1} \partial^\alpha f_\alpha$.

Lemma 5.1.11 (see Lemma 7.II in [77])**.** *Suppose that a function u solves*

$$\begin{cases} \nabla_{\mathbf{a}}^* T \nabla_{\mathbf{a}} u = \sum_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \partial^\alpha f_\alpha \text{ in } Q, \\ u \in \mathrm{W}_0^{\mathbf{a},2}(Q), \end{cases} \tag{5.6}$$

where T is a constant $\mathbf{a}$-quasielliptic tensor. Then, for any $x_0 \in \overline{Q}$, any choice of constants b_α and c_α, and any numbers $0 < \rho < r \leqslant 2$ we have

$$\sum_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \|\partial^\alpha u - (\partial^\alpha u)_\rho\|^2_{\mathrm{L}^2(I_\rho)} \leqslant C \left(\frac{\rho}{r}\right)^{|\mathbf{a}^{-1}|+2} \sum_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \|\partial^\alpha u - b_\alpha\|^2_{\mathrm{L}^2(I_r)} +$$
$$C \left(\sum_{0<\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} r^{2(1-\langle \alpha, \mathbf{a}^{-1}\rangle)} \|f_\alpha - c_\alpha\|^2_{\mathrm{L}^2(I_r)} + r^2 \|f_0\|_{\mathrm{L}^2(I_r)} \right),$$

with the constant $C = C(\nu)$ only depending on the $\mathbf{a}$-ellipticity constant of T.

Proof. Fix an arbitrary point x_0 and a number $r \in (0, 2]$ and consider the following equations

$$\begin{cases} \nabla_{\mathbf{a}}^* T \nabla_{\mathbf{a}} v = 0 \text{ in } I_r, \\ v - u \in \mathrm{W}_0^{\mathbf{a},2}(I_r), \end{cases} \tag{5.7}$$

$$\begin{cases} \nabla_{\mathbf{a}}^* T \nabla_{\mathbf{a}} w = \sum_{\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \partial^\alpha f_\alpha \text{ in } I_r, \\ w \in \mathrm{W}_0^{\mathbf{a},2}(I_r), \end{cases} \tag{5.8}$$

and let v, w be the solutions to the respective equations — note that we are in the $p = 2$ case, thus existence and uniqueness are taken care of by Lemmas 5.1.4 and 5.1.3. Since $u = v + w$ it is enough to estimate v and w separately.

Let us begin with v. We know, from Proposition 5.1.8, that v is of class $C^\infty(\overline{I_s})$ for any $s < r$. Thus, from Lemma 5.1.9 with $p = 2$, we know that for any radius $\rho \in (0, s)$, any multiindex α s.t. $\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1$, and any constant b_α we have

$$\|\partial^\alpha v - (\partial^\alpha v)_\rho\|^2_{\mathrm{L}^2(I_\rho)} \leqslant C \left(\frac{\rho}{s}\right)^{|\mathbf{a}^{-1}|+(2/\max_i a_i)} \|\partial^\alpha v - b_\alpha\|^2_{\mathrm{L}^2(I_s)}.$$

Since $s < r$ was arbitrary, and the right-hand side is continuous in s, we deduce that the inequality holds with $s = r$ as well, so that

$$\|\partial^\alpha v - (\partial^\alpha v)_\rho\|^2_{\mathrm{L}^2(I_\rho)} \leqslant C \left(\frac{\rho}{r}\right)^{|\mathbf{a}^{-1}|+(2/\max_i a_i)} \|\partial^\alpha v - b_\alpha\|^2_{\mathrm{L}^2(I_r)}. \tag{5.9}$$

To deal with w we simply apply Lemma 5.1.10 to write that, for any choice of constants c_α, we have

$$\|w\|^2_{\mathrm{W}^{\mathbf{a},2}(I_r)} \leqslant C \left(\sum_{0<\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} r^{2(1-\langle \alpha, \mathbf{a}^{-1}\rangle)} \|f_\alpha - c_\alpha\|^2_{\mathrm{L}^2(I_r)} + r^2 \|f_0\|^2_{\mathrm{L}^2(I_r)} \right). \tag{5.10}$$

Coming back to u and remembering that $u = v+w$ we may use this decomposition and the triangle inequality to write, for an arbitrary $\rho < r$,

$$\sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \|\partial^\alpha u - (\partial^\alpha u)_\rho\|^2_{\mathrm{L}^2(I_\rho)} \leqslant$$

$$C\left(\sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \|\partial^\alpha v - (\partial^\alpha v)_\rho\|^2_{\mathrm{L}^2(I_\rho)} + \sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \|\partial^\alpha w - (\partial^\alpha w)_\rho\|^2_{\mathrm{L}^2(I_\rho)}\right). \tag{5.11}$$

For the first term we use (5.9) to write

$$\sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \|\partial^\alpha v - (\partial^\alpha v)_\rho\|^2_{\mathrm{L}^2(I_\rho)} \leqslant \sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} C\left(\frac{\rho}{r}\right)^{|\mathbf{a}^{-1}|+(2/\max_i a_i)} \|\partial^\alpha v - b_\alpha\|^2_{\mathrm{L}^2(I_r)} \leqslant$$
$$\sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} C\left(\frac{\rho}{r}\right)^{|\mathbf{a}^{-1}|+(2/\max_i a_i)} \|\partial^\alpha u - b_\alpha\|^2_{\mathrm{L}^2(I_r)},$$

where the last inequality follows from the fact that $v = u$ on the boundary and Lemma 5.1.4. For the w term in (5.11) we simply write

$$\sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \|\partial^\alpha w - (\partial^\alpha w)_\rho\|^2_{\mathrm{L}^2(I_\rho)} \leqslant C\|w\|^2_{\mathrm{W}^{\mathbf{a},2}(I_\rho)} \leqslant C\|w\|^2_{\mathrm{W}^{\mathbf{a},2}(I_r)},$$

and then use (5.10).

Finally, plugging these two estimates back into (5.11) ends the proof. □

Theorem 5.1.12 (see Theorem 12.II in [77]). *Suppose that a function u solves*

$$\begin{cases} \nabla^*_{\mathbf{a}} T \nabla_{\mathbf{a}} u = \sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \partial^\alpha f_\alpha \ \text{in } Q, \\ u \in \mathrm{W}^{\mathbf{a},2}_0(Q), \end{cases}$$

where T is a constant strictly $\mathbf{a}$-quasielliptic tensor. Assume that all the functions f_α for $\alpha \neq 0$ are of class $\mathfrak{L}^{2,1}_{\mathbf{a}}(Q)$, and that $f_0 \in \mathrm{L}^\infty(Q)$. Then, for all α with $\langle\alpha,\mathbf{a}^{-1}\rangle \leqslant 1$, we have $\partial^\alpha u \in \mathfrak{L}^{2,1}_{\mathbf{a}}(Q)$ and

$$\sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \|\partial^\alpha u\|^2_{\mathfrak{L}^{2,1}_{\mathbf{a}}(Q)} \leqslant C\left(\|u\|^2_{\mathrm{W}^{\mathbf{a},2}(Q)} + \sum_{0<\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \|f_\alpha\|^2_{\mathfrak{L}^{2,1}_{\mathbf{a}}(Q)} + \|f_0\|_{\mathrm{L}^\infty}\right),$$

where the constant $C = C(\nu)$ only depends on the $\mathbf{a}$-ellipticity constant of T.

Proof. Fix some $x_0 \in \overline{Q}$ and define $\varphi(t) := \sum_{\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} \|\partial^\alpha u - (\partial^\alpha u)_t\|^2_{\mathrm{L}^2(I_t)}$ for $I_t = I_t(x_0)$. Using Lemma 5.1.11 with $b_\alpha := (\partial^\alpha u)_r$ and $c_\alpha := (f_\alpha)_r$ yields

$$\varphi(\rho) \leqslant C\left(\frac{\rho}{r}\right)^{|\mathbf{a}^{-1}|+(2/\max_i a_i)} \varphi(r) +$$
$$C\left(\sum_{0<\langle\alpha,\mathbf{a}^{-1}\rangle\leqslant 1} r^{2(1-\langle\alpha,\mathbf{a}^{-1}\rangle)} \|f_\alpha - (f_\alpha)_r\|^2_{\mathrm{L}^2(I_r)} + r^2\|f_0\|^2_{\mathrm{L}^2(I_r)}\right).$$

Since, by assumption, $f_\alpha \in \mathfrak{L}^{2,1}_{\mathbf{a}}(Q)$ we may, for any $\alpha \neq 0$, write

$$\|f_\alpha - (f_\alpha)_r\|^2_{\mathrm{L}^2(I_r)} \leqslant C r^{|\mathbf{a}^{-1}|} \|f_\alpha\|^2_{\mathfrak{L}^{2,1}_{\mathbf{a}}(Q)},$$

whereas for f_0 we simply estimate

$$r^2 \|f_0\|^2_{\mathrm{L}^2(I_r)} \leqslant r^{2+|\mathbf{a}^{-1}|} \|f_0\|^2_{\mathrm{L}^\infty(I_r)}.$$

Thus

$$\begin{aligned}\varphi(\rho) \leqslant & C \left(\frac{\rho}{r}\right)^{|\mathbf{a}^{-1}|+(2/\max_i a_i)} \varphi(r) + \\ & C \left(\sum_{0<\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} r^{|\mathbf{a}^{-1}|+2(1-\langle \alpha, \mathbf{a}^{-1}\rangle)} \|f_\alpha\|^2_{\mathfrak{L}^{2,1}_{\mathbf{a}}(Q)} + r^{2+|\mathbf{a}^{-1}|} \|f_0\|^2_{\mathrm{L}^\infty(I_r)} \right).\end{aligned}$$

With $A := \sum_{0<\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \|f_\alpha\|^2_{\mathfrak{L}^{2,1}_{\mathbf{a}}(Q)} + \|f_0\|^2_{\mathrm{L}^\infty(I_r)}$ we now have

$$\varphi(\rho) \leqslant C \left(\left(\frac{\rho}{r}\right)^{|\mathbf{a}^{-1}|+(2/\max_i a_i)} \varphi(r) + r^{|\mathbf{a}^{-1}|} A \right).$$

We may now conclude, exactly as in [77], that $\rho^{-|\mathbf{a}^{-1}|}\varphi(\rho)$ is bounded, thus ending the proof. □

Finally we are ready to prove the main result of this section.

Proof of Proposition 5.1.2. Clearly it is enough to show that the operators mapping $(f_\alpha)_{\alpha\neq 0}$ to the **a**-gradient of the solution of

$$\begin{cases} \nabla^*_{\mathbf{a}} T \nabla_{\mathbf{a}} u = \sum_{0<\langle \alpha, \mathbf{a}^{-1}\rangle \leqslant 1} \partial^\alpha f_\alpha \text{ in } Q, \\ u \in \mathrm{W}^{\mathbf{a},p}_0(Q), \end{cases} \tag{5.12}$$

and mapping f_0 to the **a**-gradient of the solution of

$$\begin{cases} \nabla^*_{\mathbf{a}} T \nabla_{\mathbf{a}} u = f_0 \text{ in } Q, \\ u \in \mathrm{W}^{\mathbf{a},p}_0(Q), \end{cases} \tag{5.13}$$

are of strong (p, p) type with norms bounded by a constant depending only on p and the strict **a**-quasiellipticity constant of T.

We know from Lemma 5.1.3 that both operators are of strong type $(2, 2)$. On the other hand, Theorem 5.1.12 tells us that the operator mapping $(f_\alpha)_\alpha$ to the **a**-gradient of the solution of (5.12) maps $\mathfrak{L}^{2,1}_{\mathbf{a}}(Q)$ to $\mathfrak{L}^{2,1}_{\mathbf{a}}(Q)$ continuously. The same estimate shows that the operator given by (5.13) is continuous as a mapping from $\mathrm{L}^\infty(Q)$ to $\mathfrak{L}^{2,1}_{\mathbf{a}}(Q)$. Hence, the Stampacchia interpolation theorem (see Theorem 2.14 in [79]) lets us conclude that both operators are of strong type (p, p), and thus so is their sum, which ends the proof. □

The last result of the linear theory we will need is the following, proven by Giusti in [77]:

Theorem 5.1.13 (see Theorem 12.3 in [77]). *Let T be a constant, strictly $\mathbf{a}$-quasielliptic tensor and suppose that $u \in \mathrm{W}^{\mathbf{a},2}(Q)$ is a weak solution of*

$$\nabla_{\mathbf{a}}^{*} T \nabla_{\mathbf{a}} u = 0 \text{ in } Q.$$

Then, for any $\theta \in (1, 1+2/|\mathbf{a}^{-1}|)$, any subdomain $\omega \Subset Q$ and any α with $\langle \alpha, \mathbf{a}^{-1} \rangle \leqslant 1$ we have $\partial^{\alpha} u \in \mathfrak{L}_{\mathbf{a}}^{2,\theta}(\omega)$ with the estimate

$$\sum_{\langle \alpha, \mathbf{a}^{-1} \rangle \leqslant 1} \|\partial^{\alpha} u\|_{\mathfrak{L}_{\mathbf{a}}^{2,\theta}(\omega)} \leqslant C \|u\|_{\mathrm{W}^{\mathbf{a},2}(Q)},$$

where the constant $C = C(\nu, \omega, \theta)$ depends only on the strict $\mathbf{a}$-quasiellipticity constant of T, the subdomain ω, and the exponent θ.

5.2 Strict a-quasiconvexity

With the linear theory at hand, we may now start building towards regularity of minimisers of variational problems. Firstly we introduce the following strengthening of $\mathbf{a}$-quasiconvexity.

Definition 5.2.1. *We say that a function $F \colon \mathbb{R}^{n \times m} \to \mathbb{R}$ satisfying $|F(W)| \leqslant C(1 + |W|^p)$ for some constant C and $p \in (1, \infty)$ is strictly $\mathrm{W}^{\mathbf{a},p}$-quasiconvex if for every $W \in \mathbb{R}^{n \times m}$ and all $\varphi \in C_c^{\infty}(Q)$ one has*

$$F(W) + \nu \|\nabla_{\mathbf{a}} \varphi\|_{\mathrm{L}^p(Q)}^p \leqslant \fint_Q F(W + \nabla_{\mathbf{a}} \varphi(x))\, \mathrm{d}x,$$

for some constant $v > 0$ independent of W and φ.

Proposition 5.2.2. *Suppose that $F \colon \mathbb{R}^{n \times m} \to \mathbb{R}$ satisfies $|F(W)| \leqslant C(1 + |W|^2)$ for some constant C and that F is strictly $\mathrm{W}^{\mathbf{a},2}$-quasiconvex. Assume furthermore that F is of class C^2 with bounded second derivative. Then, for every $W_0 \in \mathbb{R}^{n \times m}$,*

$$c \left(\sum_{\langle \alpha, \mathbf{a}^{-1} \rangle = 1} x^{2\alpha} \right) \|b\|_2^2 \leqslant \sum_{\alpha, \widetilde{\alpha} \in A_e} \sum_{1 \leqslant i,j \leqslant n} \partial^{(\alpha,i)} \partial^{(\widetilde{\alpha},j)} F(W_0) (-1)^{(|\alpha| + |\widetilde{\alpha}|)/2} x^{\alpha + \widetilde{\alpha}} b_i b_j + \\ \sum_{\beta, \widetilde{\beta} \in A_o} \sum_{1 \leqslant i,j \leqslant n} \partial^{(\beta,i)} \partial^{(\widetilde{\beta},j)} F(W_0) (-1)^{(|\beta| + |\widetilde{\beta}| + 2)/2} x^{\beta + \widetilde{\beta}} b_i b_j,$$

for some positive constant c depending only on the $\mathrm{W}^{\mathbf{a},2}$-quasiconvexity constant of F and for any $x \in \mathbb{R}^N$ and $b \in \mathbb{R}^n$. Here A_e and A_o denote the subsets of the set $\{\alpha \colon \langle \alpha, \mathbf{a}^{-1} \rangle = 1\}$ with orders of even and odd parities respectively.

Proof. We proceed as in [93] for the proof of rank-one convexity. The only difference is that we require strict $W^{\mathbf{a},2}$-quasiconvexity rather than just quasiconvexity and we allow the multiindices to have orders of different parities.

We use a test function of the form

$$\varphi\colon Q \ni \xi \mapsto t^{-1} \cos(\sum_{j=1}^{N} t^{1/a_j} x_j \xi_j) b,$$

where $x \in \mathbb{R}^N$ and $b \in \mathbb{R}^n$ are fixed. This should be multiplied by a cut-off to ensure that it is an admissible test function for strict $W^{\mathbf{a},2}$-quasiconvexity, however we omit this technical detail for the sake of clarity of presentation — it may be dealt with exactly as in [93]. For $\langle \alpha, \mathbf{a}^{-1} \rangle = 1$ with $|\alpha|$ even we have

$$\partial^\alpha \varphi(\xi) = (-1)^{|\alpha|/2} x^\alpha \cos(\sum_{1}^{N} t^{1/a_j} x_j \xi_j) b$$

and for $\langle \beta, \mathbf{a}^{-1} \rangle = 1$ with $|\beta|$ odd we, in turn, have

$$\partial^\beta \varphi(\xi) = (-1)^{(|\beta|+1)/2} x^\beta \sin(\sum_{1}^{N} t^{1/a_j} x_j \xi_j) b.$$

Using strict $W^{\mathbf{a},2}$-quasiconvexity of F we may write

$$F(W) + \nu ||\lambda \nabla_{\mathbf{a}} \varphi||_2^2 \leqslant \int_Q F(W + \lambda \nabla_{\mathbf{a}} \varphi(\xi)) \,\mathrm{d}\xi,$$

for any $W \in \mathbb{R}^{n\times m}$ — thus we test with $\lambda\varphi$ rather than just φ. Now we may multiply this inequality pointwise by a test function $\psi \in C_c^\infty(\mathbb{R}^{n\times m})$ with $\psi(W) \geqslant 0$ and integrate over $\mathbb{R}^{n\times m}$. This yields

$$\int_{\mathbb{R}^{n\times m}} F(W)\psi(W) \,\mathrm{d}W + \lambda^2 \nu \|\nabla_{\mathbf{a}} \varphi\|_2^2 \int_{\mathbb{R}^{n\times m}} \psi(W) \,\mathrm{d}W \leqslant$$

$$\int_{\mathbb{R}^{n\times m}} \int_Q F(W + \lambda \nabla_{\mathbf{a}} \varphi(\xi)) \psi(W) \,\mathrm{d}\xi \,\mathrm{d}W.$$

Using a Taylor expansion on F on the right-hand side we may rewrite it as

$$\int_{\mathbb{R}^{n\times m}} \int_Q F(W)\psi(W) + \langle F'(W), \lambda \nabla_{\mathbf{a}} \varphi(\xi) \rangle \psi(W) \,\mathrm{d}\xi \,\mathrm{d}W +$$

$$\int_{\mathbb{R}^{n\times m}} \int_Q \frac{1}{2} \left[F''(W); \lambda \nabla_{\mathbf{a}} \varphi(\xi), \lambda \nabla_{\mathbf{a}} \varphi(\xi)\right] \psi(W) + o(\lambda^2)\psi(W) \,\mathrm{d}\xi \,\mathrm{d}W.$$

The zero order term cancels out with its counterpart on the left-hand side of our inequality. For the two others we note, as in [93], that

$$Q \ni \xi \mapsto (-1)^{|\alpha|/2} x^\alpha \cos(\sum_1^N t^{1/a_j} x_j \xi_j) b$$

is equimeasurable with

$$[0,1] \ni y \mapsto (-1)^{|\alpha|/2} x^\alpha \cos(2\pi y) b,$$

and

$$Q \ni \xi \mapsto (-1)^{(|\beta|+1)/2} x^\beta \sin(\sum_1^N t^{1/a_j} x_j \xi_j) b$$

is equimeasurable with

$$[0,1] \ni y \mapsto (-1)^{(|\beta|+1)/2} x^\beta \sin(2\pi y) b.$$

From this we easily see that the first order terms integrate to 0 in ξ. The second order terms all have a factor of the form $\partial^{(\alpha,i)}\partial^{(\beta,j)}F(W)$ multiplied by an appropriate derivative of φ and are essentialy of two kinds. If the parities of α and β are different, say α is even and β is odd, we have

$$\frac{\lambda^2}{2}\int_{\mathbb{R}^{n\times m}}\int_0^1 \partial^{(\alpha,i)}\partial^{\beta,j}F(W)(-1)^{(|\alpha|+|\beta|+1)/2}x^{\alpha+\beta}\cos(2\pi y)\sin(2\pi y)b_i b_j \psi(W)\,\mathrm{d}y\,\mathrm{d}W,$$

and we see that this integrates to 0 in y. On the other hand, when the parities match, say both α and β are even, the corresponding term is

$$\frac{\lambda^2}{2}\int_{\mathbb{R}^{n\times m}}\int_0^1 \partial^{(\alpha,i)}\partial^{\beta,j}F(W)(-1)^{(|\alpha|+|\beta|)/2}x^{\alpha+\beta}\cos^2(2\pi y)b_i b_j \psi(W)\,\mathrm{d}y\,\mathrm{d}W,$$

and similarly for α, β odd. Now the integral in y is a strictly positive absolute constant, thus the last thing we need to note is that a simple calculation shows that $\|\nabla_{\mathbf{a}}\varphi\|^2 = \widetilde{c}\left(\sum_{\langle\alpha,\mathbf{a}^{-1}\rangle=1} x^{2\alpha}\right)\|b\|_2^2$. We may now plug those two facts into our inequality, divide both sides by λ^2 and take the limit as $\lambda \to 0$. This eliminates the remainder term (because we assume that F'' is bounded) and shows that the inequality in our proposition holds in the sense of distributions, and thus, given the assumption that F is C^2, in the classical sense as well. $\square$

Lemma 5.2.3. *A constant, symmetric tensor T parametrised by pairs $((\alpha, j), (\widetilde{\alpha}, \widetilde{j}))$ with $\langle \alpha, \mathbf{a}^{-1}\rangle = \langle \widetilde{\alpha}, \mathbf{a}^{-1}\rangle = 1$ and $j, \widetilde{j} \in \{1, \ldots, n\}$ is strictly quasielliptic with quasiellipticity constant ν if and only if, for all $x \in \mathbb{R}^N$ and all $b \in \mathbb{R}^n$,*

$$\nu \left(\sum_{\langle \alpha, \mathbf{a}^{-1}\rangle = 1} x^{2\alpha} \right) \|b\|_2^2 \leqslant \sum_{\alpha, \widetilde{\alpha} \in A_e} \sum_{1 \leqslant j, \widetilde{j} \leqslant n} T^{(\alpha,j)}_{(\widetilde{\alpha},\widetilde{j})} (-1)^{(|\alpha|+|\widetilde{\alpha}|)/2} x^{\alpha+\widetilde{\alpha}} b_j b_{\widetilde{j}} + \sum_{\alpha, \widetilde{\alpha} \in A_o} \sum_{1 \leqslant j, \widetilde{j} \leqslant n} T^{(\alpha,j)}_{(\widetilde{\alpha},\widetilde{j})} (-1)^{(|\alpha|+|\widetilde{\alpha}|+2)/2} x^{\alpha+\widetilde{\alpha}} b_j b_{\widetilde{j}}, \tag{5.14}$$

where we have let A_e and A_o be the subsets of $\{\alpha \colon \langle \alpha, \mathbf{a}^{-1}\rangle = 1\}$ of even and odd degrees respectively.

Proof. The 'only if' part is an immediate corollary of Proposition 5.2.2. For the 'if' direction we proceed similarly to how one proves that rank-one convexity implies quasiconvexity for quadratic forms in the classical case (see for example [40]). Indeed, the inequality (5.14) is the mixed smoothness analogue of the Legendre-Hadamard condition.

Let T be a tensor satisfying our assumptions, and let us fix an arbitrary function $u \in \mathrm{W}^{\mathbf{a},2}_0(Q)$. We may extend u by 0 to the whole of $\mathbb{R}^N$ and, using the fact that the Fourier transform is an isometry on $\mathrm{L}^2(\mathbb{R}^N)$, write

$$\begin{aligned} \int_Q \langle T\nabla_{\mathbf{a}} u, \nabla_{\mathbf{a}} u\rangle \,\mathrm{d}x &= \int_{\mathbb{R}^N} \langle T\nabla_{\mathbf{a}} u, \nabla_{\mathbf{a}} u\rangle \,\mathrm{d}x = \int_{\mathbb{R}^N} \langle \mathcal{F}(T\nabla_{\mathbf{a}} u), \mathcal{F}(\nabla_{\mathbf{a}} u)\rangle \,\mathrm{d}\xi \\ &= \int_{\mathbb{R}^N} \sum_{\substack{(\alpha,j) \\ (\widetilde{\alpha},\widetilde{j})}} T^{\alpha,j}_{\widetilde{\alpha},\widetilde{j}} (i\xi)^\alpha \mathcal{F}(u^j) \overline{(i\xi)^{\widetilde{\alpha}} \mathcal{F}(u^{\widetilde{j}})} \,\mathrm{d}\xi, \end{aligned}$$

where $\mathcal{F}(f)$ denotes the Fourier transform of f and $\overline{z}$ is the complex conjugate of z. Since T is symmetric the terms where the orders of α and $\widetilde{\alpha}$ have different parities cancel out due to the $i^{|\alpha|}(-i)^{|\widetilde{\alpha}|}$ term. Thus, we are left with exactly the expression on the left-hand side of (5.14) with $x = \xi$ and $b = \mathcal{F}(u)$. Applying the inequality (5.14) separately to real and imaginary parts of $\mathcal{F}(u)$ we may estimate with

$$\nu \int_{\mathbb{R}^N} \left(\sum_{\langle \alpha, \mathbf{a}^{-1}\rangle = 1} \xi^{2\alpha} \right) \mathcal{F}(u)\overline{\mathcal{F}(u)} \,\mathrm{d}\xi = \nu \int_{\mathbb{R}^N} \sum_{\langle \alpha, \mathbf{a}^{-1}\rangle = 1} \mathcal{F}(\partial^\alpha u)\overline{\mathcal{F}(\partial^\alpha u)} \,\mathrm{d}\xi,$$

and using Plancherel's identity again ends the proof. □

At this point let us remark that the above lemma immediately implies Proposition 5.1.8.

Proof of Proposition 5.1.8. Since the operator $\nabla_{\mathbf{a}}^* T \nabla_{\mathbf{a}}$ only involves derivatives $\partial^{\alpha+\beta}$ on the hyperplane of homogeneity given through $\langle \alpha+\beta, \mathbf{a}^{-1} \rangle = 2$ the only thing that we need to check is that its determinant is non-zero. To do so we write

$$\nabla_{\mathbf{a}}^* T \nabla_{\mathbf{a}}(-i\xi) = \sum_{\substack{\langle \alpha, \mathbf{a}^{-1} \rangle = 1 \\ \langle \widetilde{\alpha}, \mathbf{a}^{-1} \rangle = 1}} T_{\widetilde{\alpha}}^{\alpha} (-i\xi)^{\alpha} (-1)^{|\widetilde{\alpha}|} (-i\xi)^{\widetilde{\alpha}}.$$

As noted before, when α and $\widetilde{\alpha}$ have orders of different parities the $(\alpha, \widetilde{\alpha})$ term cancels the $(\widetilde{\alpha}, \alpha)$ one, due to the $(-1)^{|\widetilde{\alpha}|}$ factor coming from the adjoint $\nabla_{\mathbf{a}}^*$. Hence, with A_e and A_o as in the lemma above,

$$\nabla_{\mathbf{a}}^* T \nabla_{\mathbf{a}}(-i\xi) = \sum_{\alpha, \widetilde{\alpha} \in A_e} T_{\widetilde{\alpha}}^{\alpha} (-1)^{(|\alpha|+|\widetilde{\alpha}|)/2} \xi^{\alpha+\widetilde{\alpha}} + \sum_{\alpha, \widetilde{\alpha} \in A_o} T_{\widetilde{\alpha}}^{\alpha} (-1)^{(|\alpha|+|\widetilde{\alpha}|+2)/2} \xi^{\alpha+\widetilde{\alpha}}.$$

Therefore, the inequality (5.14) may be rewritten as

$$b^T \left(\nabla_{\mathbf{a}}^* T \nabla_{\mathbf{a}}(-i\xi) \right) b \geqslant \nu \left(\sum_{\langle \alpha, \mathbf{a}^{-1} \rangle = 1} \xi^{2\alpha} \right) \|b\|_2^2 > 0$$

for every $0 \neq b \in \mathbb{R}^n$. This means that $\nabla_{\mathbf{a}}^* T \nabla_{\mathbf{a}}(-i\xi)$ is a positive definite matrix, from which we infer that its determinant is strictly positive, thus ending the proof. ☐

5.2.1 Caccioppoli inequality

Let b be a polynomial with $\nabla_{\mathbf{a}} b \equiv \text{const}$ (so an **a**-polynomial) and let u be a minimiser of

$$v \mapsto \int_{\Omega} F(\nabla_{\mathbf{a}} v) \, \mathrm{d}x,$$

over some fixed Dirichlet class $\mathrm{W}_g^{\mathbf{a},2}(\Omega)$, where F is a strictly $\mathrm{W}^{\mathbf{a},2}$-quasiconvex integrand of class C^2 with a bounded second derivative. We wish to show:

Lemma 5.2.4. *There exists a constant $C = C(\nu, \|F''\|_{\mathrm{L}^\infty})$ depending only on the strict $\mathrm{W}^{\mathbf{a},2}$-quasiconvexity constant of F and on $\|F''\|_{\mathrm{L}^\infty}$, and such that*

$$\fint_{Q_{r/2}} |\nabla_{\mathbf{a}}(u-b)|^2 \, \mathrm{d}x \leqslant C \sum_{\langle \beta, \mathbf{a}^{-1} \rangle < 1} r^{-2(1-\langle \beta, \mathbf{a}^{-1} \rangle)} \fint_{Q_r} |\partial^{\beta}(u-b)|^2,$$

for any anisotropic box $Q_r \Subset \Omega$.

Proof. Assume for notational simplicity that the centre of Q_r is at 0 and let $r/2 \leqslant t < s \leqslant r$. Fix a cut-off function $\eta \in C_c^\infty(Q_s)$ with $\eta = 1$ on Q_t and satisfying

$$\|\partial^{\beta} \eta\|_{\mathrm{L}^\infty(Q_s)} \leqslant C \prod_{i=1}^{N} \left(s^{1/a_i} - t^{1/a_i} \right)^{-\beta_i},$$

for all $\langle \beta, \mathbf{a}^{-1} \rangle \leqslant 1$. Observe that $s^{1/a_i} - t^{1/a_i}$ is the distance between the faces of Q_s and Q_t along the x_i axis, thus such a cut-off may be constructed as a product of one dimensional cut-offs (compare with the proof of Lemma 2.5.1). Let

$$\varphi := \eta(u - b) \quad \text{ and } \quad \psi := (1 - \eta)(u - b),$$

so that

$$\nabla_{\mathbf{a}}\varphi + \nabla_{\mathbf{a}}\psi = \nabla_{\mathbf{a}}u - \nabla_{\mathbf{a}}b = \nabla_{\mathbf{a}}u - B,$$

with $B := \nabla_{\mathbf{a}}b$. We have $\varphi = 0$ on ∂Q_s, so strict $\mathbf{a}$-quasiconvexity of F implies that

$$\int_{Q_s} F(B) + \gamma|\nabla_{\mathbf{a}}\varphi|^2 \, \mathrm{d}x \leqslant \int_{Q_s} F(B + \nabla_{\mathbf{a}}\varphi) \, \mathrm{d}x = \int_{Q_s} F(\nabla_{\mathbf{a}}u - \nabla_{\mathbf{a}}\psi) \, \mathrm{d}x.$$

Estimating the last integral using, at each point, Taylor expansion of F around $\nabla_{\mathbf{a}}u$, and boundedness of F'' we may write

$$\int_{Q_s} F(B) + \gamma|\nabla_{\mathbf{a}}\varphi|^2 \, \mathrm{d}x \leqslant \int_{Q_s} F(\nabla_{\mathbf{a}}u) - F'(\nabla_{\mathbf{a}}u)\nabla_{\mathbf{a}}\psi + C|\nabla_{\mathbf{a}}\psi|^2 \, \mathrm{d}x. \tag{5.15}$$

On the other hand, u is a minimiser and $\varphi = 0$ on ∂Q_s, so that $u - \varphi$ is an admissible competitor, and we may write

$$\int_{Q_s} F(\nabla_{\mathbf{a}}u) \, \mathrm{d}x \leqslant \int_{Q_s} F(\nabla_{\mathbf{a}}u - \nabla_{\mathbf{a}}\varphi) \, \mathrm{d}x = \int_{Q_s} F(\nabla_{\mathbf{a}}\psi + B) \, \mathrm{d}x.$$

Taking a Taylor expansion around B and using boundedness of F'' yields

$$\int_{Q_s} F(\nabla_{\mathbf{a}}u) \, \mathrm{d}x \leqslant \int_{Q_s} F(B) + F'(B)\nabla_{\mathbf{a}}\psi + C|\nabla_{\mathbf{a}}\psi|^2 \, \mathrm{d}x. \tag{5.16}$$

Putting (5.15) and (5.16) together we get

$$\int_{Q_s} \gamma|\nabla_{\mathbf{a}}\varphi|^2 \, \mathrm{d}x \leqslant \int_{Q_s} \left(F'(B) - F'(\nabla_{\mathbf{a}}u)\right) \nabla_{\mathbf{a}}\psi + C|\nabla_{\mathbf{a}}\psi|^2 \, \mathrm{d}x.$$

Boundedness of F'' implies that F' is Lipschitz, which we may use to estimate the first term on the right-hand side. Moreover, we may restrict the integral on the left-hand side to Q_t, where $\varphi = u - b$, which gives

$$\int_{Q_t} |\nabla_{\mathbf{a}}(u - b)|^2 \, \mathrm{d}x \leqslant C \int_{Q_s} |B - \nabla_{\mathbf{a}}u| \cdot |\nabla_{\mathbf{a}}\psi| + |\nabla_{\mathbf{a}}\psi|^2 \, \mathrm{d}x.$$

The integrals on the right-hand side may be restricted to $Q_s \setminus Q_t$, as $\psi = 0$ on Q_t. Moreover, we may use the Cauchy-Schwarz inequality on the first term on the right-hand side and write

$$\int_{Q_t} |\nabla_{\mathbf{a}}(u - b)|^2 \, \mathrm{d}x \leqslant C \int_{Q_s \setminus Q_t} |\nabla_{\mathbf{a}}(u - b)|^2 + |\nabla_{\mathbf{a}}\psi|^2 \, \mathrm{d}x. \tag{5.17}$$

Now we need an estimate on $|\nabla_{\mathbf{a}}\psi|$. For this observe that, for any α such that $\langle \alpha, \mathbf{a}^{-1}\rangle = 1$, we have

$$|\partial^\alpha \psi| \leqslant |\partial^\alpha (u-b)| + \sum_{0<\beta\leqslant\alpha} C \left(\prod_{i=1}^N \left(s^{1/a_i} - t^{1/a_i}\right)^{-\beta_i} \right) |\partial^{\alpha-\beta}(u-b)|.$$

Thus

$$\int_{Q_s\setminus Q_t} |\nabla_{\mathbf{a}}\psi|^2 \leqslant C \int_{Q_s\setminus Q_t} |\nabla_{\mathbf{a}}(u-b)|^2 + C \sum_{\langle\alpha,\mathbf{a}^{-1}\rangle=1} \sum_{0<\beta\leqslant\alpha} \left(\prod_{i=1}^N \left(s^{1/a_i} - t^{1/a_i}\right)^{-2\beta_i} \right) |\partial^{\alpha-\beta}(u-b)|^2 \,\mathrm{d}x.$$

Using this in (5.17) we write

$$\int_{Q_t} |\nabla_{\mathbf{a}}(u-b)|^2 \,\mathrm{d}x \leqslant C \int_{Q_s\setminus Q_t} |\nabla_{\mathbf{a}}(u-b)|^2 \,\mathrm{d}x + C \int_{Q_s\setminus Q_t} \sum_{\langle\alpha,\mathbf{a}^{-1}\rangle=1} \sum_{0<\beta\leqslant\alpha} \left(\prod_{i=1}^N \left(s^{1/a_i} - t^{1/a_i}\right)^{-2\beta_i} \right) |\partial^{\alpha-\beta}(u-b)|^2 \,\mathrm{d}x.$$

We may now add $C\int_{Q_t} |\nabla_{\mathbf{a}}(u-b)|^2 \,\mathrm{d}x$ to both sides to fill the hole in the first integral on the right-hand side. Let us also extend the second integral on the right-hand side to the whole of Q_r and thus, with $\theta := \frac{C}{C+1} < 1$, write

$$\int_{Q_t} |\nabla_{\mathbf{a}}(u-b)|^2 \,\mathrm{d}x \leqslant \theta \int_{Q_s} |\nabla_{\mathbf{a}}(u-b)|^2 \,\mathrm{d}x + C \sum_{\langle\alpha,\mathbf{a}^{-1}\rangle=1} \sum_{0<\beta\leqslant\alpha} \left(\prod_{i=1}^N \left(s^{1/a_i} - t^{1/a_i}\right)^{-2\beta_i} \right) \int_{Q_r} |\partial^{\alpha-\beta}(u-b)|^2 \,\mathrm{d}x.$$

To conclude we just use the lemma below. □

Lemma 5.2.5. *Let $f\colon [r/2, r] \to [0,\infty)$ be a bounded function satisfying, for some constants c_β,*

$$f(t) \leqslant \theta f(s) + \sum_{0<\langle\beta,\mathbf{a}^{-1}\rangle\leqslant 1} c_\beta \left(\prod_{i=1}^N \left(s^{1/a_i} - t^{1/a_i}\right)^{-2\beta_i} \right),$$

for some $\theta < 1$ and any $r/2 \leqslant t < s \leqslant r$. Then there exists a constant $C = C(\theta, \mathbf{a})$ such that

$$f(r/2) \leqslant C \sum_{0<\langle\beta,\mathbf{a}^{-1}\rangle\leqslant 1} c_\beta r^{-2\langle\beta,\mathbf{a}^{-1}\rangle}.$$

Proof. Observe that for any i and small enough x we have

$$(1+x)^{1/a_i} - 1 \geqslant \frac{1}{2a_i}x,$$

which results simply from comparing the derivatives, as long as x is small enough so that $(1+x)^{1-1/a_i} \geqslant 1/2$. Hence we deduce that

$$\left(s^{1/a_i} - t^{1/a_i}\right) = t^{1/a_i}\left((1+\frac{s-t}{t})^{1/a_i} - 1\right) \geqslant C(\frac{r}{2})^{1/a_i}\frac{s-t}{t}, \tag{5.18}$$

with $0 < C$ dependent only on $\mathbf{a}$ and provided that $\frac{s-t}{t}$ is small enough. Finally, using this estimate in our assumption we may write

$$f(t) \leqslant \theta f(s) + C \sum_{0<\langle\beta,\mathbf{a}^{-1}\rangle\leqslant 1} c_\beta(\frac{r}{2})^{-2\langle\beta,\mathbf{a}^{-1}\rangle}(\frac{s-t}{t})^{-2|\beta|}.$$

Now set $t_k := r(1-\frac{\tau^k}{2})$ for $k = 0, 1, \ldots$ and a constant $\tau < 1$ to be determined later. We have $t_0 = r/2$ and $t_k \to r$ as $k \to \infty$ and $t_{k+1} - t_k = \frac{r}{2}(1-\tau)\tau^k$. Plugging $s = t_{k+1}$ and $t := t_k$ in the last inequality we get

$$f(t_k) \leqslant \theta f(t_{k+1}) + C \sum_{0<\langle\beta,\mathbf{a}^{-1}\rangle\leqslant 1} c_\beta(\frac{r}{2})^{-2\langle\beta,\mathbf{a}^{-1}\rangle}\left(\frac{\frac{r}{2}(1-\tau)\tau^k}{r(1-\frac{\tau^k}{2})}\right)^{-2|\beta|}.$$

Simplifying further we may write

$$f(t_k) \leqslant \theta f(t_{k+1}) + C(1-\tau)^{-2\max|\beta|} \sum_{0<\langle\beta,\mathbf{a}^{-1}\rangle\leqslant 1} c_\beta r^{-2\langle\beta,\mathbf{a}^{-1}\rangle}\tau^{-k2|\beta|}.$$

Starting with $k = 0$ and iteratively using this inequality m times we get

$$f(t_0) \leqslant \theta^m f(t_m) + C(1-\tau)^{-2\max|\beta|} \sum_{0<\langle\beta,\mathbf{a}^{-1}\rangle\leqslant 1} c_\beta r^{-2\langle\beta,\mathbf{a}^{-1}\rangle} \sum_{k=0}^{m-1} \theta^k \tau^{-k2|\beta|}.$$

Now we want to pass to the limit as $m \to \infty$, and all we need to ensure is the convergence of the series $\sum_{k=0}^{\infty} \theta^k \tau^{-k2|\beta|}$ for β's such that $\langle\beta,\mathbf{a}^{-1}\rangle \leqslant 1$. This is easily done by taking τ close enough to 1 so that $\theta < \tau^{-2|\beta|}$ for all relevant β's. For such a choice of τ, as a function of θ and $\mathbf{a}$, the series is convergent and its sum is again a function of θ and $\mathbf{a}$, so that in the limit we get

$$f(t_0) \leqslant C(\theta, a) \sum_{0<\langle\beta,\mathbf{a}^{-1}\rangle\leqslant 1} c_\beta r^{-2\langle\beta,\mathbf{a}^{-1}\rangle}.$$

The last point we wish to note is that, with $\tau < 1$, for every k we have

$$\frac{t_{k+1}-t_k}{t_k} = \frac{\frac{r}{2}(1-\tau)\tau^k}{r(1-\frac{\tau^k}{2})} \leqslant (1-\tau)\tau^k \leqslant (1-\tau),$$

so that for τ sufficiently close to 1 the use of the estimate (5.18) is justified, and thus the proof is complete. $\square$

5.3 Partial regularity of minimisers

We continue under the assumption that F is a quadratic growth integrand with bounded second derivatives. Assume that the domain Ω satisfies the weak $\mathbf{a}$-horn condition and that u is a minimiser of $\int_\Omega F(\nabla_{\mathbf{a}} u)\,\mathrm{d}x$ over some Dirichlet class $\mathrm{W}^{\mathbf{a},2}_{u_b}(\Omega)$, where $u_b \in \mathrm{W}^{\mathbf{a},2}(\mathbb{R}^N)$ is a given boundary datum. In this section we are interested in establishing higher regularity results for u, in particular Hölder continuity of its $\mathbf{a}$-gradient, up to a set of measure zero.

5.3.1 The a-harmonic approximation

To begin with, let us fix a number $M > 0$ and an arbitrary box $Q_R \subset \Omega$ such that, with $W_0 := (\nabla_{\mathbf{a}} u)_{Q_R} = \fint_{Q_R} \nabla_{\mathbf{a}} u\,\mathrm{d}x$ denoting the average of $\nabla_{\mathbf{a}} u$ over Q_R, we have $|W_0| < M$. Fix a polynomial P_{W_0} satisfying $\nabla_{\mathbf{a}} P_{W_0} \equiv W_0$ and denote $\widetilde{u} := u - P_{W_0}$. Let us take $h \in \mathrm{W}^{\mathbf{a},2}_{\widetilde{u}}(Q_R)$ to be the unique solution of

$$\begin{cases} \nabla_{\mathbf{a}}^* F''(W_0)\nabla_{\mathbf{a}} h = 0 \text{ in } Q_R, \\ h - \widetilde{u} \in \mathrm{W}^{\mathbf{a},2}_0(Q_R). \end{cases} \tag{5.19}$$

We recall that Lemma 5.1.4 shows that the function h is well-defined and, due to Proposition 5.1.8, C^∞ smooth in the interior of Q_R. Furthermore, directly from the definition of a weak solution we obtain that

$$\int_{Q_R} \langle F''(W_0)\nabla_{\mathbf{a}} h, \nabla_{\mathbf{a}}\varphi\rangle = 0$$

for all $\varphi \in C_c^\infty(Q_R)$. Since u is a minimiser of our functional it solves the Euler-Lagrange equations, thus for all $\varphi \in C_c^\infty(Q_R)$ we have

$$\int_{Q_R} \langle F'(\nabla_{\mathbf{a}} u), \nabla_{\mathbf{a}}\varphi\rangle\,\mathrm{d}x = 0.$$

Similarly,

$$\int_{Q_R} \langle F'(W_0), \nabla_{\mathbf{a}}\varphi\rangle\,\mathrm{d}x = 0,$$

because $F'(W_0)$ is constant and φ is zero on the boundary. Note here that, in fact, a standard density argument shows that we may test these two equations with any $\varphi \in \mathrm{W}^{\mathbf{a},2}_0(Q_R)$, not necessarily of class $C_c^\infty(Q_R)$. Thus, for any $\varphi \in \mathrm{W}^{\mathbf{a},2}_0(Q_R)$ we may write

$$\int_{Q_R} \langle F''(W_0)\nabla_{\mathbf{a}}\widetilde{u}, \nabla_{\mathbf{a}}\varphi\rangle\,\mathrm{d}x = \int_{Q_R} \langle F''(W_0)\nabla_{\mathbf{a}}\widetilde{u} - (F'(\nabla_{\mathbf{a}} u) - F'(W_0)), \nabla_{\mathbf{a}}\varphi\rangle\,\mathrm{d}x =$$
$$\int_{Q_R} -\int_0^1 \langle [F''(W_0 + t\nabla_{\mathbf{a}}\widetilde{u}) - F''(W_0)]\,\nabla_{\mathbf{a}}\widetilde{u}, \nabla_{\mathbf{a}}\varphi\rangle\,\mathrm{d}t\,\mathrm{d}x, \tag{5.20}$$

where the last equality is just the fundamental theorem of calculus.

We now wish to estimate the difference $F''(W_0 + t\nabla_{\mathbf{a}}\widetilde{u}) - F''(W_0)$ in terms of $\nabla_{\mathbf{a}}\widetilde{u}$. We assume F'' to be continuous, hence it is uniformly continuous on the ball $B_{2M}(0) \subset \mathbb{R}^{n\times m}$, so it admits there a finite modulus of continuity, that we may assume to be concave and increasing. Finally, composing this modulus of continuity with the square root we ensure the existence of a concave and increasing function $\overline{\omega}_M$ such that

$$|F''(W_0 + t\nabla_{\mathbf{a}}\widetilde{u}(x)) - F''(W_0)| \leqslant \overline{\omega}_M(|t\nabla_{\mathbf{a}}\widetilde{u}(x)|^2)$$

whenever

$$|W_0 + t\nabla_{\mathbf{a}}\widetilde{u}(x)| \leqslant 2M.$$

To get rid of this last restriction we note that when $|W_0 + t\nabla_{\mathbf{a}}\widetilde{u}(x)| > 2M$ then necessarily $|t\nabla_{\mathbf{a}}\widetilde{u}(x)| > M$ as well. However, F'' is assumed to be bounded, thus, increasing the function $\overline{\omega}_M$ if necessary (so that it is larger than $2\|F''\|_{\mathrm{L}^\infty}$ at M^2), we ensure that there exists a concave function $\omega_M\colon [0,\infty) \to [0,\infty)$ bounded above by a constant depending only on $\|F''\|_{\mathrm{L}^\infty}$, satisfying $\omega_M(0) = 0$, and such that

$$|F''(W_0 + t\nabla_{\mathbf{a}}\widetilde{u}(x)) - F''(W_0)| \leqslant \omega_M(|t\nabla_{\mathbf{a}}\widetilde{u}(x)|^2)$$

for all t and x. Note that the function ω_M only depends on M and F'', in particular it does not depend on W_0.

Plugging this estimate into (5.20) yields

$$\left|\int_{Q_R} \langle F''(W_0)\nabla_{\mathbf{a}}\widetilde{u}, \nabla_{\mathbf{a}}\varphi\rangle \,\mathrm{d}x\right| \leqslant \int_{Q_R}\int_0^1 \omega_K(t^2|\nabla_{\mathbf{a}}\widetilde{u}|^2)\cdot|\nabla_{\mathbf{a}}\widetilde{u}|\cdot|\nabla_{\mathbf{a}}\varphi|\,\mathrm{d}t\,\mathrm{d}x \leqslant \int_{Q_R} \omega_K(|\nabla_{\mathbf{a}}\widetilde{u}|^2)\cdot|\nabla_{\mathbf{a}}\widetilde{u}|\cdot|\nabla_{\mathbf{a}}\varphi|\,\mathrm{d}x.$$

Recall that $\nabla_{\mathbf{a}}\widetilde{u} = \nabla_{\mathbf{a}}u - W_0$, but $F''(W_0)W_0$ is a constant, so it does not contribute to the integral on the left-hand side. Neither would h, as it solves the linear equation, thus we may replace P_{W_0} by h to get

$$\int_Q \langle F''(W_0)\nabla_{\mathbf{a}}\widetilde{u}, \nabla_{\mathbf{a}}\varphi\rangle\,\mathrm{d}x = \int_Q \langle F''(W_0)\nabla_{\mathbf{a}}(u-h), \nabla_{\mathbf{a}}\varphi\rangle\,\mathrm{d}x.$$

Hence, setting $\psi := u - h$, we have

$$\left|\int_{Q_R} \langle F''(W_0)\nabla_{\mathbf{a}}\psi, \nabla_{\mathbf{a}}\varphi\rangle\,\mathrm{d}x\right| \leqslant \int_{Q_R} \omega_K(|\nabla_{\mathbf{a}}\widetilde{u}|^2)\cdot|\nabla_{\mathbf{a}}\widetilde{u}|\cdot|\nabla_{\mathbf{a}}\varphi|\,\mathrm{d}x. \tag{5.21}$$

5.3.2 Choosing the test function

We now wish to use the inequality (5.21) to obtain bounds on ψ in terms of its derivatives, as they will quantify the deviation of u from the smooth function h. The strategy we employ here relies on constructing a particular test function φ to be used in the aforementioned inequality and is based on the approach of Gmeineder and Kristensen in [82].

In the following it will be more convenient to work on the unit cube $Q = Q_1$, so let us change variables in (5.21). We may assume for simplicity that the centre of Q_R is at 0, as this is just a matter of a translation. We use the anisotropic scaling of our pattern of homogeneity and define, for a function $w\colon Q_R \to \mathbb{R}^n$, the function $w_R\colon Q \to \mathbb{R}^n$ by

$$w_R(x) := R^{-1} w(R \odot x).$$

Then (5.21) may be rewritten as

$$\left| \int_Q \langle F''(W_0) \nabla_{\mathbf{a}} \psi_R, \nabla_{\mathbf{a}} \varphi_R \rangle \, \mathrm{d}x \right| \leqslant \int_Q \omega_K(|\nabla_{\mathbf{a}} \widetilde{u}_R|^2) \cdot |\nabla_{\mathbf{a}} \widetilde{u}_R| \cdot |\nabla_{\mathbf{a}} \varphi_R| \, \mathrm{d}x, \tag{5.22}$$

for any $\varphi_R \in \mathrm{W}_0^{\mathbf{a},2}(Q)$.

Now our goal is to construct a family of φ's that will allow us to get bounds on all the non-maximal derivatives of ψ. Thus, let us fix a multiindex β with $\langle \beta, \mathbf{a}^{-1} \rangle < 1$ and an exponent $p \in (2, \infty)$ to be determined shortly, and consider the equation

$$\begin{cases} \nabla_{\mathbf{a}}^* (F''(W_0))^* \nabla_{\mathbf{a}} \varphi_R^\beta = (\partial^\beta)^* \partial^\beta \psi_R \text{ in } Q, \\ \varphi_R^\beta \in \mathrm{W}_0^{\mathbf{a},p}(Q). \end{cases} \tag{5.23}$$

Recall from Proposition 5.1.2 that φ_R^β is well-defined and satisfies

$$\|\varphi_R^\beta\|_{\mathrm{W}_0^{\mathbf{a},p}(Q)} \leqslant C \|(\partial^\beta)^* \partial^\beta \psi_R\|_{\mathrm{W}^{-\mathbf{a},p}(Q)}$$

whenever the latter is finite. We claim that there exists an exponent $p > 2$ and a constant C such that for all β's with $\langle \beta, \mathbf{a}^{-1} \rangle < 1$ one has

$$\|(\partial^\beta)^* \partial^\beta \psi_R\|_{\mathrm{W}^{-\mathbf{a},p}(Q)} \leqslant C \|\partial^\beta \psi_R\|_{\mathrm{L}^2(Q)}.$$

To prove this note that

$$\|(\partial^\beta)^* \partial^\beta \psi_R\|_{\mathrm{W}^{-\mathbf{a},p}(Q)} = \sup_{\|\zeta\|_{\mathrm{W}_0^{\mathbf{a},q}(Q)} \leqslant 1} \int_Q \partial^\beta \psi_R \partial^\beta \zeta \, \mathrm{d}x, \tag{5.24}$$

where q is such that $p^{-1}+q^{-1}=1$. Consider the maximal value of $\langle\beta,\mathbf{a}^{-1}\rangle$ over those multiindices β that satisfy $\langle\beta,\mathbf{a}^{-1}\rangle<1$. Since the set of admissible β's is finite this maximum is strictly smaller than 1. Thus, there exists a $q<2$ such that

$$\left(\frac{1}{q}-\frac{1}{2}\right)|\mathbf{a}^{-1}|+\max_{\langle\beta,\mathbf{a}^{-1}\rangle<1}\langle\beta,\mathbf{a}^{-1}\rangle<1.$$

Now the improved integrability result from Theorem 2.2.9 tells us that, with this choice of q, we have

$$\|\partial^\beta\zeta\|_{\mathrm{L}^2(Q)}\leqslant C\|\zeta\|_{\mathrm{W}_0^{\mathbf{a},q}(Q)},$$

for every admissible β, where the constant C may depend on β. However, this may be easily remedied by simply taking the maximum of all such constants across the set of β's considered. Applying the Cauchy-Schwarz inequality to the right-hand side of (5.24) we may write

$$\begin{aligned}\sup_{\|\zeta\|_{\mathrm{W}_0^{\mathbf{a},q}(Q)}\leqslant 1}\int_Q\partial^\beta\psi_R\partial^\beta\zeta\,\mathrm{d}x\leqslant&\sup_{\|\zeta\|_{\mathrm{W}_0^{\mathbf{a},q}(Q)}\leqslant 1}\|\partial^\beta\psi_R\|_{\mathrm{L}^2}\|\partial^\beta\zeta\|_{\mathrm{L}^2}\leqslant\\&\sup_{\|\zeta\|_{\mathrm{W}_0^{\mathbf{a},q}(Q)}\leqslant 1}C\|\partial^\beta\psi_R\|_{\mathrm{L}^2}\|\zeta\|_{\mathrm{W}_0^{\mathbf{a},q}}=C\|\partial^\beta\psi_R\|_{\mathrm{L}^2}.\end{aligned}$$

Thus indeed, with this choice of q, we obtain a $p>2$ such that

$$\|(\partial^\beta)^*\partial^\beta\psi_R\|_{\mathrm{W}^{-\mathbf{a},p}(Q)}\leqslant C\|\partial^\beta\psi_R\|_{\mathrm{L}^2(Q)},$$

and therefore also

$$\|\varphi_R^\beta\|_{\mathrm{W}_0^{\mathbf{a},p}(Q)}\leqslant C\|\partial^\beta\psi_R\|_{\mathrm{L}^2(Q)},\tag{5.25}$$

for all admissible β's.

Using φ_R^β as a test function in (5.22) we get

$$\left|\int_Q\langle\psi_R,(\nabla_{\mathbf{a}})^*(F''(W_0))^*\varphi_R^\beta\rangle\,\mathrm{d}x\right|\leqslant\int_Q\omega_K(|\nabla_{\mathbf{a}}\widetilde{u}_R|^2)\cdot|\nabla_{\mathbf{a}}\widetilde{u}_R|\cdot|\nabla_{\mathbf{a}}\varphi_R^\beta|\,\mathrm{d}x,$$

and using the fact that φ_R^β solves (5.23) this may be re-written as

$$\int_Q|\partial^\beta\psi_R|^2\,\mathrm{d}x\leqslant\int_Q\omega_K(|\nabla_{\mathbf{a}}\widetilde{u}_R|^2)\cdot|\nabla_{\mathbf{a}}\widetilde{u}_R|\cdot|\nabla_{\mathbf{a}}\varphi_R^\beta|\,\mathrm{d}x.$$

Using Hölder's inequality with exponents $(r,2,p)$ with $p>2$ obtained above and r such that $\frac{1}{r}+\frac{1}{2}+\frac{1}{p}=1$ we may write

$$\int_Q|\partial^\beta\psi_R|^2\,\mathrm{d}x\leqslant\left(\int_Q\omega_K(|\nabla_{\mathbf{a}}\widetilde{u}_R|^2)^r\right)^{1/r}\left(\int_Q|\nabla_{\mathbf{a}}\widetilde{u}_R|^2\,\mathrm{d}x\right)^{1/2}\left(\int_Q|\nabla_{\mathbf{a}}\varphi_R^\beta|^p\,\mathrm{d}x\right)^{1/p}.$$

Using (5.25), dividing both sides by $\|\partial^\beta \psi_R\|_{\mathrm{L}^2(Q)}$, and squaring yields

$$\int_Q |\partial^\beta \psi_R|^2 \,\mathrm{d}x \leqslant C \left(\int_Q \omega_K(|\nabla_{\mathbf{a}} \widetilde{u}_R|^2)^r \right)^{2/r} \int_Q |\nabla_{\mathbf{a}} \widetilde{u}_R|^2 \,\mathrm{d}x.$$

Using the L^∞ bound on ω_K to get rid of the r-th power inside the first integral gives

$$\int_Q |\partial^\beta \psi_R|^2 \,\mathrm{d}x \leqslant C \left(\int_Q \omega_K(|\nabla_{\mathbf{a}} \widetilde{u}_R|^2) \right)^{2/r} \int_Q |\nabla_{\mathbf{a}} \widetilde{u}_R|^2 \,\mathrm{d}x,$$

and now we may use concavity of ω_K to apply Jensen's inequality and get

$$⨍_Q |\partial^\beta \psi_R|^2 \,\mathrm{d}x \leqslant C \left(\omega_K \left(⨍_Q |\nabla_{\mathbf{a}} \widetilde{u}_R|^2 \right) \right)^{2/r} ⨍_Q |\nabla_{\mathbf{a}} \widetilde{u}_R|^2 \,\mathrm{d}x.$$

Furthermore, we may write $\widetilde{\omega}_K(t) := (\omega_K(t))^{2/r}$ and note that $\widetilde{\omega}_K(t)$ is still a bounded continuous function that satisfies $\widetilde{\omega}_K(0) = 0$, whereby our inequality simplifies to

$$⨍_Q |\partial^\beta \psi_R|^2 \,\mathrm{d}x \leqslant C \widetilde{\omega}_K \left(⨍_Q |\nabla_{\mathbf{a}} \widetilde{u}_R|^2 \right) ⨍_Q |\nabla_{\mathbf{a}} \widetilde{u}_R|^2 \,\mathrm{d}x.$$

Finally, re-scaling this inequality back to Q_R, we obtain

$$R^{-(2-2\langle \beta, \mathbf{a}^{-1} \rangle)} ⨍_{Q_R} |\partial^\beta (u-h)|^2 \,\mathrm{d}x \leqslant C \widetilde{\omega}_M \left(⨍_{Q_R} |\nabla_{\mathbf{a}} \widetilde{u}|^2 \right) ⨍_{Q_R} |\nabla_{\mathbf{a}} \widetilde{u}|^2 \,\mathrm{d}x. \tag{5.26}$$

5.3.3 Excess decay estimates

Here we let

$$\mathrm{E}(x_0, r) := ⨍_{Q_r(x_0)} |\nabla_{\mathbf{a}} u - (\nabla_{\mathbf{a}} u)_{Q_r(x_0)}|^2 \,\mathrm{d}x$$

be the excess of u on $Q_r(x_0)$. For brevity, we often omit the centre x_0, as it is kept fixed throughout the argument. Our goal is to obtain estimates on $\mathrm{E}(x_0, r)$ as a function of r.

Let us take $\tau \in (0,1)$ to be determined later and an arbitrary **a**-polynomial b, i.e., a polymomial with $\nabla_{\mathbf{a}} b \equiv \text{const}$. Caccioppoli's inequality from Lemma 5.2.4 applied on $Q_{\frac{\tau R}{2}}$ reads

$$⨍_{Q_{\frac{\tau R}{2}}} |\nabla_{\mathbf{a}} (u-b)|^2 \,\mathrm{d}x \leqslant C \sum_{\langle \beta, \mathbf{a}^{-1} \rangle < 1} (\tau R)^{-2(1-\langle \beta, \mathbf{a}^{-1} \rangle)} ⨍_{Q_{\tau R}} |\partial^\beta (u-b)|^2 \,\mathrm{d}x.$$

For the left-hand side we observe that $\nabla_{\mathbf{a}} b$ is a constant and, amongst all the constants, $(\nabla_{\mathbf{a}} u)_{Q_{\frac{\tau R}{2}}}$ minimises the value of $z \mapsto ⨍_{Q_{\frac{\tau R}{2}}} |\nabla_{\mathbf{a}} u - z|^2 \,\mathrm{d}x$. Thus

$$\mathrm{E}(x_0, \frac{\tau R}{2}) \leqslant ⨍_{Q_{\frac{\tau R}{2}}} |\nabla_{\mathbf{a}} (u-b)|^2 \,\mathrm{d}x,$$

for any **a**-polynomial b. For the right-hand side we may use the triangle inequality on each of the integral terms to write

$$\fint_{Q_{\tau R}} |\partial^\beta (u-b)|^2 \,\mathrm{d}x \leqslant 2\fint_{Q_{\tau R}} |\partial^\beta (u-h)|^2 \,\mathrm{d}x + 2\fint_{Q_{\tau R}} |\partial^\beta (h-b)|^2 \,\mathrm{d}x.$$

Thus, we need to estimate the following two sums:

$$\mathrm{I}_1 := \sum_{\langle\beta,\mathbf{a}^{-1}\rangle<1} (\tau R)^{-2(1-\langle\beta,\mathbf{a}^{-1}\rangle)} \fint_{Q_{\tau R}} |\partial^\beta (u-h)|^2 \,\mathrm{d}x,$$

$$\mathrm{I}_2 := \sum_{\langle\beta,\mathbf{a}^{-1}\rangle<1} (\tau R)^{-2(1-\langle\beta,\mathbf{a}^{-1}\rangle)} \fint_{Q_{\tau R}} |\partial^\beta (h-b)|^2 \,\mathrm{d}x.$$

Let us begin with I_1. Extending the integral to the whole of Q_R and noting that $\frac{|Q_R|}{|Q_{\tau R}|} = \tau^{-|\mathbf{a}^{-1}|}$ lets us write

$$\mathrm{I}_1 \leqslant \sum_{\langle\beta,\mathbf{a}^{-1}\rangle<1} \tau^{-2(1-\langle\beta,\mathbf{a}^{-1}\rangle)-|\mathbf{a}^{-1}|} R^{-2(1-\langle\beta,\mathbf{a}^{-1}\rangle)} \fint_{Q_R} |\partial^\beta (u-h)|^2 \,\mathrm{d}x.$$

Using the estimate (5.26) yields

$$\mathrm{I}_1 \leqslant C \sum_{\langle\beta,\mathbf{a}^{-1}\rangle<1} \tau^{-2(1-\langle\beta,\mathbf{a}^{-1}\rangle)-|\mathbf{a}^{-1}|} \widetilde{\omega}_M \left(\fint_{Q_R} |\nabla_\mathbf{a} \widetilde{u}|^2\right) \fint_{Q_R} |\nabla_\mathbf{a} \widetilde{u}|^2 \,\mathrm{d}x.$$

Remembering that $\nabla_\mathbf{a}\widetilde{u} = \nabla_\mathbf{a} u - (\nabla_\mathbf{a} u)_{Q_R}$ this may be written as

$$\mathrm{I}_1 \leqslant C \sum_{\langle\beta,\mathbf{a}^{-1}\rangle<1} \tau^{-2(1-\langle\beta,\mathbf{a}^{-1}\rangle)-|\mathbf{a}^{-1}|} \widetilde{\omega}_M \left(\mathrm{E}(x_0,R)\right) \mathrm{E}(x_0,R).$$

Finally, putting an upper bound on τ, say $\tau \leqslant 1/2$, we may write

$$\mathrm{I}_1 \leqslant C\tau^{-2-|\mathbf{a}^{-1}|} \widetilde{\omega}_M \left(\mathrm{E}(x_0,R)\right) \mathrm{E}(x_0,R).$$

To deal with I_2, let us start by rescaling h to h^R, so that we work on Q_τ and Q instead of $Q_{\tau R}$ and Q_R. Let us pick b to be the polynomial given by Corollary 2.5.3 with $p = 2$ and $r = \tau$ and apply the inequality given therein. This lets us estimate

$$\mathrm{I}_2 \leqslant C\fint_{Q_\tau} |\nabla_\mathbf{a} h^R - \left(\nabla_\mathbf{a} h^R\right)_{Q_\tau}|^2 \,\mathrm{d}x,$$

with an absolute constant C. Since we assume $\tau \leqslant 1/2$, we may estimate this last quantity in terms of a Campanato norm of $\nabla_\mathbf{a} h$ on $Q_{1/2}$. To this end let us fix an arbitrary $\theta \in (1, 1 + \frac{2}{|\mathbf{a}^{-1}|})$. Then Theorem 5.1.13 tells us that

$$\|\nabla_\mathbf{a} h^R\|^2_{\mathfrak{L}^{2,\theta}_\mathbf{a}(Q_{1/2})} \leqslant C\|h^R\|^2_{\mathrm{W}^{\mathbf{a},2}(Q)}.$$

In fact, since we only care about the Campanato norm of the highest derivatives of h^R, we may estimate with the homogeneous norm $\|\nabla_{\mathbf{a}} h^R\|_{\mathrm{L}^2(Q)}$ rather than the full $\mathrm{W}^{\mathbf{a},2}(Q)$ norm. To see this, observe that for any polynomial P with $\nabla_{\mathbf{a}} P \equiv 0$ the function $h^R - P$ still satisfies $\nabla_{\mathbf{a}}^* F''(W_0)\nabla_{\mathbf{a}}(h-P) = 0$ in Q, and obviously

$$\|\nabla_{\mathbf{a}} h^R\|^2_{\mathfrak{L}^{2,\theta}_{\mathbf{a}}(Q_{1/2})} = \|\nabla_{\mathbf{a}}(h^R - P)\|^2_{\mathfrak{L}^{2,\theta}_{\mathbf{a}}(Q_{1/2})} \leqslant C\|h^R - P\|^2_{\mathrm{W}^{\mathbf{a},2}(Q)},$$

where the last inequality is just Theorem 5.1.13 applied to $h^R - P$. Letting P be the polynomial given by Proposition 2.5.2 we have

$$\|h^R - P\|^2_{\mathrm{W}^{\mathbf{a},2}(Q)} \leqslant C\|\nabla_{\mathbf{a}} h^R\|^2_{\mathrm{L}^2(Q)},$$

and we may finally write

$$\mathrm{I}_2 \leqslant C\tau^{(\theta-1)|\mathbf{a}^{-1}|}\|\nabla_{\mathbf{a}} h^R\|^2_{\mathfrak{L}^{2,\theta}_{\mathbf{a}}(Q_{1/2})} \leqslant C\tau^{(\theta-1)|\mathbf{a}^{-1}|}\|\nabla_{\mathbf{a}} h^R\|^2_{\mathrm{L}^2(Q)}.$$

Rescaling back yields

$$\mathrm{I}_2 \leqslant C\tau^{(\theta-1)|\mathbf{a}^{-1}|}R^{-|\mathbf{a}^{-1}|}\|\nabla_{\mathbf{a}} h\|^2_{\mathrm{L}^2(Q_R)}.$$

Remembering that h solves (5.19) we note that Lemma 5.1.4 implies that

$$\|\nabla_{\mathbf{a}} h\|^2_{\mathrm{L}^2(Q_R)} \leqslant C\|\nabla_{\mathbf{a}}\widetilde{u}\|^2_{\mathrm{L}^2(Q_R)},$$

hence,

$$\mathrm{I}_2 \leqslant C\tau^{(\theta-1)|\mathbf{a}^{-1}|}\fint_{Q_R} |\nabla_{\mathbf{a}}\widetilde{u}|^2 \,\mathrm{d}x = C\tau^{(\theta-1)|\mathbf{a}^{-1}|}\mathrm{E}(x_0, R).$$

Finally, putting the bounds on I_1 and I_2 together we have shown:

Lemma 5.3.1. *For any fixed $\theta < 1 + \frac{2}{|\mathbf{a}^{-1}|}$, any $M > 0$, any $x_0 \in \Omega$ and any $R > 0$ such that $Q_R \subset \Omega$ and $\left|\fint_{Q_R} \nabla_{\mathbf{a}} u \,\mathrm{d}x\right| < M$ we have*

$$\mathrm{E}(x_0, \tau R) \leqslant C\left(\tau^{-2-|\mathbf{a}^{-1}|}\widetilde{\omega}_M\left(\mathrm{E}(x_0, R)\right) + \tau^{(\theta-1)|\mathbf{a}^{-1}|}\right)\mathrm{E}(x_0, R), \tag{5.27}$$

for any $\tau < 1/4$. Here $C = C(\theta, M, \nu)$ depends only on these three parameters, in particular it does not depend on x_0 or R, whereas $\widetilde{\omega}_M$ depends only on M and F'', and $\|\widetilde{\omega}_M\|_{\mathrm{L}^\infty} \leqslant 2\|F''\|_{\mathrm{L}^\infty}$.

Note that, for cosmetic reasons, we have changed from $\mathrm{E}(x_0, \frac{\tau R}{2})$ to $\mathrm{E}(x_0, \tau R)$. Clearly this is only a matter of increasing the constant C by an absolute factor and going from the restriction $\tau < 1/2$ to $\tau < 1/4$. We easily deduce the following:

Corollary 5.3.2. *For any fixed $\theta < 1 + \frac{2}{|\mathbf{a}^{-1}|}$ and any $M > 0$ there exists a number $\tau < 1/4$ and a positive number $\varepsilon_0 = \varepsilon_0(\theta, M, \nu, F'')$ such that, for any $x_0 \in \Omega$ and any $R > 0$ such that $Q_R(x_0) \subset \Omega$ with $\left|\fint_{Q_R} \nabla_{\mathbf{a}} u \, \mathrm{d}x\right| < M$ and $\mathrm{E}(x_0, R) < \varepsilon_0$ we have*

$$\mathrm{E}(x_0, \tau R) \leqslant \tau^{(\theta-1)|\mathbf{a}^{-1}|}\mathrm{E}(x_0, R). \tag{5.28}$$

Proof. Fix some $\theta' \in (\theta, 1 + 2/|\mathbf{a}^{-1}|)$ to use in Lemma 5.3.1. Since the constant C in (5.27) does not depend on τ we may find $\tau \in (0, 1/4)$ small enough so that

$$C\tau^{(\theta'-1)|\mathbf{a}^{-1}|} \leqslant \frac{1}{2}\tau^{(\theta-1)|\mathbf{a}^{-1}|}.$$

Note that $\tau = \tau(\theta, M, \nu)$, as it depends on the constant C. With τ fixed, we may now find a positive number $\varepsilon_0 > 0$ such that

$$C\tau^{-2-|\mathbf{a}^{-1}|}\widetilde{\omega}_M(\varepsilon) < \frac{1}{2}\tau^{(\theta-1)|\mathbf{a}^{-1}|}$$

for any $\varepsilon \in [0, \varepsilon_0)$. This is because the function $\widetilde{\omega}_M(\cdot)$ is continuous and $\widetilde{\omega}_M(0) = 0$. Therefore, our ε_0 depends on $\widetilde{\omega}_M$ and τ, i.e., it is a function of θ, M, ν, and F''. Putting the two estimates together yields the desired inequality. □

5.3.4 Iteration and conclusion

Now our goal is to iterate the inequality (5.28) to obtain, for an arbitrary natural number k, estimates on $\mathrm{E}(x_0, \tau^k R)$ in terms of $\mathrm{E}(x_0, R)$. For that we need to ensure that the assumptions of Corollary 5.3.2 are satisfied on each consecutive box $Q_{\tau^k R}$.

Let us now fix an arbitrary number $M > 0$ and let $\varepsilon_0(\theta, M+1, \nu, F'')$ be given by Corollary 5.3.2. Note that we take ε_0 to correspond to $M + 1$ rather than M in order to allow for some room in the iterative argument. Take an $\varepsilon \in (0, \varepsilon_0)$ to be determined later and suppose that the box $Q_R = Q_R(x_0)$ satisfies the assumptions of Corollary 5.3.2 with M and ε, i.e., assume that

$$\left|\fint_{Q_R} \nabla_{\mathbf{a}} u \, \mathrm{d}x\right| < M \quad \text{and} \quad \mathrm{E}(x_0, R) < \varepsilon.$$

Thus, we may apply Corollary 5.3.2 on Q_R, and we wish to check whether we can also do so on $Q_{\tau R}$ to begin our iteration. Clearly,

$$\mathrm{E}(x_0, \tau R) \leqslant \tau^{(\theta-1)|\mathbf{a}^{-1}|}\mathrm{E}(x_0, R) \leqslant \mathrm{E}(x_0, R) < \varepsilon,$$

so we just need a bound on $|(\nabla_{\mathbf{a}} u)_{Q_{\tau R}}|$. To this end, let us note that, using the triangle inequality, we may write

$$\left|\fint_{Q_{\tau R}} \nabla_{\mathbf{a}} u \, \mathrm{d}x\right| \leqslant |(\nabla_{\mathbf{a}} u)_{Q_R}| + \left|\fint_{Q_{\tau R}} \nabla_{\mathbf{a}} u - (\nabla_{\mathbf{a}} u)_{Q_R} \, \mathrm{d}x\right|.$$

We assumed that $|(\nabla_{\mathbf{a}} u)_{Q_R}| < M$, and for the second term we may enlarge the domain of integration to the whole of Q_R, taking into account the scaling factor $\tau^{-|\mathbf{a}^{-1}|}$ due to the fact that we work with average integrals. This gives

$$\mathrm{E}(x_0, \tau R) \leqslant M + \tau^{-|\mathbf{a}^{-1}|} \fint_{Q_R} |\nabla_{\mathbf{a}} u - (\nabla_{\mathbf{a}} u)_{Q_R}| \,\mathrm{d}x.$$

Finally, using the Hölder inequality between the L^1 and L^2 norms on Q_R with normalised measure yields

$$\mathrm{E}(x_0, \tau R) \leqslant M + \tau^{-|\mathbf{a}^{-1}|} \sqrt{\mathrm{E}(x_0, R)} \leqslant M + \tau^{-|\mathbf{a}^{-1}|} \sqrt{\varepsilon},$$

where we used the assumption that $\mathrm{E}(x_0, R) < \varepsilon$. Thus, requiring that $\varepsilon < \tau^{2|\mathbf{a}^{-1}|}$ yields $\left| \fint_{Q_{\tau R}} \nabla_{\mathbf{a}} u \,\mathrm{d}x \right| < M + 1$, and Corollary 5.3.2 may be applied on $Q_{\tau R}(x_0)$.

To apply the result on $Q_{\tau^2 R}$ we note that $\mathrm{E}(x_0, \tau^2 R) \leqslant \tau^{2(\theta-1)|\mathbf{a}^{-1}|} \mathrm{E}(x_0, R) < \varepsilon$, and for the average, similarly as before, we write

$$\begin{aligned} \left| \fint_{Q_{\tau^2 R}} \nabla_{\mathbf{a}} u \,\mathrm{d}x \right| \leqslant & |(\nabla_{\mathbf{a}} u)_{Q_R}| + \fint_{Q_{\tau R}} |\nabla_{\mathbf{a}} u - (\nabla_{\mathbf{a}} u)_{Q_R}| \,\mathrm{d}x + \\ & \fint_{Q_{\tau^2 R}} |\nabla_{\mathbf{a}} u - (\nabla_{\mathbf{a}} u)_{Q_{\tau R}}| \,\mathrm{d}x \leqslant \\ & M + \tau^{-|\mathbf{a}^{-1}|} \left(\sqrt{\mathrm{E}(x_0, R)} + \sqrt{\mathrm{E}(x_0, \tau R)} \right) \leqslant \\ & M + \tau^{-|\mathbf{a}^{-1}|} \left(\sqrt{\varepsilon} + \sqrt{\tau^{(\theta-1)|\mathbf{a}^{-1}|} \varepsilon} \right). \end{aligned}$$

In general, if we know we may apply our result $k-1$ times then we have

$$\begin{aligned} \left| \fint_{Q_{\tau^k R}} \nabla_{\mathbf{a}} u \,\mathrm{d}x \right| \leqslant & M + \tau^{-|\mathbf{a}^{-1}|} \sum_{j=0}^{k-1} \sqrt{\mathrm{E}(x_0, \tau^j R)} < \\ & M + \tau^{-|\mathbf{a}^{-1}|} \sqrt{\varepsilon} \sum_{j=0}^{k-1} \tau^{\frac{j}{2}(\theta-1)|\mathbf{a}^{-1}|} < \\ & M + \tau^{-|\mathbf{a}^{-1}|} \sqrt{\varepsilon} \sum_{j=0}^{\infty} \tau^{\frac{j}{2}(\theta-1)|\mathbf{a}^{-1}|}. \end{aligned}$$

Thus, if we require that $\sqrt{\varepsilon} < \tau^{|\mathbf{a}^{-1}|} \left(\sum_{j=0}^{\infty} \tau^{\frac{j}{2}(\theta-1)|\mathbf{a}^{-1}|} \right)^{-1}$ then we may apply our result indefinitely, as for every k we have $\left| \fint_{Q_{\tau^k R}} \nabla_{\mathbf{a}} u \,\mathrm{d}x \right| < M+1$ and $\mathrm{E}(x_0, \tau^k R) < \varepsilon$, hence

$$\mathrm{E}(x_0, \tau^k R) \leqslant \tau^{k(\theta-1)|\mathbf{a}^{-1}|} \mathrm{E}(x_0, R),$$

for every $k \in \mathbb{N}$. Finally, to show that

$$\mathrm{E}(x_0, \rho R) \leqslant C \rho^{(\theta-1)|\mathbf{a}^{-1}|} \mathrm{E}(x_0, R)$$

for an arbitrary $\rho \in (0,1)$ and C independent of ρ, it is enough to observe that for a given $\rho \in (0,1)$ there exists some $k \geqslant 0$ such that

$$\tau^{k+1} \leqslant \rho \leqslant \tau^k,$$

and thus one can interpolate by taking the k as above and noting that

$$\begin{aligned}\mathrm{E}(x_0, \rho R) \leqslant & \frac{|Q_{\tau^k R}|}{|Q_{\rho R}|}\mathrm{E}(x_0, \tau^k R) \leqslant \\ & \tau^{-|\mathbf{a}^{-1}|}\tau^{(k+1)(\theta-1)|\mathbf{a}^{-1}|}\tau^{\theta|\mathbf{a}^{-1}|}\mathrm{E}(x_0, R) \leqslant \\ & \tau^{(\theta-1)|\mathbf{a}^{-1}|}\rho^{(\theta-1)|\mathbf{a}^{-1}|}\mathrm{E}(x_0, R),\end{aligned}$$

which gives $C = \tau^{(\theta-1)|\mathbf{a}^{-1}|}$, thus $C = C(M, \theta, F'')$. All in all, we have shown the following:

Lemma 5.3.3. *For any fixed $\theta < 1 + \frac{2}{|\mathbf{a}^{-1}|}$ and $M > 0$ there exists some constant $\varepsilon = \varepsilon(M, \theta, \nu, F'') > 0$ such that for any $x_0 \in \Omega$ and $R > 0$ satisfying*

$$\begin{cases}\operatorname{dist}(x_0, \partial\Omega) > R, \\ \left|\fint_{Q_R(x_0)} \nabla_{\mathbf{a}} u \,\mathrm{d}x\right| < M, \\ \mathrm{E}(x_0, R) < \varepsilon(M, \theta, \nu, F''),\end{cases}$$

we have

$$\mathrm{E}(x_0, \rho R) \leqslant C\rho^{(\theta-1)|\mathbf{a}^{-1}|}\mathrm{E}(x_0, R),$$

with $C = C(M, \theta, \nu, F'')$.

For a fixed θ and any $M, R > 0$ let us denote by $\Omega_{M,R}$ the set of those $x \in \Omega$ for which $Q_R(x)$ satisfies the conditions of Lemma 5.3.3.

Lemma 5.3.4. *Each $\Omega_{M,R}$ is open.*

Proof. This is an immediate consequence of the continuity of the three functions

$$x \mapsto \operatorname{dist}(x, \partial\Omega), \quad x \mapsto \fint_{Q_R(x)} \nabla_{\mathbf{a}} u \,\mathrm{d}y, \quad x \mapsto \mathrm{E}(x, R).$$

□

Lemma 5.3.5. *The family $\{\Omega_{M,R}\}_{M,R}$ covers Ω up to a set of Lebesgue measure zero, i.e., for almost every $x \in \Omega$ there exist numbers $M, R > 0$ such that $x \in \Omega_{M,R}$.*

Proof. By the Lebesgue differentiation theorem the set of points that are L^2-Lebesgue (thus also L^1-Lebesgue) points of $\nabla_{\mathbf{a}} u$ has full measure in Ω. Thus, for a given point x in this set we may take $M := 2|\nabla_{\mathbf{a}} u(x)|$, where $\nabla_{\mathbf{a}} u(x)$ is understood as the value of the Lebesgue representative of $\nabla_{\mathbf{a}} u$ at x. Since $x \in \Omega$ is an L^2-Lebesgue point of $\nabla_{\mathbf{a}} u$ the three conditions of Lemma 5.3.3 must be satisfied for $R > 0$ small enough, which ends the proof. □

Proposition 5.3.6. *For almost every $x \in \Omega$ there exists an open box $x \in Q_x \Subset \Omega$ on which $\nabla_{\mathbf{a}} u$ is Hölder continuous with respect to the metric $\delta_{\mathbf{a}}$ with exponent $\frac{\theta-1}{2}|\mathbf{a}^{-1}|$.*

Proof. By Lemma 5.3.5 we know that, for almost every $x \in \Omega$, there exists some $\Omega_{M,R}$ with $x \in \Omega_{M,R}$. The set $\Omega_{M,R}$ is open, so we may choose an open box $x \in Q_x \Subset \Omega_{M,R}$. The bound from Lemma 5.3.3 implies that $\nabla_{\mathbf{a}} u \in \mathfrak{L}^{2,\theta}_{\mathbf{a}}(Q_x)$, and since Q_x is clearly of type (A) with respect to $\mathbf{a}$, we deduce from Lemma 2.6.3 that $\nabla_{\mathbf{a}} u$ is Hölder continuous with exponent $\frac{\theta-1}{2}|\mathbf{a}^{-1}|$ on Q_x with respect to the metric $\delta_{\mathbf{a}}$. □

Since the above proposition holds with any $\theta \in (1, 1 + 2/|\mathbf{a}^{-1}|)$, we have finally proven the following:

Theorem 5.3.7. *Let $\Omega \subset \mathbb{R}^N$ be a bounded domain satisfying the weak $\mathbf{a}$-horn condition and suppose that $F\colon \mathbb{R}^{n\times m} \to \mathbb{R}$ is a strictly $W^{\mathbf{a},2}$-quasiconvex integrand of class C^2. Assume furthermore that $|F''(W)| \leqslant C_1$ for some $C_1 > 0$ and all $W \in \mathbb{R}^{n\times m}$, so that also $|F(W)| \leqslant C_2(1 + |W|^2)$ for some $C_2 > 0$ and all $W \in \mathbb{R}^{n\times m}$. Let u be a minimiser of the induced functional $I_F(u) := \int_\Omega F(\nabla_{\mathbf{a}} u)\,dx$ over the class $W^{\mathbf{a},2}_{u_b}(\Omega)$ for some fixed boundary datum $u_b \in W^{\mathbf{a},2}(\mathbb{R}^N)$. Then, for any $\alpha < 1$, there exists an open set $\Omega_g \subset \Omega$ with $|\Omega \setminus \Omega_g| = 0$ and such that, on Ω_g, $\nabla_{\mathbf{a}} u$ is locally Hölder continuous with exponent α with respect to the metric $\delta_{\mathbf{a}}$.*

Remark 9. Let us note that the previous result is phrased in terms of the anisotropic metric $\delta_{\mathbf{a}}$, which is the natural setting since the result is obtained through Campanato bounds on the family of anisotropic boxes. Nevertheless, Theorem 5.3.7 may be rephrased in terms of the standard Euclidean metric thanks to Lemma 2.6.4. Then, the conclusion is that $\nabla_{\mathbf{a}} u$ is locally Hölder continuous with exponent $\alpha\left(\frac{a_i}{\max_j a_j}\right)$ in the x_i variable.

Bibliography

[1] E. Acerbi, N. Fusco, Semicontinuity problems in the calculus of variations, *Archive for Rational Mechanics and Analysis*, 86.2 (1984), 125-145.

[2] E. Acerbi, N. Fusco, A regularity theorem for minimizers of quasiconvex integrals, *Archive for Rational Mechanics and Analysis*, 99.3 (1987), 261-281.

[3] E. Acerbi, N. Fusco, Local regularity for minimizers of non convex integrals, *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 16.4 (1989), 603-636.

[4] J.-J. Alibert, G. Bouchitté, Non-Uniform Integrability and Generalized Young Measures, *Journal of Convex Analysis*, 4 (1997), 129-148.

[5] J.-J. Alibert, B. Dacorogna, An example of a quasiconvex function that is not polyconvex in two dimensions, *Archive for rational mechanics and analysis*, 117.2 (1992), 155-166.

[6] F. J. Almgren Jr., Existence and regularity almost everywhere of solutions to elliptic variational problems among surfaces of varying topological type and singularity structure, *Annals of Mathematics (2)*, 87, (1968), 321-391.

[7] F. J. Almgren Jr., Existence and regularity almost everywhere of solutions to elliptic variational problems with constraints, *Memoirs of the American Mathematical Society*, 154.4 (1976).

[8] E. J. Balder, A general approach to lower semicontinuity and lower closure in optimal control theory, *SIAM journal on control and optimization*, 22.4 (1984), 570-598.

[9] J. M. Ball, Convexity conditions and existence theorems in nonlinear elasticity, *Archive for rational mechanics and Analysis*, 63.4 (1976), 337-403.

[10] J. M. Ball, A version of the fundamental theorem for Young measures in *PDEs and continuum models of phase transitions, Springer, Berlin, Heidelberg*, (1989), 207-215.

[11] J. M. Ball, J. C. Currie, P. J. Olver, Null Lagrangians, weak continuity, and variational problems of arbitrary order *Journal of Functional Analysis*, 41.2 (1981), 135-174.

[12] J. M. Ball, B. Kirchheim, J. Kristensen, Regularity of quasiconvex envelopes *Calculus of Variations and Partial Differential Equations*, 11.4 (2000), 333-359.

[13] J. M. Ball, F. Murat, $W^{1,p}$-quasiconvexity and variational problems for multiple integrals, *Journal of Functional Analysis*, 58.3 (1984), 225-253.

[14] J. M. Ball, K. Zhang, Lower semicontinuity of multiple integrals and the biting lemma, *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 114.3-4 (1990), 367-379.

[15] G. C. Barozzi, Un problema al contorno non omogeneo in un dominio angoloso per equazioni fortemente quasi-ellittiche in due variabili (I), *Rendiconti del Seminario Matematico della Universita di Padova*, 44 (1970), 27-63.

[16] G. C. Barozzi, Un problema al contorno non omogeneo in un dominio angoloso per equazioni fortemente quasi-ellittiche in due variabili (II), *Rendiconti del Seminario Matematico della Universita di Padova*, 44 (1970), 319-337.

[17] J. J. Benedetto, W. Czaja, Integration and modern analysis *Springer Science & Business Media*, 2010.

[18] A. Bensoussan, J. Frehse, Regularity results for nonlinear elliptic systems and applications *Springer Science & Business Media*, 2013.

[19] H. Berliocchi, J.-M. Lasry, Intégrandes normales et mesures paramétrées en calcul des variations *Bulletin de la Société Mathématique de France*, 101 (1973), 129-184.

[20] O. V. Besov, Concerning the theory of embedding and continuing classes of differentiable functions, *Mathematical Notes*, 1.2 (1967), 156-161.

[21] O. V. Besov, Growth of a mixed derivative of a function of $C^{(l_1,l_2)}$, *Mathematical notes of the Academy of Sciences of the USSR*, 15.3 (1974), 201-206.

[22] O. V. Besov, V. P. Il'in, Natural extension of the class of regions in embedding theorems *Sbornik: Mathematics*, 4.4 (1968), 445-456.

[23] O. V. Besov, V. P. Il'in, S. M. Nikolskii, Integral representations of functions and imbedding theorems Vol. 1 *Winston & sons*, 1978.

[24] O. V. Besov, V. P. Il'in, S. M. Nikolskii, Integral representations of functions and imbedding theorems Vol. 2 *Winston & sons*, 1978.

[25] J. Boman, Supremum norm estimates for partial derivatives of functions of several real variables, *Illinois Journal of Mathematics*, 16.2 (1972), 203-216.

[26] A. Braides, I. Fonseca, G. Leoni, A-quasiconvexity: relaxation and homogenization, *ESAIM: Control, Optimisation and Calculus of Variations*, 5 (2000), 539-577.

[27] H. Brezis, Functional analysis, Sobolev spaces and partial differential equations, *Springer Science & Business Media*, 2010.

[28] V. I. Burenkov, Imbedding and extension theorems for classes of differentiable functions of several variables defined on the entire spaces, *Itogi Nauki i Tekhniki. Seriya Matematicheskii Analiz*, 3 (1966), 71-155.

[29] V. I. Burenkov, B. L. Fain, On the extension of functions from anisotropic spaces with preservation of class in Doklady Akademii Nauk 228.3 *Russian Academy of Sciences*, 1976.

[30] G. Buttazzo, Semicontinuity, relaxation and integral representation in the calculus of variations, *Longman*, 1989.

[31] R. Caccioppoli, Limitazioni integrali per le soluzioni di un equazione lineare ellitica a derivate parziali, *Giornale di Matematiche di Battaglini*, 4.80 (1951), 186-212.

[32] F. Cagnetti, k-quasi-convexity reduces to quasi-convexity *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 141.4 (2011), 673-708.

[33] A. P. Calderón, A. Torchinsky, Parabolic maximal functions associated with a distribution, *Advances in Mathematics*, 16.1 (1975), 1-64.

[34] S. Campanato, Hölder continuity of the solutions of some non-linear elliptic systems, *Advances in Mathematics*, 48.1 (1983), 16-43.

[35] M. Carozza, N. Fusco, G. Mingione, Partial regularity of minimizers of quasiconvex integrals with subquadratic growth, *Annali di Matematica Pura ed Applicata*, 175.1 (1998), 141-164.

[36] A. Cavallucci, Sulle proprieta differenziali delle soluzioni delle equazioni quasi-ellittiche, *Annali di Matematica Pura ed Applicata*, 67.1 (1965): 143-167.

[37] C. Y. Chen, J. Kristensen, On coercive variational integrals *Nonlinear Analysis: Theory, Methods & Applications*, 153 (2017), 213-229.

[38] B. Dacorogna, Quasiconvexity and relaxation of nonconvex problems in the calculus of variations, *Journal of Functional Analysis*, 46.1 (1982), 102-118.

[39] B. Dacorogna, Weak Continuity and Weak Lower Semicontinuity of Non-Linear Functionals, *Lecture Notes in Mathematics, Springer-Verlag*, 1982.

[40] B. Dacorogna, Direct methods in the calculus of variations, *Springer Science & Business Media*, 2007.

[41] B. Dacorogna, N. Fusco, Semi-continuité des fonctionnelles avec contraintes du type $\det \nabla u > 0$, *Bollettino della Unione Matemàtica Italiana. Serie VI. B*, (1985), 179-189.

[42] B. Dacorogna, P. Marcellini, A counterexample in the vectorial calculus of variations, *Material instabilities in continuum mechanics*, (1988), 77-83.

[43] B. Dacorogna, P. Marcellini, Implicit Partial Differential Equations, *Birkhäuser Basel*, 1999.

[44] G. Dal Maso, I. Fonseca, G. Leoni, M. Morini, Higher-order quasiconvexity reduces to quasiconvexity, *Archive for rational mechanics and analysis*, 171.1 (2004), 55-81.

[45] E. De Giorgi, Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari, *Memorie della Reale Accademia delle Scienze di Torino. Classe di Scienze Fisiche, Matematiche et Naturali*, 3 (1957), 25-43.

[46] E. De Giorgi, Frontiere orientate di misura minima, *Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960-61 Editrice Tecnico Scientifica, Pisa*, (1961).

[47] E. De Giorgi, Un esempio di estremali discontinue per un problema variazionale di tipo ellittico *Bollettino dell'Unione Matematica Italiana*, 4.1 (1968), 135-137.

[48] E. De Giorgi, Congetture riguardanti alcuni problemi di evoluzione *Duke Mathematical Journal*, 81.2 (1996), 255-268.

[49] F. dell'Isola, T. Lekszycki, M. Pawlikowski, R. Grygoruk, L. Greco, Designing a light fabric metamaterial being highly macroscopically tough under directional extension: first experimental evidence, *Zeitschrift für angewandte Mathematik und Physik*, 66.6 (2015), 3473-3498.

[50] G. V. Demidenko, Quasielliptic operators and Sobolev type equations, *Siberian Mathematical Journal*, 49.5 (2008), 842-851.

[51] G. V. Demidenko, Quasielliptic operators and Sobolev type equations. II, *Siberian Mathematical Journal*, 50.5 (2009): 838-845.

[52] G. V. Demidenko, S. V. Upsenskii, Partial differential equations and systems not solvable with respect to the highest-order derivative *CRC Press*, 2003.

[53] R. J. DiPerna, Compensated compactness and general systems of conservation laws *Transactions of the American Mathematical Society*, 292.2 (1985), 383-420.

[54] R. J. DiPerna, A. J. Majda, Oscillations and concentrations in weak solutions of the incompressible fluid equations, *Communications in Mathematical Physics*, 108.4 (1987), 667-689.

[55] G. Dolzmann, J. Kristensen, Higher integrability of minimizing Young measures, *Calculus of Variations and Partial Differential Equations*, 22.3 (2005), 283-301.

[56] T. Dupont, R. Scott, Polynomial approximation of functions in Sobolev spaces, *Mathematics of Computation*, 34.150 (1980), 441-463.

[57] I. Ekeland, Nonconvex minimization problems, *Bulletin of the American Mathematical Society*, 1.3 (1979), 443-474.

[58] I. Ekeland, R. Temam, Convex analysis and variational problems *Siam*, 1999.

[59] V. A. Eremeyev, F. dell'Isola, C. Boutin, D. Steigmann, Linear pantographic sheets: existence and uniqueness of weak solutions, *Journal of Elasticity*, 132.2 (2018), 175-196.

[60] L. Euler, Elementa calculi variationum, *Novi commentarii academiae scientiarum Petropolitanae*, (1766), 51-93.

[61] L. C. Evans, Quasiconvexity and partial regularity in the calculus of variations, *Archive for rational mechanics and analysis*, 95.3 (1986), 227-252.

[62] L. C. Evans, R. F. Gariepy, Some remarks concerning quasiconvexity and strong convergence, *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 106.1-2 (1987), 53-61.

[63] I. Fonseca, M. Kružík, Oscillations and concentrations generated by A-free mappings and weak lower semicontinuity of integral functionals, *ESAIM: Control, Optimisation and Calculus of Variations*, 16.2 (2010), 472-502.

[64] I. Fonseca, G. Leoni, Modern Methods in the Calculus of Variations: L^p Spaces *Springer Science & Business Media*, 2007.

[65] I. Fonseca, G. Leoni, S. Müller, A-quasiconvexity: weak-star convergence and the gap, *Annales de l'IHP, Analyse non linéaire*, 21.2 (2004), 209-236.

[66] I. Fonseca, J. Malý, Relaxation of multiple integrals below the growth exponent, *Annales de l'Institut Henri Poincare (C) Non Linear Analysis, Elsevier Masson*, 14.3 (1997), 309-338.

[67] I. Fonseca, S. Müller, A-quasiconvexity, lower semicontinuity, and Young measures, *SIAM Journal on Mathematical Analysis*, 30.6 (1999), 1355-1390.

[68] J. Frehse, A note on the Hölder continuity of solutions of variational problems, *in In Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg, Springer Berlin/Heidelberg*, 43.1 (1975), 59-63.

[69] J. Friberg, Multi-quasielliptic polynomials, *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 21.2 (1967), 239-260.

[70] D. V. Gennadii, S. V. Upsenskii, Partial differential equations and systems not solvable with respect to the highest-order derivative *CRC Press*, 2003.

[71] P. Gérard, Microlocal defect measures, *Communications in Partial Differential Equations*, 16.11 (1991), 1761-1794.

[72] M. Giaquinta, Multiple integrals in the calculus of variations and nonlinear elliptic systems *Princeton University Press*, 1983.

[73] M. Giaquinta, E. Giusti, Nonlinear elliptic systems with quadratic growth, *Manuscripta Mathematica*, 24.3 (1978), 323-349.

[74] M. Giaquinta, E. Giusti, On the regularity of the minima of variational integrals, *Acta Mathematica*, 148.1 (1982), 31-46.

[75] M. Giaquinta, G. Modica, Partial regularity of minimizers of quasiconvex integrals, *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*, 3.3 (1986), 185-208.

[76] D. Gilbarg, N. S. Trudinger, Elliptic partial differential equations of second order *Springer*, 2015.

[77] E. Giusti, Equazioni quasi ellittiche e spazi $\mathfrak{L}^{p,\theta}(\Omega, \delta)$ (I), *Annali di Matematica Pura ed Applicata*, 75.1 (1967), 313-353.

[78] E. Giusti, Equazioni quasi ellittiche e spazi $\mathfrak{L}^{p,\theta}(\Omega, \delta)$ (II), *Annali di Matematica Pura ed Applicata*, 21.3 (1967), 353-372.

[79] E. Giusti, Direct methods in the calculus of variations *World Scientific*, 2003.

[80] E. Giusti, M. Miranda, Un esempio di soluzioni discontinue per un problema di minimo relativo ad un integrale regolare del calcolo delle variazioni, *Bollettino dell'Unione Matematica Italiana*, 4.1 (1968), 219-226.

[81] E. Giusti, M. Miranda, Sulla regolarita delle soluzioni deboli di una classe di sistemi ellittici quasi-lineari, *Archive for Rational Mechanics and Analysis*, 31.3 (1968), 173-184.

[82] F. Gmeineder, J. Kristensen, Partial Regularity for BV Minimizers, *Archive for Rational Mechanics and Analysis*, 232.3 (2019), 1429-1473.

[83] Y. Grabovsky, From microstructure-independent formulas for composite materials to rank-one convex, non-quasiconvex functions, *Archive for Rational Mechanics and Analysis*, 227.2 (2018), 607-636.

[84] O. A. Hafsa, J.-P. Mandallena, Relaxation theorems in nonlinear elasticity, *Annales de l'Institut Henri Poincare (C) Non Linear Analysis, Elsevier Masson*, 25.1 (2008), 135-148.

[85] D. Hilbert, Mathematische probleme, *Nachrichten von der Koniglichen Gesellschaft der Wissenschaften zu Göttingen*, (1900), 253-297.

[86] G. N. Hile, Fundamental solutions and mapping properties of semielliptic operators, *Mathematische Nachrichten*, 279.13-14 (2006), 1538-1564.

[87] G. N. Hile, C. P. Mawata, C. Zhou, A priori bounds for semielliptic operators, *Journal of Differential Equations*, 176.1 (2001): 29-64.

[88] L. Hörmander, Linear Partial Differential Operators, *Springer-Verlag*, 1969.

[89] L. Hörmander, The Analysis of Linear Partial Differential Operators II, *Springer-Verlag*, 1990.

[90] V. P. Il'in, Conditions of validity of inequalities between L^p-norms of partial derivatives of functions of several variables, *Trudy Matematicheskogo Instituta imeni VA Steklova*, 96 (1968), 205-242.

[91] T. Iwaniec, A. Lutoborski, Integral estimates for null Lagrangians, *Archive for Rational Mechanics and Analysis*, 125.1 (1993), 25-79.

[92] A. Kałamajska, M. Kružík, Oscillations and concentrations in sequences of gradients, *ESAIM: Control, Optimisation and Calculus of Variations*, 14.1 (2008), 71-104.

[93] K. Kazaniecki, D. M. Stolyarov, M. Wojciechowski, Anisotropic Ornstein noninequalities, *Analysis & PDE*, 10.2 (2017), 351-366.

[94] D. Kinderlehrer, P. Pedregal, Characterizations of Young measures generated by gradients, *Archive for rational mechanics and analysis*, 115.4 (1991), 329-365.

[95] D. Kinderlehrer, P. Pedregal, Gradient Young measures generated by sequences in Sobolev spaces, *Journal of Geometric Analysis*, 4.1 (1994), 59-90.

[96] B. Kirchheim, J. Kristensen, On rank one convex functions that are homogeneous of degree one, *Archive for rational mechanics and analysis*, 221.1 (2016), 527-558.

[97] V. I. Kolyada, On embedding theorems, *Nonlinear Analysis, Function Spaces and Applications*, (2007), 35-94.

[98] V. I. Kolyada, F. J. Pérez, Estimates of difference norms for functions in anisotropic Sobolev spaces, *Mathematische Nachrichten*, 267.1 (2004), 46-64.

[99] J. Kristensen, Lower semicontinuity of quasi-convex integrals in BV, *Calculus of Variations and Partial Differential Equations*, 7.3 (1998), 249-261.

[100] J. Kristensen, Lower semicontinuity in spaces of weakly differentiable functions, *Mathematische Annalen*, 313.4 (1999), 653-710.

[101] J. Kristensen, On the non-locality of quasiconvexity, *Annales de l'Institut Henri Poincare (C) Non Linear Analysis, Elsevier Masson*, 16.1 (1999), 1-13.

[102] J. Kristensen, A necessary and sufficient condition for lower semicontinuity, *Nonlinear Analysis: Theory, Methods & Applications*, 120 (2015), 43-56.

[103] J. Kristensen, Nonlinear analysis & applications, Lecture notes for a course given at the University of Oxford, 2015.

[104] J. Kristensen, Young measures for oscillation and concentration, Lecture notes for a course given at the University of Parma, 2017.

[105] J. Kristensen, G. Mingione, The singular set of minima of integral functionals, *Archive for Rational Mechanics and Analysis*, 180.3 (2006), 331-398.

[106] J. Kristensen, G. Mingione, The singular set of Lipschitzian minima of multiple integrals, *Archive for Rational Mechanics and Analysis*, 184.2 (2007), 341-369.

[107] J. Kristensen, F. Rindler, Characterization of generalized gradient Young measures generated by sequences in $\mathrm{W}^{1,1}$ and BV, *Archive for Rational Mechanics and Analysis*, 197.2 (2010), 539-598.

[108] J. Kristensen, A. Taheri, Partial regularity of strong local minimizers in the multi-dimensional calculus of variations, *Archive for rational mechanics and analysis*, 170.1 (2003), 63-89.

[109] M. Kružík, T. Roubíček, Explicit Characterization of L^p-Young Measures, *Journal of mathematical analysis and applications*, 198.3 (1996), 830-843.

[110] M. Kružík, T. Roubíček, On the measures of DiPerna and Majda, *Mathematica Bohemica*, 122.4 (1997), 383-399.

[111] K. Kuratowski, C. Ryll-Nardzewski A general theorem on selectors, *Bulletin de l'Académie polonaise des sciences. Série des sciences mathématiques, astronomiques, et physiques*, 13.1 (1965), 397-403.

[112] O. A. Ladyzhenskaya, N. N. Ural'tseva, Linear and quasilinear elleptic equations *Academic Press*, 1968.

[113] M. Lavrentieff, Sur quelques problèmes du calcul des variations, *Annali di Matematica Pura ed Applicata*, 4.1 (1927), 7-28.

[114] H. Lebesgue, Intégrale, longueur, aire, *Annali di Matematica Pura ed Applicata*, 7.1 (1902), 231-359.

[115] J. L. Lions, E. Magenes, Non-homogeneous boundary value problems and applications. Vol. 1 *Springer Science & Business Media*, 2012.

[116] P. Marcellini, Approximation of quasiconvex functions, and lower semicontinuity of multiple integrals, *Manuscripta Mathematica*, 51.1 (1985), 1-28.

[117] P. Marcellini, On the definition and the lower semicontinuity of certain quasiconvex integrals, *Annales de l'IHP, Analyse non linéaire*, 3.5 (1986), 391-409.

[118] P. Marcellini, Alcuni recenti sviluppi nei problemi 19-simo e 20-simo di Hilbert, *Bollettino della Unione Matematica Italiana-A*, 2 (1997), 323-352.

[119] P. Marcellini, C. Sbordone, Relaxation of non convex variational problems, *Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti*, 63.5 (1977), 341-344.

[120] P. Marcellini, C. Sbordone, Semicontinuity problems in the calculus of variations, *Nonlinear Analysis: Theory, Methods & Applications*, 4.2 (1980), 241-257.

[121] V.G. Maz'ya, Examples of nonregular solutions of quasilinear elliptic equations with analytic coefficients, *Functional Analysis and its Applications*, 2.3 (1968), 230-234.

[122] E. J. McShane, Generalized curves, *Duke Mathematical Journal*, 6.3 (1940), 513-536.

[123] E. J. McShane, Necessary conditions in the generalized curve problem of the calculus of variations, *Duke Mathematical Journal*, 7 (1940), 1-27.

[124] L. Mejlbro, F. Topsøe, A precise Vitali theorem for Lebesgue measure, *Mathematische Annalen*, 230.2 (1977), 183-193.

[125] N. G. Meyers, Quasi-convexity and lower semi-continuity of multiple variational integrals of any order, *Transactions of the American Mathematical Society*, 119.1 (1965), 125-149.

[126] G. Mingione, Regularity of minima: an invitation to the dark side of the calculus of variations, *Applications of mathematics*, 51.4 (2006), 355-425.

[127] G. Mingione, D. Mucci, Integral functionals and the gap problem: sharp bounds for relaxation and energy concentration, *SIAM journal on mathematical analysis*, 36.5 (2005), 1540-1579.

[128] C. Mooney, O. Savin, Some singular minimizers in low dimensions in the calculus of variations, *Archive for Rational Mechanics and Analysis*, 221.1 (2016), 1-22.

[129] C. B. Morrey, Quasiconvexity and the lower semicontinuity of multiple integrals, *Pacific journal of mathematics*, 2.1 (1952), 25-53.

[130] C. B. Morrey, Partial regularity results for non-linear elliptic systems, *Journal of Mathematics and Mechanics*, 17.7 (1968), 649-670.

[131] C. B. Morrey, Multiple integrals in the calculus of variations *Springer Science & Business Media*, 2009.

[132] S. Müller, Rank-one convexity implies quasiconvexity on diagonal matrices, *International Mathematics Research Notices*, 20 (1999), 1087-1095.

[133] S. Müller, Variational models for microstructure and phase transitions in Calculus of variations and geometric evolution problems *Springer, Berlin, Heidelberg*, (1999), 85-210.

[134] S. Müller, V. Šverák, Convex integration for Lipschitz mappings and counterexamples to regularity, *Annals of mathematics*, (2005), 715-742.

[135] F. Murat, Compacité par compensation, *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 5.3 (1978), 489-507.

[136] F. Murat, Compacité par compensation: condition nécessaire et suffisante de continuité faible sous une hypothese de rang constant, *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 8.1 (1981), 69-102.

[137] J. Nash, Continuity of solutions of parabolic and elliptic equations, *American Journal of Mathematics*, 80.4 (1958), 931-954.

[138] J. Nečas, Example of an irregular solution to a nonlinear elliptic system with analytic coefficients and conditions for regularity in Theory of nonlinear operators, *Proceedings of the Fourth International Summer School, Academy of Science, Berlin, 1975*, (1977), 197-206.

[139] S. M. Nikol'skii, Inequalities for entire functions of finite degree and their application in the theory of differentiable functions of several variables, *Trudy Matematicheskogo Instituta imeni VA Steklova*, 38 (1951), 244-278.

[140] P. Pedregal, Jensen's inequality in the calculus of variations, *Differential Integral Equations*, 7.1 (1994) 57-72.

[141] P. Pedregal, Parametrized measures and variational principles, *Birkhäuser*, (1997).

[142] A. Pełczyński, Boundedness of the canonical projection for Sobolev spaces generated by finite families of linear differential operators, Analysis at Urbana, vol. I, London Mathematical Society Lecture Note Series 137 *Cambridge University Press*, 1989, 395-415.

[143] A. Pełczyński, K. Senator, On isomorphisms of anisotropic Sobolev spaces with classical Banach spaces and a Sobolev type embedding theorem, *Studia Mathematica*, 84.2 (1986), 169-215.

[144] A. Pełczyński, K. Senator, Addendum to the paper 'On isomorphisms of anisotropic Sobolev spaces with classical Banach spaces and a Sobolev type embedding theorem', *Studia Mathematica*, 84.2 (1986), 217-218.

[145] A. Prosinski, Closed $\mathcal{A}$-p Quasiconvexity and Variational Problems with Extended Real-Valued Integrands, *ESAIM: Control, Optimisation and Calculus of Variations*, 24.4 (2018), 1605-1624.

[146] A. Prosinski, B. Raiţă, On the well-posedness of some variational problems, (*in preparation*).

[147] F. Rindler, Directional oscillations, concentrations, and compensated compactness via microlocal compactness forms, *Archive for Rational Mechanics and Analysis*, 215.1 (2015), 1-63.

[148] F. Rindler, Calculus of Variations *Springer International Publishing*, 2018.

[149] S. Saks, Theory of the Integral *Hafner Publishing Company*, 1937.

[150] F. Santambrogio, Optimal transport for applied mathematicians *Birkäuser*, 2015.

[151] J. Serrin, A new definition of the integral for non-parametric problems in the calculus of variations, *Acta Mathematica*, 102.1 (1959), 23-32.

[152] J. Serrin, On the definition and properties of certain variational integrals, *Transactions of the American Mathematical Society*, 101.1 (1961), 139-167.

[153] V. Shevchik, An a priori estimate for a model semi-elliptic differential operator, *Journal of mathematical analysis and applications*, 268.2 (2002), 385-399.

[154] L. N. Slobodeckii, Generalized Sobolev spaces and their application to boundary problems for partial differential equations, *Leningradskii Gosudarstvennyi Pedagogiceskii Institut imeni A. I. Gercena. Ucenye Zapiski*, 197 (1958), 54-112.

[155] L. N. Slobodeckii, S. L. Sobolev's spaces of fractional order and their application to boundary problems for partial differential equations, *Doklady Akademii Nauk SSSR*, 118 (1958), 243-246.

[156] V. A. Solonnikov, Inequalities for functions of the classes $\overrightarrow{\mathrm{W}}_p(\mathbb{R}^n)$, *Journal of Mathematical Sciences*, 3.4 (1975), 549-564.

[157] V. Šverák, Rank-one convexity does not imply quasiconvexity, *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 120.1 (1992), 185-189.

[158] V. Šverák, X. Yan, A singular minimizer of a smooth strongly convex functional in three dimensions, *Calculus of Variations and Partial Differential Equations*, 10.3 (2000), 213-221.

[159] V. Šverák, X. Yan, Non-Lipschitz minimizers of smooth uniformly convex functionals, *Proceedings of the National Academy of Sciences of the United States of America*, 99.24 (2002), 15269.

[160] M. A. Sychev, A new approach to Young measure theory, relaxation and convergence in energy, *Annales de l'Institut Henri Poincare (C) Non-Linear Analysis, Elsevier Masson*, 16.06 (1999), 773-812.

[161] M. A. Sychev, Characterization of homogeneous gradient Young measures in case of arbitrary integrands, *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 29.3 (2000), 531-548.

[162] L. Székelyhidi Jr, The regularity of critical points of polyconvex functionals, *Archive for rational mechanics and analysis*, 172.1 (2004), 133-152.

[163] L. Székelyhidi Jr., E. Wiedemann, Young measures generated by ideal incompressible fluid flows, *Archive for Rational Mechanics and Analysis*, 206.1 (2012), 333-366.

[164] L. Tartar, Compensated compactness and applications to partial differential equations, *Nonlinear analysis and mechanics, Heriot-Watt symposium, Pitman*, 4 (1979), 136-211.

[165] L. Tartar, The compensated compactness method applied to systems of conservation laws in Systems of nonlinear partial differential equations, *Springer, Dordrecht*, 1983, 263-285.

[166] L. Tartar, Étude des oscillations dans les équations aux dérivées partielles non linéaires in Systems of nonlinear partial differential equations, *Springer, Berlin, Heidelberg*, 1984, 384-412.

[167] L. Tartar, H-measures, a new approach for studying homogenisation, oscillations and concentration effects in partial differential equations, *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 115.3-4 (1990), 193-230.

[168] L. Tartar, On mathematical tools for studying partial differential equations of continuum physics: H-measures and Young measures in *Developments in partial differential equations and applications to mathematical physics, Springer, Boston, MA*, (1992), 201-217.

[169] L. Tonelli, Sur une méthode directe du calcul des variations, *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, 39.1 (1915), 233-264.

[170] L. Tonelli, La semicontinuita nel calcolo delle variazioni, *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, 44.1 (1920), 167-249.

[171] H. Triebel, A priori estimates and boundary value problems for semi-elliptic differential equations: A model case, *Communications in Partial Differential Equations*, 8.15 (1983), 1621-1664.

[172] M. Troisi, Problemi al contorno con condizioni omogenee per le equazioni quasi-ellittiche, *Annali di Matematica Pura ed Applicata*, 90.1 (1971), 331-412.

[173] E. Turco, I. Giorgio, A. Misra, F. dell'Isola, King post truss as a motif for internal structure of (meta) material with controlled elastic properties, *Royal Society open science*, 4.10 (2017), 171153.

[174] L. R. Volevich, A class of hypoelliptic systems in Doklady Akademii Nauk 134.6 *Russian Academy of Sciences*, (1960), 1275-1278.

[175] L. R. Volevich, Local properties of solutions of quasi-elliptic systems, *Matematicheskii Sbornik*, (1962), 3-52.

[176] M. Wagner, Mehrdimensionale Steuerungsprobleme mit quasikonvexen Integranden, *Habilitation Thesis, BTU Cottbus*, (2006).

[177] M. Wagner, On the lower semicontinuous quasiconvex envelope for unbounded integrands (I), *ESAIM: Control, Optimisation and Calculus of Variations*, 15.1 (2009), 68-101.

[178] M. Wagner, On the lower semicontinuous quasiconvex envelope for unbounded integrands (II): Representation by generalized controls, *Journal of Convex Analysis*, 16 (2009), 441-472.

[179] K. O. Widman, Hölder continuity of solutions of elliptic systems, *Manuscripta Mathematica*, 5.4 (1971), 299-308.

[180] L. C. Young, Generalized curves and the existence of an attained absolute minimum in the calculus of variations, *Comptes Rendus de la Société des Sciences et des Lettres de Varsovie*, 30 (1937), 212-234.

[181] L. C. Young, Generalized surfaces in the calculus of variations, *Annals of mathematics*, (1942), 84-103.

[182] L. C. Young, Generalized surfaces in the calculus of variations II, *Annals of mathematics*, (1942), 530-544.

[183] L. C. Young, Lectures on the calculus of variations and optimal control theory *American Mathematical Society*, 2000.

[184] K. Zhang, A construction of quasiconvex functions with linear growth at infinity, *Annali della Scuola Normale Superiore di Pisa — Classe di Scienze*, 19.3 (1992), 313-326.