Problems of Stability and Well-Posedness in the Calculus of Variations and Related PDEs



Panas Kalayanamit
The Queen's College
University of Oxford

under the supervision of
Prof. Jan Kristensen

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

Hilary 2022

Contents

Abstract 1

Acknowledgement 2

Introduction 3

1 Preliminaries 10
- 1.1 Well-posedness and Multivalued Functions . . . 12
- 1.2 The Direct Method in the Calculus of Variations and Quasiconvexity 22
- 1.3 Young Measure and Lower Semicontinuity . . . 31

2 Generic Well-posedness with Respect to Linear Perturbations 37
- 2.1 Baire's Category Theorem and σ-Small Sets . . . 40
- 2.2 Convex Analysis and Geometries of Banach Spaces . . . 44
- 2.3 Applications of Convex Analysis to Linearly Perturbed Variational Integrals . . . 51

3 Well-Posedness of Strictly Quasiconvex Variational Problems with Varying Dirichlet Data 58
- 3.1 Caccioppoli Inequalities and Uniform Integrability . . . 61
- 3.2 Upper Semicontinuity of the Solution Map $\mathcal{S}$ . . . 88

4 Rectifiability of the Sets of Extremals for a Singularly Perturbed Variational Integrals 94
- 4.1 Extension Operators and Higher Integrability . . . 97
- 4.2 Nonlinear Gårding Inequality . . . 109
- 4.3 The Case $\Phi(|z|) = E_p(z)$ for $p > 1$ . . . 113
- 4.4 Regularized Solutions . . . 121
- 4.5 Partial Convexity and the Sets of Extremals of J_ε . . . 125

5 Further Developments and Related Results 132

A Appendix 137
- A.1 Weak Topologies and Compactness . . . 137
- A.2 Convergence and Compactness in $\mathrm{L}^p(\Omega)$ . . . 139
- A.3 Bochner Integrals, Vector Measures and the Radon–Nikodym Property (RNP) . . . 142

B References 148

Index 157

Abstract

We present a collection of three closely related topics regarding the stability and well-posedness of minimization problems in the calculus of variations, namely the generic Tykhonov well-posedness with respect to linear perturbations, the generalized Hadamard well-posedness for a class of Dirichlet quasiconvex variational problems, and the d-rectifiability of the set of minimizers for a singularly perturbed quasiconvex variational integral. Regarding the first topic on generic Tykhonov well-posedness, we prove that by introducing a linear perturbation term ξ to a variational integral of the form $J(u) = \int_\Omega F(x, u, \nabla u)\,dx$ satisfying a mild growth condition on the negative part of the integrand F, and a coercivity condition for J, the perturbed functional $J - \xi$ admits a unique strong minimizer in the Dirichlet class $\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$ for all ξ in a dense G_δ-subset of $\mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)^*$ for any $p \geq 1$. The proof for the case $p = 1$ is particularly distinctive and is based on a duality between the lower semicontinuity of J on sets with the Radon–Nikodym Property and the generic differentiability of its Fenchel conjugate J^*. As for the second topic on generalized Hadamard well-posedness, we show that under the p-growth condition and the uniform strict quasiconvex condition for the integrand F of J, the solution map $\mathcal{S}\colon g \mapsto \operatorname{argmin}\{J(u) : u \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)\}$ enjoys some nice semicontinuity property as a multivalued function $\mathcal{S}\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \rightrightarrows \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$. Lastly, we study the set $\mathcal{M}$ of all the minimizers for a singularly perturbed strongly quasiconvex variational integral $J_\varepsilon(u) = \int_\Omega F(\nabla u)\,dx + \varepsilon \int_\Omega |\nabla^2 u|^2\,dx$ from the Dirichlet class $\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$ and show that when F satisfies the p-growth condition for a $p \in \left[2, \frac{n+2}{n-2}\right)$ when $n > 2$ (no condition on $p \in [1,\infty)$ for $n = 2$), $\mathcal{M}$ is a d-rectifiable set (in Federer's sense), or, more precisely, $\mathcal{M}$ is bi-Lipschitz equivalent to a compact subset of $\mathbb{R}^d$. This result is based on an extension of a technique from a recent work by Campos Cordero and Kristensen on questions on uniqueness.

Acknowledgement

First and foremost, I would like to express my deepest gratitude to the best supervisor I could have hoped for, Prof. Jan Kristensen. The quality of this research would not have been anywhere near what it is now without his valuable insights and kind support throughout my years in Oxford. Secondly, I would like to thank each and every professors from the OxPDE research group who imparted their vast knowledge upon us, the PDE CDT students. A huge thank you to Prof. Melanie Rupflin, my mini-project advisor, in particular for all the great discussions during the short period of time we worked together.

I would also like to thank all the mathematicians I knew from Chaing Mai University who made my mathematical journey possible. A special thank to Asst. Prof. Somlak Utudee, my first advisor, who introduced me to the beauty of differential equations. Thank you to Assoc. Prof. Smai Yodintra, Assoc. Prof. Sanun Supasai and Asst. Prof. Mullika Tawonatiwas for shaping me into the mathematician I am today. Last but not least, I am forever indebted to Assoc. Prof. Chulin Likasiri for everything she has done for me.

Outside of my academic life, there are countless friends who have played significant parts in my adventure so far, both directly and indirectly. I would like to thank all my PDE CDT cohorts who shared unforgettable experiences with me during the last few years, especially my good friend Timo Sprekeler. For those friends who live far away, let me name a few in particular: thank you Ukrit Keyen, Hariphan Philathong, Boriwat Pijarana and Sataban Srisuriyatada.

My greatest appreciation goes to my family. It is impossible to overstate the contributions from my parents and my brother who always give me supports and encouragements. Also, I would like to thank my late grandfather for being a big inspiration for me and for all the good memories he left behind.

I am truly grateful for the DPST Scholarship. Without their funding, my DPhil studies at the University of Oxford would not have been possible.

Introduction

The topic of well-posedness and stability in the calculus of variations has gathered immense interest since the early days of the calculus of variations, encompassing important issues such as uniqueness of solutions, or the effects of various kinds of perturbations to a variational problem. Calculus of variations originated from problems in physics and geometry, so the stability of solutions to variational problems is not merely a theoretical curiosity, but is also of practical interest. Well-posedness and stability are words used to describe a broad range of concepts that have many different (but closely related) meanings, and we will be studying a few of these notions in this thesis. The first topic we shall encounter is Tykhonov well-posedness of variational integrals under linear perturbations, then we shall present the continuity (in the sense of multivalued functions) of the solution map of a quasiconvex variational integral with varying Dirichlet boundary data. We then study the structure of the set of minimizers to a singularly perturbed variational integral and the regularity of these minimizers. Lastly, we will identify some plausible research directions that can be built on the work done in this thesis.

As a branch of mathematics that has its root in a field of practical science, it is not surprising that many topics of interest in the calculus of variations are deeply connected to real-world problems. Over the years, calculus of variations has proved itself to be a valuable tool for solving problems from various disciplines, both within and outside of mathematics. A minimizer (or an extremal) of a variational integral could correspond to an equilibrium point of a physical process, a stable configuration of some system of interest, or an optimal method to reach some desired outcome. Hence understanding how minimizers of such integrals behave under various kinds of perturbations is of critical importance.

The study of how the minimizers of an integral functional vary in response to perturbations of various kinds is commonly known as the study of well-posedness of a minimization problem. The idea could be dated back to many classical works by well-known mathematicians, notably Hadamard's work on the well-posedness of problems from mathematical physics. Nowadays, the word 'well-posedness' has become a broad term that assumes many diverse meanings in different areas of mathematics, but they all encapsulate the idea that solutions of a problem (or a family of interconnected problems) are 'nice' in some specific senses.

Tykhonov well-posedness under linear perturbations of a variational integral $J\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \to \overline{\mathbb{R}}$ of the form $J(u) = \int_\Omega F(x,u,\nabla u)\,dx$ is the first type of well-posedness covered in this thesis. Here $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ denotes the space of $\mathbb{R}^N$-valued Sobolev functions with exponent $p \in [1,\infty)$ on a domain $\Omega \subset \mathbb{R}^n$, and $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$. Regarding the well-posedness under linear perturbations, our motivation is from a result in [91] by Stegall, which is commonly known today as the Stegall

Variational Principle. The theorem states that, under the assumption of strong lower semicontinuity and a coercivity condition, an extended real-valued function J on a suitable Banach space X can be perturbed by a small linear functional ξ so that the perturbed functional attains a strong minimum. So, in a sense, a 'bad' minimization problem is only a small linear perturbation away from being a good one. This is somewhat surprising since usually one would need to assume weak sequential lower semicontinuity to conclude that J has a minimizer. There are other variational principles of a similar kind (cf. [13], [39], [61], [72]), most notably the Ekeland Variational Principle in [38], which has seen much usage in the calculus of variations.

The second main topic is about the continuity property of the solution map $\mathcal{S}\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \rightrightarrows \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ that sends a function g to the set of minimizers $\operatorname{argmin}\{J(u) : u \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)\}$, where the integrand $F(x,u,z)$ of J now is assumed quasiconvex in the last variable (the symbol $\rightrightarrows$ signifies that $\mathcal{S}$ here is a multivalued function, a notion to be defined later). Here, the behaviour of the set of all minimizers of a variational integral $J(u) = \int_\Omega F(x,u,\nabla u)\,dx$ in the Dirichlet class $\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$ is described by analysing $\mathcal{S}(g)$ as a multivalued function between topological spaces. Our ultimate goal is to use the results from this chapter to prove the generic uniqueness of the problem

$$\text{minimize } (J-\xi)(u) \quad \text{over} \quad u \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N),$$

a topic that we will come back to discuss in the final chapter.

Our third main topic is about the structure of the sets of extremals and minimizers of a singularly perturbed quasiconvex variational integral J_ε. The reason for adding a singular perturbation term to our original variational integral is to regularize the extremals and naturally obtain a uniform Hölder bound. This chapter also contains a couple of results that could be of independent interest, such as the nonlinear Gårding inequality for quasiconvex integrals and the existence of linear extension operators with certain desirable properties.

The final chapter is a discussion about the potential developments of some of the results from earlier chapters. The obstacles to certain unrealized goals in this thesis are identified, together with some proposals on how these problems could be overcome. Apart from that, we make a couple of observations on how some results that appear in the thesis could be strengthened, together with a few propositions that may be helpful for that purpose. This chapter concludes the main body of this thesis.

Outlines

Here we will give a short summary of each chapter in this thesis. The main theorems from the chapters will not be stated fully here since that will be done later at the beginning of each chapter separately. This section acts as a short guide on how the contents of this thesis fit together and provides an outline of the whole work. The discussions in this section are sometimes a bit vague and one should consult the respective parts of this thesis for the full definition of a term, or the statement of a proposition or a theorem.

1. Preliminaries

This chapter contains the background results and summarizes the basic notation and technical terms that are used throughout the rest of the thesis. The results found in this chapter are entirely from the existing literature, with perhaps shorter or simplified proofs. I make no claim to the originality of anything in this chapter.

The first topic in this chapter is *multivalued functions and well-posedness*, where the definition and properties of multivalued functions and a few notions of well-posedness for general minimization problems are introduced. A *multivalued function* $T\colon X \rightrightarrows Y$ is a function from X into the power set of Y, i.e. $T\colon X \to 2^Y$. When X and Y are topological spaces, we can generalize the usual definition of continuity for functions from X into Y to various different continuity notions for multivalued functions, serving various different purposes. Some general theorems about these continuity properties are presented. The use of multivalued functions allows us to state a few definitions and results regarding well-posedness more clearly. We also describe some relevant forms of well-posedness that are studied in this thesis. In particular, Tykhonov well-posedness, Hadamard well-posedness and their generalizations are introduced, together with some examples.

The next part of this chapter is a short exposition on the *direct method in the calculus of variations and quasiconvexity*. Here, we examine the conditions that ensure the existence of a minimizer for a variational integral of type $J(u) = \int_\Omega F(x, u, \nabla u)\,dx$, viewed as a nonlinear functional on a Sobolev space $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$ of vector-valued functions. The growth conditions for the integrand F is one of the main issues, together with the *quasiconvexity* (in the last variable) of $F(x, u, z)$. Other closely related convexity notions such as *polyconvexity* and *rank-one convexity* are also briefly mentioned.

Finally, this chapter ends with a brief introduction to *Young measures* and some of their important properties. Primarily, we focus on the relationship between quasiconvexity and homogeneous gradient Young measure, and its implication on the (sequential) lower semicontinuity of the variational integral $J(u) = \int_\Omega F(x, u, \nabla u)\,dx$ in the weak topology of $\mathrm{W}^{1,p}$.

2. Generic Well-Posedness with Respect to Linear Perturbations

The central topic of this chapter is about how a variational integral, viewed as an $\overline{\mathbb{R}}$-valued nonlinear functional $J\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \to \overline{\mathbb{R}}$, can be perturbed by a linear functional $\xi \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)^*$ so that the minimization problem of the perturbed functional $J-\xi$ behaves better than the original unperturbed problem. The validity of such perturbations is presented in a general setting using techniques from convex analysis, in particular the properties of the subdifferential ∂J^* of the Fenchel conjugate of the functional J, where ∂J^* is viewed as a multivalued function. We then apply the result to variational integrals J of the form $J(u) = \int_\Omega F(x,u,\nabla u)\,dx$. Note that no quasiconvexity assumption is needed on F in this chapter.

This chapter begins with an overview of some known results from convex analysis related to generic uniqueness, in particular the properties of subdifferentials and Fenchel conjugates. The generic Tykhonov well-posedness of the problem $\min(X; J-\xi)$, i.e.

$$\text{minimize } (J-\xi)(u) \quad \text{over} \quad u \in X,$$

is transformed into a problem about the (sub)differentiability of J^* at $\xi \in X^*$, which at a deeper level is connected to the geometries of X and X^*. This is a compilation of established results, no original work is being added in this part.

We then apply the general result about the generic well-posedness of the problem $\min(X; J-\xi)$ to the case when J is a variational integral $J(u) = \int_\Omega F(x,u,\nabla u)\,dx$, defined for $u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ with $p \in [1,\infty)$. The case $p > 1$ is more or less a direct application of the theorems from the previous section, but the case $p = 1$ is more involved since the space $X = \mathrm{W}^{1,1}(\Omega;\mathbb{R}^N)$ does not have the Radon-Nikodym Property. In particular, we prove that when the integrand F of J satisfies a certain coercivity condition, the problem

$$\text{minimize } (J-\xi)(u) \quad \text{over} \quad u \in \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N),$$

has a unique solution for almost every $\xi \in \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)^*$ and that the solution is well-behaved. The modifications needed to deal with minimization problems in more general Dirichlet classes $\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$ and constrained minimization problems are also discussed.

3. Well-Posedness of Strictly Quasiconvex Variational Problems with Varying Dirichlet Data

In this chapter, the well-posedness of the problem $\min(\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N); J)$, i.e.

$$\text{minimize } \int_\Omega F(x,u,\nabla u)\,dx \quad \text{over} \quad u \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N),$$

where the integrand $F(x, u, z)$ of J is strictly quasiconvex and has p-growth in the z variable, with varying Dirichlet boundary data $g \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$ is studied. The strict quasiconvexity assumption is used to obtain a uniform integrability estimate for the family

$$\left\{ F(\cdot, u_j, \nabla u_j) : u_j \text{ minimizes } J \text{ in } \mathrm{W}^{1,p}_{g_j}(\Omega; \mathbb{R}^N) \right\}$$

(where g_j is a $\mathrm{W}^{1,p}$-convergent sequence), which is then used to prove the semi-continuity of the solution map $\mathcal{S}\colon \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N) \rightrightarrows \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$, where $\mathcal{S}(g) := \operatorname{argmin}\{J(u) : u \in \mathrm{W}^{1,p}_g(\Omega; \mathbb{R}^N)\}$.

In order to get the well-posedness result for the quasiconvex variational problem, we first establish Caccioppoli-type inequalities both for cubes in the interior of our domain Ω and the cubes near the boundary $\partial\Omega$. The technique used here is a variant of the usual 'hole-filling' argument, but with some additional subtleties arising because our Caccioppoli inequalities must hold uniformly for minimizers in different Dirichlet classes.

Through a Gehring-type lemma, the Caccioppoli inequalities are then used in proving a uniform integrability estimate for the family

$$\left\{ |\nabla u_j|^p : u_j \text{ minimizes } J \text{ in } \mathrm{W}^{1,p}_{g_j}(\Omega; \mathbb{R}^N) \right\}$$

whenever g_j is strongly $\mathrm{W}^{1,p}$ convergent, which then implies the uniform integrability of $\{F(\cdot, u_j, \nabla u_j) : u_j \text{ minimizes } J \text{ in } \mathrm{W}^{1,p}_{g_j}(\Omega; \mathbb{R}^N)\}$. This approach is inspired by an estimate due to D'Onofrio and Iwaniec [34] that was used in a different context, and our main uniform integrability estimate is a suitable modification of their work.

Lastly, this chapter concludes by utilizing the above uniform integrability result to prove that the solution map $\mathcal{S}$ is a USCO (upper semi-continuous with compact values) map, hence showing that the problem $\min(\mathrm{W}^{1,p}_g(\Omega; \mathbb{R}^N); J)$ is Hadamard well-posed in the generalized sense (with respect to the parameter g). The use of Young measures in the proof may not be strictly necessary, but it does help to clarify the role that uniform integrability plays regarding the continuity of the nonlinear functional J. Most importantly, it is plausible that the results from this chapter might be crucial for a generic uniqueness result that we have not been able to prove yet. This will be discussed further in the fifth chapter.

4. Rectifiability of Sets of Extremals for Singularly Perturbed Quasiconvex Variational Integrals

In this chapter, we examine the properties of the set of all minimizers for a singularly perturbed variational integral of the form

$$J_\varepsilon(u) = \int_\Omega F(\nabla u)\,dx + \varepsilon \int_\Omega |\nabla^2 u|^2\,dx,$$

where F is a strongly quasiconvex integrand ($\nabla^2 u$ denotes the Hessian of u). Viewing the variational integral as a functional $J_\varepsilon\colon \mathrm{W}^{2,2}(\Omega;\mathbb{R}^N) \to \mathbb{R}$, we show that under suitable assumptions on the integrand F, the set $\mathcal{M}$ of all the minimizers of J_ε subjected to a $\mathrm{W}^{2,2}$ Dirichlet boundary condition is d-rectifiable (in Federer's sense). In particular, this means that $\mathcal{M}$ can be parametrized by a Lipschitz function $f\colon K \to \mathcal{M}$, where K is a compact set in some Euclidean space $\mathbb{R}^d$. The results in this chapter come from a joint work with Prof Jan Kristensen.

The nonlinear Gårding inequality that is crucial for understanding the structure of the set of extremals of J_ε is based on a recent work by Campos Cordero and Kristensen, where they establish the result for the quadratic $p = 2$ case. The proof of this inequality requires many ingredients, and one of them is the existence of an extension operator that extends Sobolev functions $u \in \mathrm{W}^{1,p}(\partial Q;\mathbb{R}^N)$ to $u \in \mathrm{W}^{1,p}(\mathbb{R}^n\backslash\overline{Q};\mathbb{R}^N)$ (Q is a cube in $\mathbb{R}^n$) which improves the integrability of $\Phi(|\nabla u|)$, where $\Phi\colon \mathbb{R}_+ \to \mathbb{R}_+$ is particular type of Young function. The first part of this chapter covers this topic and some closely related extension results.

Using the extension results, the Gårding inequality is then proved for variational integrals of the form $J(u) = \int_\Omega F(\nabla u)\,dx$ with strongly quasiconvex integrand F satisfying the p-growth condition. The proofs are divided into two cases, namely the case when $p > 1$ and when $p = 1$ since they require slightly different techniques to prove certain relevant estimates. A crucial step for the main results in this chapter is the strengthening of the Gårding inequality so that it implies strict convexity of the restriction of the functional J to the affine subsets $u + Y \subset \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$, where $u \in C^1(\overline{\Omega};\mathbb{R}^N)$ and $Y \subset \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)$ is some subspace of finite codimension. The proof of this strengthened form of the Gårding inequality utilizes a simultaneous Schauder basis of Sobolev spaces on Ω and a modified version of the standard Rayleigh quotient used in the eigenvalue problem for the Dirichlet Laplacian.

Next, the variational integral J is perturbed into J_ε by adding a small singular perturbation term, this regularizes the extremals of the variational integral so that they are of the class $C^{1,\alpha}$, which makes them smooth enough that we can apply the Gårding inequality. Here we prove that for a sufficiently smooth domain $\Omega \subset \mathbb{R}^n$ and any sufficiently smooth Dirichlet boundary datum $g \in \mathrm{W}^{3,r}(\Omega;\mathbb{R}^N)$, the extremals of J_ε in the class $\mathrm{W}^{3,r}_g(\Omega;\mathbb{R}^N)$ are as smooth as g. The techniques used here are a

variation of the usual elliptic estimates and a bootstrapping argument, which uses the fact that F is a rank-one convex integrand with p-growth in a nontrivial way.

Lastly, we end this chapter with the proof of the d-rectifiability of the sets of extremals and minimizers of J_ε, the main result of this chapter. This theorem is important because it shows that the set of minimizers of the perturbed variational integral J_ε is 'structured' even when the minimizers are not unique. The proof of this result uses the Schauder estimates for the extremals of the variational integral J_ε to get a strengthened form of Gårding's inequality that holds uniformly for all the minimizers of J_ε. This theorem on d-rectifiability may imply an even stronger well-posedness result, in particular the generic uniqueness of the solution map $\mathcal{S}\colon \mathrm{W}^{3,r}(\Omega;\mathbb{R}^N) \rightrightarrows \mathrm{W}^{2,2}(\Omega;\mathbb{R}^N)$ associated with this problem. A further discussion of this point is found in the last chapter.

5. Further Developments and Related Results

This last chapter is mostly a discussion of the possibility of developing the results from chapters 3 and 4 into a generic uniqueness result (similar to the ones in chapter 2), together with some related results. We identify the extra work that needs to be done on top of the existing results in order to reach the unrealized goal, and suggest some potential strategies that may lead to the desired outcome.

Regarding the generic uniqueness of minimizers of a quasiconvex variational integral $J(u) = \int_\Omega F(x,u,\nabla u)\,dx$ with varying Dirichlet boundary data, we state some results from the theory of multivalued functions that could be applied (or suitably modified so that it could be applied) to our problem. The continuity of the solution map $\mathcal{S}$ from chapter 3 need to be supplemented with some further properties, and the knowledge of the structure of the sets of minimizers of J_ε (from chapter 4) could make it easier to prove the corresponding theorem for the singularly perturbed case.

Appendix

In the appendix, some additional facts about functional analysis, measure theory and integration that are used throughout the thesis can be found.

1 Preliminaries

In this chapter, the basic notations and some background results are stated to be used as references for the rest of this thesis. There are three main sections in this chapter, namely *well-posedness and multivalued functions*, *the direct method in calculus of variations and quasiconvexity*, and *Young measures*. Most results are stated without proofs but with references, only some key results are accompanied by proofs. It should be emphasized that the terms and results in this chapters are not really new and I do not claim any of them to be my own original work. The proofs of some propositions in this chapter are modifications or simplified versions of known proofs.

In the study of well-posedness and stability in the calculus of variations, we seek to understand the behaviour of the minimizers (and the extremals) of variational integrals under various forms of perturbations. The primary variational integrals that appear in this work can be viewed as nonlinear integral functionals $J\colon S \to \overline{\mathbb{R}}$ of the general form

$$J(u) = \int_\Omega F(x, u(x), \nabla u(x))\, dx. \tag{1.1}$$

Here $\Omega \subset \mathbb{R}^n$ is a bounded open set with sufficiently smooth boundary, the integrand $F\colon \Omega \times \mathbb{R}^N \times \mathbb{R}^{N\times n} \to \overline{\mathbb{R}}$ will be assumed to satisfy certain natural assumptions regarding its growth, continuity, and measurability. The domain S of our functional J is a suitable subset of some function space of vector valued functions $u\colon \Omega \to \mathbb{R}^N$. Throughout the rest of this work, we will use the notation $\overline{\mathbb{R}}$ to denote the extended real line **excluding** $-\infty$, i.e. $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$.

Some examples of variational integrals J that fit into our framework are as follows.

Example 1.1.

1. Let $\Omega = [a, b]$, $S = \{u \in C^1([a, b]) : u(a) = \alpha, u(b) = \beta\}$ and J be

$$J(u) = \int_a^b \sqrt{\frac{1 + u'(x)^2}{u(x)}}\, dx.$$

   The minimization problem corresponding to this functional (and the domain S) is a classical problem named the *brachistochrone problem*. This is a problem of historical importance in the calculus of variations.

2. Let $\Omega \subset \mathbb{R}^3$, $S = \{u \in C^1(\overline{\Omega}) : u|_{\partial\Omega} = g\}$ and J be

$$J(u) = \int_\Omega \sqrt{1 + |\nabla u(x)|^2}\, dx.$$

   The minimizers of this functional within the Dirichlet class S is the (regular) minimal surfaces whose boundaries are given by a function g. This is an

example of a variational integral whose integrand $F(z) = \sqrt{1+|z|^2}$ satisfies the linear growth condition

$$F(z) \leq c(1+|z|).$$

We can generalize this by replacing $\mathbb{R}^3$ with $\mathbb{R}^{n+1}$ to get a variational principle for regular minimal hypersurfaces, or relaxing J and enlarging the class S even further to allow more general types of minimal hypersurfaces.

3. $\Omega \subset \mathbb{R}^3$, $S = \mathrm{W}^{1,2}_g(\Omega;\mathbb{R}^3)$, and J be

$$J(u) = \int_\Omega \left[|\nabla u(x)|^2 + h(\det \nabla u)\right]\, dx,$$

where h is a convex function. Here S is the Dirichlet class

$$\mathrm{W}^{1,2}_g(\Omega;\mathbb{R}^3) = \{u \in \mathrm{W}^{1,2}(\Omega;\mathbb{R}^3) : u - g \in \mathrm{W}^{1,2}_0(\Omega;\mathbb{R}^3)\}$$

with the boundary datum g. This is an example of a neo-Hookian energy for hyperelastic bodies. The integrand F in this case is strictly/strongly quasiconvex, which is a main assumption we will be working with from chapter 3 onward.

4. $\Omega \subset \mathbb{R}^n$, $S = \mathrm{W}^{1,2}_g(\Omega)$, and J be

$$J(u) = \int_\Omega \left[u(x)f(x) + \frac{1}{2}|\nabla u(x)|^2\right] dx,$$

where f is a smooth function. The extremals (critical points) of J weakly satisfy the corresponding Euler-Lagrange equation, which in this case is the Poisson equation

$$\Delta u = f.$$

This type of relationship between (elliptic) PDEs and minimization problems is of crucial importance for the result in chapter 5.

To study the well-posedness of the minimization problems in the calculus of variations, it is useful to be able to inspect these problems from a broader perspective, and to perform an in-depth analysis with the right tools. The first section of this chapter will be about terms and notation related to well-posedness for general variational problems and how they can be formulated as continuity properties of associated multivalued functions. After that, we will be more concrete and discuss topics that are more specific to variational integrals, namely the relationship between quasiconvexity (and other related notions) and the direct method in the calculus of variation. Lastly, Young measures are introduced as an effective tool to study the semicontinuity properties of a variational integral.

1.1 Well-posedness and Multivalued Functions

Well-posedness of a mathematical problem is a general concept that assumes many technical meanings in various contexts, even when we limit the scope of the discussion to the calculus of variations. Given a minimization problem, one can ask how the behaviour of the problem would change if we introduce a *perturbation* into it. The problem is said to be *well-posed* if small perturbations do not drastically affect its behaviour. To make the previous statement precise, we need to quantify the perturbation we are talking about and specify the type of behaviour that we wish to show is robust under small perturbations.

A minimization problem in the calculus of variations is a special case of a more general class of abstract variational problems:

$$\text{minimize } I(x) \quad \text{over} \quad x \in X, \tag{1.2}$$

where X is a topological space and $I\colon X \to \overline{\mathbb{R}}$ is a given function (recall that $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$). Indeed, well-posedness for this latter, more general, class of problems is a topic of interest in its own right. For simplicity, the problem (1.2) will be abbreviated as **the problem** $\min(X; I)$.

Our minimization problems for variational integrals are more specific and hence have more structure which allows us to use tools and techniques not usually available for the more general version, and usually take the form

$$\text{minimize } J(u) \quad \text{over} \quad u \in S, \tag{1.3}$$

where S is a subset of $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$ and J is the variational integral

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

Of course, in this case the minimization problem is abbreviated as the problem $\min(S; J)$.

The types of perturbations on the minimization problem (1.3) that we will discuss in this work can be categorized as follows.

1. *Linear perturbation of J*: Here the set S in (1.3) is fixed but we perturb the integral J (viewed as a nonlinear functional) by adding a linear functional ξ. For example, we can take $S = \mathrm{W}^{1,p}_0(\Omega; \mathbb{R}^N)$ and the perturbed functional could be
$$(J - \xi)(u) = \int_\Omega F(x, u, \nabla u)\, dx + \int_\Omega h : \nabla u\, dx$$
for some $h \in \mathrm{L}^{p'}(\Omega; \mathbb{R}^{N \times n})$, where p' denotes the conjugate exponent of p, i.e. $\frac{1}{p} + \frac{1}{p'} = 1$.

2. *Boundary value perturbation*: The variational integral J is fixed but we vary the set $S = \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$ in (1.3), where

$$\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N) = \{u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) : u - g \in \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)\}$$

is the Dirichlet class associated with the boundary datum $g \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$, by perturbing the function g.

3. *Singular perturbation*: In order to regularize the problem, a singular perturbation term is added to J. In particular, we are interested in the perturbed variational integral of the form

$$J_\varepsilon = \int_\Omega F(\nabla u)\,dx + \varepsilon \int_\Omega |\nabla^2 u|^2\,dx,$$

where $\varepsilon > 0$ is some fixed value (presumably small). Here $\nabla^2 u$ denotes the Hessian of u. In this case, we want $S = \mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$ where

$$\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N) = \{u \in \mathrm{W}^{2,2}(\Omega;\mathbb{R}^N) : u - g \in \mathrm{W}^{2,2}_0(\Omega;\mathbb{R}^N)\}$$

so that the second integral is finite.

These three types of perturbations are the main contents of chapters 2, 3 and 4, respectively. Of course, there are countless other ways that the problem (1.3) can be perturbed, but only the types of perturbations stated above are the ones that are covered in this thesis.

For perturbed problems in the same class, one can associate with them precise notions of well-posedness. The types of well-posedness considered here can be classified into two main categories, namely *Tykhonov-type well-posedness* and *Hadamard-type well-posedness*.

1. *Tykhonov-type well-posedness*: A given problem $\min(X; I)$ is well-posed in Tykhonov sense if whenever a element x is an approximate minimizer of I, i.e. the value $I(x)$ is close to the infimum of I on X, then x is also a good approximation of the minimizer of I in some specific sense. For example, if the problem $\min(X; I)$ admits a unique minimizer $x_{\min}$, then x could be close to the unique minimizer $x_{\min}$ in the norm topology.

2. *Hadamard-type well-posedness*: For a class of perturbed problems whose perturbations are determined by some parameter, we say that this class is well-posed in Hadamard sense if the solutions of these perturbed problems also vary continuously (in some sense) with this parameter. For example, in the problem $\min(\mathrm{W}^{1,p}_g(\Omega), J)$ we can view g as a parameter and suppose that u_j is the unique minimizer of J in $\mathrm{W}^{1,p}_{g_j}(\Omega)$. If $g_j \to g_0$ implies that $u_j \to u_0$, then we say that this class of problems is well-posed.

The descriptions for these two categories of well-posedness given above are deliberately kept vague. There are several different ways we can interpret the words 'close', 'approximate' and 'continuous', leading to different notions of well-posedness. The need for such diversity partly stems from the fact that the set of all minimizers of a given problem need not be a singleton, it could be many-valued or even an empty set.

For these reasons, it is natural to introduce the notion of *multivalued function* (also known as *set-valued functions* or *multifunctions*) and *convergence of sets* in order to properly formulate and study well-posedness.

Definition 1.2 (Multivalued Function)**.** *Given two sets X and Y, a* multivalued function $T\colon X \rightrightarrows Y$ *is a function (in the usual sense) $T\colon X \to 2^Y$ whose range is contained in the power set of Y. We allow T to take $\emptyset$ as its value and call the set of points $x \in X$ such that $T(x)$ is nonempty the* effective domain *of T, denoted by*

$$\operatorname{Dom} T := \{x \in X : T(x) \neq \emptyset\}$$

whereas its range *is the set* $\operatorname{Range} T = \bigcup_{x\in X} T(x)$.

- *A multivalued function is* proper *if* $\operatorname{Dom} T \neq \emptyset$.
- *The* graph *of T is the subset of $X \times Y$ defined similarly to the graph of a usual function*

  $$\operatorname{Graph} T := \{(x, y) \in X \times Y : y \in T(x)\}.$$

- *The* inverse *of T is another multivalued function $T^{-1}\colon Y \rightrightarrows X$ defined as*

  $$T^{-1}(y) := \{x \in X : y \in T(x)\}.$$

  It is easy to see that $\operatorname{Dom} T^{-1} = \operatorname{Range} T$ *and vice-versa.*

- *When X and Y are topological spaces (or linear spaces) a multivalued function $T\colon X \rightrightarrows Y$ is* compact-valued *(or* closed, open, bounded, convex*, resp.) if $T(x)$ is compact (closed, open, bounded, convex, resp.) for every $x \in X$. T is* locally compact *at x if there exists $U \in \mathcal{N}(x)$ such that $T(U)$ is compact ($\mathcal{N}(x)$ denotes the family of all neighbourhoods of x). We define locally bounded similarly.* //

There are many important multivalued functions that arise naturally in the study of the problem $\min(X; I)$.

Example 1.3.

1. For a given problem $\min(X; I)$, the set of all of its solutions (i.e. minimizers) are of uttermost importance. The solution set

$$\operatorname{argmin}(X; I) := \{x \in X : I(x) = \inf_X I\} \tag{1.4}$$

of the problem $\min(X; I)$ could be an empty set, a singleton, or a set containing many points (even an infinite set in many instances). This gives rise to many important multivalued functions, e.g. the *solution map*

$$\mathcal{S}(g) := \operatorname{argmin}(\mathrm{W}^{1,p}_g(\Omega); J)$$

where J is of the form (1.1), or the map

$$\mathcal{T}(\xi) := \operatorname{argmin}(X; I - \xi).$$

These two multivalued functions play vital roles in the study of how the minimizers vary as we perturb the boundary data or add a linear perturbation to our functional.

2. In addition to the behaviour of the exact solutions of the problem $\min(X; I)$, the behaviour of the *approximate solutions* is just as important. The set of approximate solutions of $\min(X; I)$ can be defined rigorously for $\varepsilon \geq 0$ as

$$\varepsilon\text{-argmin}(X; I) := \{x \in X : I(x) \leq \varepsilon + \inf_X I\}. \tag{1.5}$$

In particular, if $\varepsilon = 0$, then $\operatorname{argmin}(X; I) = 0\text{-argmin}(X; I)$.

Besides the obvious multivalued functions $\mathcal{S}_\varepsilon(g)$ and $\mathcal{T}_\varepsilon(\xi)$ that would result from replacing argmin with ε-argmin in the above definitions, the multivalued function $T\colon [0, \infty) \rightrightarrows X$ defined by

$$T(\varepsilon) := \varepsilon\text{-argmin}(X; I) \tag{1.6}$$

also plays an important role.

3. In chapter 2, we will be using tools from convex analysis to help us in proving well-posedness. One of the indispensable tools is the so-called *subdifferential*, which is defined for a function $f\colon X \to \overline{\mathbb{R}}$ as

$$\partial f(x) = \{x^* \in X^* : \langle x^*, y - x\rangle \leq f(y) - f(x) \text{ for all } y \in X\}.$$

We can see that ∂f is a multivalued function $\partial f\colon X \rightrightarrows X^*$.

Now that we have seen examples for multivalued functions that are relevant to the study of well-posedness, it is time to address what it means for a multivalued function to be *continuous*.

Suppose that X and Y are both topological spaces, the concept of continuity of a single-valued function $f\colon X \to Y$ can be generalized to a multivalued function $T\colon X \rightrightarrows Y$ as follows. Recall that $\mathcal{N}(x)$ denotes the family of all neighbourhoods of x in X.

Definition 1.4 (USC and LSC)**.** *Let X and Y be topological spaces, $x_0 \in X$ and $T\colon X \rightrightarrows Y$ be a multivalued function. We say that T is*

1. USC (Upper semicontinuous) at x_0 *if for any $U \in \mathcal{N}(T(x_0))$, there exists $V \in \mathcal{N}(x_0)$ such that $T(V) \subset U$.*

2. LSC (Lower semicontinuous) at x_0 *if for any $y \in T(x_0)$ and $U \in \mathcal{N}(y)$, there exists $V \in \mathcal{N}(x_0)$ such that $T(x) \cap U \neq \emptyset$ for every $x \in V$.*

T is said to be a USC (resp. LSC) multivalued function if it is USC (resp. LSC) at all points of its domain X. //

The reason for using the upper case USC and LSC to denote the Upper and Lower semicontinuity, respectively, of a multifunction $T\colon X \rightrightarrows Y$ is to distinguish them from the abbreviations usc and lsc reserved for the usual upper and lower semicontinuity of a normal single-valued function $f\colon X \to Y$.

It may not be obvious from the definition why Upper semicontinuity and Lower semicontinuity of a multivalued function are named this way since they do not generalize upper semicontinuity and lower semicontinuity of a single-valued function. In fact, if $T\colon X \rightrightarrows Y$ is single-valued (i.e. $T(x)$ is a singleton for all $x \in X$) Upper and Lower semicontinuity both reduce to the usual continuity. The reason for them being called Upper and Lower semicontinuity is related to another characterization using upper and lower inverse of a multivalued function. For any subset A of Y, we define the *upper inverse* of T by

$$T^u(A) := \{x \in X : T(x) \subset A\}.$$

Similarly, the *lower inverse* of T is defined by

$$T^l(A) := \{x \in X : T(x) \cap A \neq \emptyset\}.$$

Note that the lower inverse T^l essentially coincides with the inverse T^{-1} defined in Definition 1.2. Then

$$\begin{aligned} T \text{ is USC} &\iff T^u(U) \text{ is open for every open set } U \text{ in } Y, \text{ and} \\ T \text{ is LSC} &\iff T^l(U) \text{ is open for every open set } U \text{ in } Y. \end{aligned}$$

There are classes of USC multivalued functions that are of sufficient importance to be given special abbreviations since they appear naturally in many situations.

Definition 1.5 (USCO and CUSCO mappings)**.** *Let (X, τ) and (Y, τ') be topological spaces and $T\colon X \rightrightarrows Y$ be a multivalued function.*

- *T is called a τ-τ'* USCO mapping *if it is τ-τ' USC (at all points $x \in X$) and is τ'-*compact-valued*.*
- *Suppose that Y is a vector space, we say that T is a τ-τ'* CUSCO mapping *if it is both τ-τ' USCO and* convex-valued*.*
- *An USCO mapping (resp. CUSCO mapping) is called a* minimal USCO mapping *(resp.* minimal CUSCO mapping*) if its graph does not strictly contain a graph of any other USCO (resp. CUSCO) mapping.*

The prefix τ-τ' that specifies the topologies on X and Y is often dropped if there is no ambiguity regarding which topologies we are using. //

Here are some examples to illustrate each continuity notion of a multivalued function.

Example 1.6. Let $X = Y = \mathbb{R}$ be endowed with their usual topologies. Suppose that the multivalued functions $T_i\colon \mathbb{R} \rightrightarrows \mathbb{R}$ are defined as follows:

1. $$T_1(x) = \begin{cases} \{1\} & ; x < 0 \\ \{-1, 0, 1\} & ; x = 0 \\ \{0\} & ; x > 0. \end{cases}$$
   Then T_1 is USC but not LSC. It is also USCO but not CUSCO.

2. $$T_2(x) = \begin{cases} \{-1, 1\} & ; x < 0 \\ \{0, 1\} & ; x = 0 \\ \{0\} & ; x > 0. \end{cases}$$
   Then T_2 is LSC but not USC.

3. $$T_3(x) = \begin{cases} [-1, 1] & ; x \leq 0 \\ \{1\} & ; x > 0. \end{cases}$$
   Then T_3 is CUSCO but it is not minimal CUSCO since its graph strictly contains the graph of
   $$T_4(x) = \begin{cases} \{-1\} & ; x < 0 \\ [-1, 1] & ; x = 0 \\ \{1\} & ; x > 0, \end{cases}$$

which is a CUSCO map. In fact, T_4 is a minimal CUSCO map but is still not minimal USCO since its graph is strictly larger than the graph of

$$T_5(x) = \begin{cases} \{-1\} & ; x < 0 \\ \{-1, 1\} & ; x = 0 \\ \{1\} & ; x > 0, \end{cases}$$

which is (minimal) USCO.

It is worth noting that T_4 is actually the subdifferential of the convex function $f(x) = |x|$ on $\mathbb{R}$, i.e. $T_4(x) = \partial f(x)$. This is the reason why it enjoys so many nice properties.

With the precise notions of (semi)continuities for a multivalued functions at our disposal, it is now possible to give exact definitions for the types of well-posedness that are relevant to this thesis.

The first kind of well-posedness we will consider is the Tykhonov-type well-posedness, which ensures that the approximate solutions to the problem $\min(X; I)$ are close to the exact solutions of $\min(X; I)$. This can be made rigorous by requiring that any minimizing sequence converges to the unique minimizer of $\min(X; I)$.

Definition 1.7 (Tykhonov Well-posedness)**.** *Let X be a topological space and $I \colon X \to \overline{\mathbb{R}}$. The problem* $\min(X; I)$ (1.2) *is called* Tykhonov well-posed *if there exists a unique minimizer $x_{\min} \in X$ of I, i.e.*

$$I(x_{\min}) = \inf_X I \quad \textit{and} \quad I(x_{\min}) < I(x) \quad \textit{for all} \quad x \in X \backslash \{x_{\min}\}$$

and any minimizing sequence (x_j) of $\min(X; I)$*, i.e. a sequence such that*

$$I(x_j) \to \inf_X I,$$

we have $x_j \to x_{\min}$ as $j \to \infty$. Such a minimizer is called the strong minimizer *of* $\min(X; I)$*.* //

Remark 1.8. Tykhonov well-posedness of the problem $\min(X; I)$ can be characterized by the conditions

- The multivalued function $T(\varepsilon) = \varepsilon\text{-argmin}\,(X; I)$ (defined at (1.6)) is USC at $\varepsilon = 0$, and
- $T(0) = \operatorname{argmin}(X; I)$ is a singleton.

The proof of this is easy (see [35, Proposition I.22] for the proof). //

Another Tykhonov-type well-posedness is called *Tykhonov well-posedness in the generalized sense*, which can be viewed as Tykhonov well-posedness that allows the non-uniqueness of argmin $(X; I)$.

Definition 1.9 (Generalized Tykhonov well-posedness)**.** *Let X be a topological space and $I\colon X \to \overline{\mathbb{R}}$ be a given function. The problem* $\min(X; I)$ *is called* Tykhonov well-posed in the generalized sense *if* argmin $(X; I)$ *is nonempty and for every minimizing sequence (x_j) in X, there exists a subsequence (x_{j_k}) of (x_j) and $x_{\min} \in$* argmin $(X; I)$ *such that $x_{j_k} \to x_{\min}$ as $k \to \infty$.* //

Remark 1.10. Generalized Tykhonov well-posedness of the problem $\min(X; I)$ can be characterized by conditions similar to the above if we assume further that X is a *metric space* and that I is *lsc* and *bounded below*.

- The multivalued function $T(\varepsilon) = \varepsilon$-argmin $(X; I)$ is USC at $\varepsilon = 0$, and
- $T(0) =$ argmin $(X; I)$ is compact.

See [35, Proposition I.36] for the proof. //

Here are some examples of minimization problems that are well-posed in a (generalized) Tykhonov sense.

Example 1.11. Let us consider a minimization problem $\min(X; I)$ of the form (1.2).

1. Suppose X is a Banach space and $I(x) = \|x\|$. It is easy to see that the problem $\min(X; \|\cdot\|)$ is Tykhonov well-posed and $x = 0$ is the strong minimizer of this problem.

2. Similarly, suppose X is a Banach space, C is a compact subset of X and $I(x) = \operatorname{dist}(x, C)$ (which is a continuous function). We can verify that $T(\varepsilon) = \varepsilon$-argmin $(X; I)$ is USC at $\varepsilon = 0$, hence by Remark 1.10 it follows that $\min(X; \operatorname{dist}(\cdot, C))$ is Tykhonov well-posed in the generalized sense.

3. Looking a little bit ahead, we let X be a topological space and suppose that I satisfies the conditions in Proposition 1.17 that are needed to apply the direct method. Then it is not hard to see that the problem $\min(X; I)$ is Tykhonov well-posed in the generalized sense.

   In particular, this applies to the problem $\min(\mathrm{W}^{1,p}(\Omega); J)$ provided that $\mathrm{W}^{1,p}(\Omega)$ is endowed with the weak topology and the integrand of the functional J satisfies the bounds in Proposition 1.32 below.

The second type of well-posedness is the Hadamard well-posedness, which is about the stability of $\min(X; I)$ under a small perturbation of the function I. The main idea of this type of well-posedness dates back to Hadamard when he studied how solutions to problems that come from mathematical physics depend continuously on the initial data. For our purpose, we will only consider two closely related kinds of Hadamard well-posedness.

Definition 1.12 (Hadamard well-posedness)**.** *Let X be a topological space and $I\colon X \to \overline{\mathbb{R}}$ be an extended real-valued function belonging to a given function space $\mathcal{F}$, where $\mathcal{F}$ is endowed with some topology.*

- *The problem* $\min(X; I)$ *is called* solution Hadamard well-posed *(with respect to $\mathcal{F}$) if there exists a neighbourhood U of I in $\mathcal{F}$ such that* $\operatorname{argmin}(X; J)$ *is a singleton for all $J \in U$ and the mapping $J \mapsto \operatorname{argmin}(X; J)$ is continuous at I.*
- *Similarly, the problem* $\min(X; I)$ *is called* value Hadamard well-posed *(with respect to the function space $\mathcal{F}$) if the mapping $J \mapsto \inf_X J$ is continuous at I.* //

We can generalize Hadamard well-posedness in the same manner as how Tykhonov well-posedness was generalized, i.e. by allowing $\operatorname{argmin}(X; I)$ to be multi-valued and replace the convergence of points in X with a suitable notion of convergence of subsets. Since in later section we will deal with a particular type of $\mathcal{F}$ where each function in it is parametrized by an element of a Banach space, the definition of generalized Hadamard well-posedness below will be given only for this special case.

Definition 1.13 (generalized Hadamard well-posedness)**.** *Let X be a topological space and suppose that we have a family of parametrized functions of the form $I_p\colon X \to \overline{\mathbb{R}}$, where the parameter p comes from a subset K of some Banach space Y.*

We say that the problem $\min(X; I_p)$ *is* solution Hadamard well-posed in the generalized sense *at p_0 (with respect to K) if the multivalued mapping $p \mapsto \operatorname{argmin}(X; I_p)$ is USC at p_0.* //

If $\operatorname{argmin}(X; I_p)$ is single-valued in a neighbourhood of p_0 in K, then solution Hadamard well-posedness in the generalized sense reduces to the previous definition. One can see that value Hadamard well-posedness follows easily from solution Hadamard well-posedness.

Example 1.14. Let us consider the class of problems $\min(X; I_p)$, where the parameter p comes from a Banach space Y.

1. Let $X = Y = \mathbb{R}^n$ be equipped with its usual topology and let $I_p(x) = |x - p|$. Then it is easy to see that the problem $\min(X; I_p)$ is solution Hadamard well-posed for all $p_0 \in \mathbb{R}^n$ (here $\operatorname{argmin}(X; I_p) = \{p\}$).

2. Let Ω be the unit ball in $\mathbb{R}^n$, $X = \mathrm{W}_0^{1,2}(\Omega)$ and $Y = \mathrm{L}^2(\Omega)$. Consider the functional
$$I_g(v) = \int_\Omega \frac{1}{2}|\nabla v(x)|^2 + v(x)g(x)\,dx,$$
where $g \in \mathrm{L}^2(\Omega)$. It is not hard to verify that the direct method applies in this case (here the integrand is convex) so this problem has a minimizer. We have seen from Example 1.1, 3) that I_g is differentiable and that its minimizer (weakly) satisfies the Poisson's equation $\Delta u = g$, so the standard elliptic estimate gives
$$\|u\|_{\mathrm{W}^{1,2}} \leq C\,\|g\|_{\mathrm{L}^2}$$
which, together with linearity, implies that the problem $\min(X; I_g)$ is solution Hadamard well-posed at all $g \in \mathrm{L}^2(\Omega)$.

3. Let $X = Y = \mathbb{R}^n$ (with its usual topology) and let $I_p(x) = \operatorname{dist}(x, B(p; 1))$, where $B(p; 1)$ is the open ball in $\mathbb{R}^n$ centred at p with radius 1. Then the problem $\min(X; I_p)$ is solution Hadamard well-posed for all $p_0 \in \mathbb{R}^n$ (here $\operatorname{argmin}(X; I_p) = \{x \in \mathbb{R}^n : |x - p| \leq 1\}$).

There is another aspect of well-posedness called *genericity*, which is to say that well-posedness may not always hold for a certain class of perturbations, but holds *almost always* (in some specific sense).

To motivate the study of generic well-posedness, let us consider the following possible situations that we may encounter when a perturbation is introduced to a minimization problem. As before, we let X be a topological space, $\{I_p : p \in K\}$ be a family of functions parametrized by $p \in K$ (K is a subset of a Banach space) and $\min(X; I_0)$ be the original minimization problem. Consider the following three situations:

1. We can solve (X, I_0) and want to understand the properties and solvability of (X, I_p) for p near 0.

2. We cannot solve (X, I_0) but we know how to solve (X, I_p) for some small p.

3. We are interested in the global properties of (X, I_p) such as being able to describe the subset of K in which the problem (X, I_p) is well-posed.

In situation (1.), Lower semicontinuity of the solution map $\mathcal{S}(p) = \operatorname{argmin}(X; I_p)$ can help in the study of the problem. Situation (2.) is closely related to the so-called *variational principles*, where we perturb the function I_0 slightly to get a much

better problem (X, I_p) with nice properties, and the solution of (X, I_p) usually says something about the original problem (X, I_0). It is apparent that by accomplishing the task in (3.), we also gain knowledge of situations (1.) and (2.). However, such complete characterization is often very challenging or even intractable, hence the next best thing is to show that there is a large subset (in some specific sense) $D \subset K$ such that our class of problems $\min(X; I_p)$ is well-posed for all $p \in D$.

Definition 1.15 (Generic Well-posedness)**.** *Let X be a topological space and $\{I_p : p \in K\}$ be a family of functions $I_p \colon X \to \overline{\mathbb{R}}$ parametrized by $p \in K$, where K is a subset of a Banach space Y.*

- *The class of problems* $\min(X; I_p)$ *is said to be* generically Tykhonov well-posed in K *if there exists a dense G_δ subset of K such that* $\min(X; I_p)$ *is Tykhonov well-posed for all p in this subset.*
- *Likewise, the class of problems* $\min(X; I_p)$ *is said to be* generically Hadamard well-posed in K *if there exists a dense G_δ subset of K such that* $\min(X; I_p)$ *is Hadamard well-posed with respect to the class $\mathcal{F} = \{I_p : p \in K\}$.* //

Suppose that $\min(X; I_p)$ is generically well-posed in K, where $D \subset K$ is a dense G_δ, then $K \backslash D$ is a meagre set (in the sense of Baire) provided that K is open (or closed) in Y. It is possible to say more about the smallness of $K \backslash D$ in some cases that we will see later on. In chapter 2, we will explore some important results in convex analysis that are used to prove generic Hadamard and Tykhonov well-posedness of the problem $\min(\mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^N); J - \xi)$, where J is a fixed nonlinear integral functional of the form (1.1) and ξ denotes a linear functional taken from the dual space $\mathrm{W}_0^{1,p}(\Omega; \mathbb{R}^N)^*$.

1.2 The Direct Method in the Calculus of Variations and Quasiconvexity

The properties of a variational integral of the form $J(u) = \int_\Omega F(x, u, \nabla u)\, dx$, including the existence of minimizers of J, which is of uttermost importance in the study of well-posedness, depend crucially on its integrand F. In this section we discuss the issues related to the direct method in calculus of variations, which is a technique for obtaining a minimizer of a variational problem, and other closely related topics. In particular, the relationship between quasiconvexity and the direct method is considered in detail.

While this thesis focuses on perturbed variational problems in various forms, understanding the methods for solving the unperturbed problems is still indispensable. In the calculus of variations, one of the most important methods for finding

a minimizer of a variational integral is called the *direct method*. This method can be stated in a very general setting as an existence theorem of a minimizer to an extended real-valued function f on a topological space (X, τ). Since we even allow our function to take $+\infty$ as its values, the following definition rules out a degenerated case.

Definition 1.16. *A function $f\colon X \to \overline{\mathbb{R}}$ is called **proper** if f is not identically $+\infty$ on X.* //

The general statement of the direct method is as follows.

Proposition 1.17 (Direct Method)**.** *Let (X, τ) be a topological space and $I\colon X \to \overline{\mathbb{R}}$ be a proper function on X. Then the function I has a minimizer if the following conditions are satisfied:*

1. ***coercivity**: For each $c \in \mathbb{R}$, the sublevel set*

$$S_c := \{x \in X : I(x) \leq c\}$$

is sequentially precompact, i.e. for any sequence (x_j) in S_c, (x_j) has a subsequence that τ-converges in X.

2. ***sequential τ-lower semicontinuity**: For each sequence $x_j \to x$, we have*

$$I(x) \leq \liminf_{j\to\infty} I(x_j).$$

This result is sometimes called the Weierstrass' theorem even though the property of 'semi-continuity' introduced by Baire was probably not available at the time.

Proof. Let (x_j) be a minimizing sequence of I (by properness, we may assume that $I(x_j) < \infty$ for all j), then $\{I(x_j) : j \in \mathbb{N}\}$ is bounded above by some $M > 0$ and hence $(x_j) \subset S_M$. The coercivity assumption implies that we can extract a subsequence (not relabelled) of (x_j) such that $x_j \to x$ for some $x \in X$. The sequential lower semicontinuity of I then implies that

$$I(x) \leq \liminf_{j\to\infty} I(x_j),$$

hence x is a minimizer of the function I. □

The two conditions in Proposition 1.17 are competing with one another in a sense. Intuitively, if we impose a strong topology τ on X, then it is more likely that I is τ-swlsc (sequentially weakly semicontinuous) but less likely that S_c is compact. On the other hand, if the topology τ is weak, then S_c is more likely to be compact while I is less likely to be swlsc. For example, a Banach space (X, τ) equipped with its normed topology has so few compact sets, making the coercivity condition very hard to satisfy. For this reason, we usually equip X (or its subset) with its weak (or weak*) topology in order to apply the Direct method.

Example 1.18. We take $X = \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ with $p > 1$ and equip X with its weak topology. Of course, here our function $J\colon X \to \overline{\mathbb{R}}$ is an integral functional of the form (1.1) with $F(x,u,z)$ as its integrand. Since X is reflexive, the coercivity condition holds provided that S_c is a bounded set for each $c \in \mathbb{R}$ (this follows from the Banach–Alaoglu theorem and Eberlein–Šmulian theorem, which tells us that precompactness is the same as sequential precompactness). The argument extends to the case that X is a closed affine subset such as $\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$ of $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$, where

$$\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N) := \{u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) : u - g \in \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)\}$$

for some g in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$.

To get the desired conclusion that S_c is bounded, we can assume that the integrand $F(x,u,z)$ is *normal* and *p-coercive* i.e.

$$\alpha + \beta|z|^p \leq F(x,v,z) \tag{1.7}$$

for some $\alpha \in \mathbb{R}$ and $\beta > 0$. The weak sequential lower semicontinuity is related to the quasiconvexity of $F(x,u,z)$ in the z variable (to be discussed below).

The proof that the p-coercivity assumption (1.7) does imply coercivity of J (in the sense of Proposition 1.17) can be found in the appendix. The definition of *normal integrands*, together with another important class of integrands called the *Carathéodory integrands*, is as follows.

Definition 1.19 (Normal and Carathéodory integrands)**.** *Let Ω be an open subset in $\mathbb{R}^n$ and $F\colon \Omega\times\mathbb{R}^d \to \overline{\mathbb{R}}$ be a given function. We say that F is a* normal integrand *if the following conditions are satisfied:*

1. *For each $x \in \Omega$, the function $F(x,\cdot)$ is lsc.*
2. *F is $(\mathcal{L}^n \otimes \mathcal{B})$-measurable.*

Here $\mathcal{B}$ denotes the Borel σ-algebra of $\mathbb{R}^d$ whereas $\mathcal{L}^n$ denotes the σ-algebra of all (n-dimensional) Lebesgue measurable subset of Ω. The integrand F is called proper *if $F(x,\cdot)$ is proper for all $x \in \Omega$.*

Suppose that the integrands F and $-F$ are both proper and normal, we say then that F is a Carathéodory integrand*. Equivalently, F is Carathéodory if the following conditions, known as the Carathéodory conditions, are satisfied:*

1. *For each $x \in \Omega$, the function $F(x,\cdot)$ is continuous.*
2. *For each $\xi \in \mathbb{R}^d$, the function $F(\cdot,\xi)$ is $\mathcal{L}^n$-measurable.*

A statement that an integrand $F\colon \Omega\times\mathbb{R}^N\times\mathbb{R}^{N\times n} \to \overline{\mathbb{R}}$ is normal (or Carathéodory) is to be interpreted that we take $\mathbb{R}^d$ in the above to be $\mathbb{R}^d = \mathbb{R}^N \times \mathbb{R}^{N\times n}$. //

Remark 1.20. Suppose that $u\colon \Omega \to \mathbb{R}^N$ is a measurable function; then it follows directly from the definition of the normal integrand (and the fact that $x \mapsto (x, u(x))$ is $\mathcal{L}^n$-$(\mathcal{L}^n \otimes \mathcal{B})$-measurable) that $x \mapsto F(x, u(x))$ is measurable provided that F is normal (or Carathédory). It should be noted that the Carathéodory conditions can be deduced easily from the information that F and $-F$ are normal but the converse is far from obvious and relies on the fact that $\mathcal{L}^n$ is a complete measure. Also, the equivalence between the definition of Carathéodory integrand from normal integrand and the definition using Carathéodory conditions is not trivial. For a detailed discussion about normal and Carathéodory integrands, together with the proofs of those statements, we refer to Rockafellar [88]. //

There are several other conditions beside (1.7) that ensure weak coercivity (in the sense of Proposition 1.17). We refer to [21] and [28] for in-depth treatments of the direct methods in the calculus of variations and [48] for further discussion regarding coercivity of variational integrals. It should be noted that (1.7) is sufficient to guarantee that the integral function J (with F as its integrand) is strongly continuous in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$, i.e. continuous with respect to the norm topology, but not enough to prove that J is weakly continuous. In fact, a much weaker lower bound is sufficient to guarantee strong continuity. The proof of the following can be found in the Appendix A.

Proposition 1.21. *For a functional $J\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \to \overline{\mathbb{R}}$ of the form (1.1), where $p \geq 1$, J is strongly lsc (lower semicontinuous) in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ provided its integrand F is normal and satisfies the lower bound*

$$F(x, u, z) \geq -h(x) - \alpha|u|^t - \beta|z|^p, \tag{1.8}$$

where $\alpha, \beta \geq 0$, $1 \leq t \leq p^*$ and $h \in \mathrm{L}^1(\Omega)$. Here p^* is the exponent from the Sobolev embedding theorem $\mathrm{W}^{1,p} \hookrightarrow \mathrm{L}^{p^*}$, i.e. $p^* = \frac{np}{n-p}$ if $p < n$ and $p^* = \infty$ if $p \geq n$.

Applying the above proposition to F and $-F$, the following corollary is immediate.

Corollary 1.22. *If $J\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \to \overline{\mathbb{R}}$ is a functional of the form (1.1), where $p \geq 1$, then J is continuous in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ provided its integrand F is Carathéodory and satisfies*

$$|F(x, u, z)| \leq h(x) + \alpha|u|^t + \beta|z|^p, \tag{1.9}$$

where $\alpha, \beta \geq 0$, $1 \leq t \leq p^*$ and $g \in \mathrm{L}^1(\Omega)$.

The sequential weak lower semicontinuity of a variational integral J is a strictly stronger condition than being sequentially semicontinuous in the strong topology. Indeed, we shall give a counterexample of a variational integral J whose integrand satisfies the lower bound (1.8) in the space $\mathrm{W}^{1,p}(\Omega;\mathbb{R})$ for any $p \geq 1$, but J is not sequential weak lower semicontinuous.

Example 1.23. Let $\Omega = (0,1)$, $p \in [1,\infty)$ and J be defined for any $u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R})$ by

$$J(u) = \int_0^1 (u'(x)^2 - 1)^2 \, dx,$$

i.e. $F(x,u,v) = (v^2-1)^2$. Clearly F satisfies the lower bound (1.8) since $F \geq 0$, so J is strongly continuous.

Let $f\colon \mathbb{R} \to \mathbb{R}$ be

$$f(x) = \begin{cases} x & ; x \in [0,1/2] \\ 1-x & ; x \in [1/2,1] \end{cases}$$

and extends it periodically to all $x \in \mathbb{R}$ so that $f(x+1) = f(x)$. Let (u_j) be a sequence of functions in $\mathrm{W}^{1,p}(\Omega;\mathbb{R})$ defined by

$$u_j(x) = \frac{1}{n} f(nx).$$

Observe that $u_j'(x) = \pm 1$ a.e., hence $J(u_j) = 0$ for every $j \in \mathbb{N}$. We can show that $u_j \overset{w}{\rightharpoonup} 0$ in $\mathrm{W}^{1,p}(\Omega;\mathbb{R})$ but

$$J(0) = 1 > 0 = \liminf_{n\to\infty} J(u_j).$$

This shows that J is not sequentially weakly lower semicontinuous.

Sequential weak lower semicontinuity is directly related to *convexity* in the last variable of $F(x,u,z)$. Indeed, a quick look at the integrand $F(z) = (z^2-1)^2$ in the previous example reveals that $F(z)$ is not convex in z.

Recall that n is the dimension of the domain $\Omega \subset \mathbb{R}^n$ and N is the dimension of the target space of $u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$. In the *scalar case* $\min\{n,N\} = 1$, convexity is both necessary and sufficient for sequential weak lower semicontinuity. More generally, convexity in the last variable of $F(x,u,z)$ always implies sequential weak lower semicontinuity of the corresponding variational integral J on $\mathrm{W}^{1,p}$. The following well-known results give the precise statement of this fact (the proof of it can be found in [28, Theorem 1.3], for example)

Proposition 1.24. *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with Lipschitz boundary. Let $F\colon \Omega \times \mathbb{R}^N \times \mathbb{R}^{N\times n} \to \mathbb{R}$ be a non-negative continuous integrand of a variational integral J of the form (1.1). Then*

1. *If the function $F(x,u,\cdot)$ is convex for all $x,u \in \Omega \times \mathbb{R}^N$, then J is swlsc in $\mathrm{W}^{1,p}$.*

2. *If J is swlsc in $\mathrm{W}^{1,p}$ and either $n = 1$ or $N = 1$, then its integrand $F(x,u,\cdot)$ is convex for all $x,u \in \Omega \times \mathbb{R}^N$.*

The assumptions on F can be significantly relaxed, e.g. F need only be Carathéodory instead of continuous, and still yield essentially the same result (see [28, Corollary 3.24]).

In the vectorial case where $n > 1$ and $N > 1$, convexity in the last variable of $F(x, u, z)$ is no longer a necessary condition for the variational integral J to be sequentially weakly lower semicontinuous (but it is still a sufficient condition). Indeed, many important classes of integrands that fail to satisfy this assumption still give rise to a swlsc functional. It turns out that the 'correct' condition for the integrand $F(x, u, \cdot)$ in order that J be swlsc in $\mathrm{W}^{1,p}$ is a weaker form of convexity called *quasiconvexity*. Recall that for a convex function $f\colon \mathbb{R}^{N\times n} \to \mathbb{R}$, Jensen's inequality says that

$$F(z) \leq \int_\Omega F(z + h(\omega))\, d\mu(\omega)$$

for any integrable function $h\colon \Omega \to \mathbb{R}^{N\times n}$ and any probability measure μ on Ω such that $\int_\Omega h(\omega)\, d\mu(\omega) = 0$. A quasiconvex function can be viewed as a function that satisfies a restricted form of Jensen's inequality.

Definition 1.25 (quasiconvexity)**.** *Let $F\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ be a locally bounded Borel-measurable function. F is called* quasiconvex *at $z \in \mathbb{R}^{N\times n}$ if*

$$F(z) \leq \frac{1}{|B(0,1)|} \int_{B(0,1)} F(z + \nabla\varphi(x))\, dx \tag{1.10}$$

holds for all $\varphi \in \mathrm{W}^{1,\infty}_0(B(0,1); \mathbb{R}^N)$. Here $B(0,1)$ is the open unit ball in $\mathbb{R}^n$.

F is said to be a quasiconvex integrand if it is quasiconvex at every point $z \in \mathbb{R}^{N\times n}$. //

Clearly, if F is convex, then Jensen inequality immediately implies that it is quasiconvex since $\int_{B(0,1)} \nabla\varphi(x)\, dx = 0$ for any $\varphi \in \mathrm{W}^{1,\infty}_0(B(0,1); \mathbb{R}^N)$.

Remark 1.26.

1. The domain of integration $B(0,1)$ in the definition of quasiconvexity can be replaced by any bounded open domain to get an equivalent definition of quasiconvexity (see [48]).

2. Under the *p-growth* assumption, i.e.

   $$|F(z)| \leq c(1 + |z|^p)$$

   for some $p \in [1, \infty)$ and $c > 0$, the inequality (1.10) holds for any $\varphi \in \mathrm{W}^{1,p}_0(\Omega; \mathbb{R}^N)$.

3. Quasiconvexity coincides with the usual convexity if $N = 1$ or $n = 1$. //

The definition of quasiconvexity was introduced by Morrey who showed in [79] that *quasiconvexity* is essentially equivalent to sequential weak lower semicontinuity of the integral functional J of the form (1.1), modulo some technical assumptions. Nowadays, quasiconvexity is recognized as a central convexity condition of the calculus of variations. The following sequential weak lower semicontinuity result for a functional J is due to Acerbi and Fusco ([1, Theorem II.4]).

Theorem 1.27 (Acerbi–Fusco)**.** *Let $p \in [1, \infty)$, Ω be a Lipschitz domain, J be an integral functional of the form (1.1) and $F\colon \Omega \times \mathbb{R}^N \times \mathbb{R}^{N\times n} \to \mathbb{R}_+$ be a non-negative Carathéodory integrand with p-*growth*, i.e.*

$$F(x, u, z) \leq c(1 + |u|^p + |z|^p). \tag{1.11}$$

Then the functional J is swlsc on $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$ if and only if $F(x, u, \cdot)$ is quasiconvex for almost every $x \in \Omega$ and every $u \in \mathbb{R}^N$.

An alternative proof of this theorem using Young measures can also be found in [87]. This theorem can be modified in many ways, such as replacing the non-negativity condition in Proposition 1.27 by a much more lenient lower bound as in [28, Theorem 8.11] or [48, Theorem 5.7], which is

$$F(x, u, z) \geq -h(x) - \alpha|u|^t - \beta|z|^r \tag{1.12}$$

for some $\alpha, \beta \geq 0$, $1 \leq t < p^*$, $1 \leq r < p$ and some $h \in \mathrm{L}^1(\Omega)$.

Remark 1.28. It is worth noting that sequential weak lower semicontinuity of $f\colon X \to Y$ is generally a strictly weaker property than weak lower semicontinuity (since the weak topology on X is not completely determined by sequences). Consider, for example, the operator $T\colon l^2 \to l^2$ defined by

$$T(\alpha) = \left(\alpha^1, \frac{\alpha^2}{\sqrt{2}}, \frac{\alpha^3}{\sqrt{3}}, \dots\right)$$

for any $\alpha = (\alpha^1, \alpha^2, \alpha^3, \dots) \in l^2$ (each component of α is indexed by a superscript). This operator is sequentially w-$\|\cdot\|$ continuous but not w-$\|\cdot\|$ continuous. Indeed, for any sequence $\alpha_j \overset{w}{\rightharpoonup} 0$ in l^2, there exists $C > 0$ such that $\|\alpha_j\|^2 \leq C$, hence

$$\|T\alpha_j\|^2 = \sum_{i=1}^{\infty} \frac{|\alpha_j^i|^2}{i} \leq \sum_{i=1}^{m} |\alpha_j^i|^2 + \frac{1}{m+1} \sum_{i=1}^{\infty} |\alpha_j^i|^2 \leq \sum_{i=1}^{m} |\alpha_j^i|^2 + \frac{C}{m+1}.$$

Given any $\varepsilon > 0$, we can make $\|T\alpha_j\| \leq \varepsilon$ by choosing sufficiently large m and let j be sufficiently large. This shows that T is sequentially w-$\|\cdot\|$ continuous. However, $0 \in l^2$ is in the weak closure of the set $A = \{\sqrt{j}e_j : j \in \mathbb{N}\} \subset l^2$ but the inverse image under T of an open ball $B_{l^2}(0, 1/2)$ cannot contain an element of A since $\left\|T(\sqrt{j}e_j)\right\| = \|e_j\| = 1$. This shows that T is not w-$\|\cdot\|$ continuous.

Lastly, we compose T with the map $-\|\cdot\| : l^2 \to \mathbb{R}$ to get a map that is sequentially weakly continuous, but not weakly lower semicontinuous (at $0 \in l^2$). //

Despite its importance with regard to the sequential lower semicontinuity of variational integrals, quasiconvexity is an elusive condition that is hard to verify in general. This is largely due to the fact that quasiconvexity is not a local condition (in a sense that was made precise by Kristensen [59]). Nevertheless, there is an important subclass of quasiconvex functions that is much easier to deal with, called the class of *polyconvex functions*. Polyconvexity was introduced by Ball in [10] and had quickly gained recognition as another convexity condition of great significance in the field of the calculus of variations, in particular the modern theory of hyperelastic bodies.

Definition 1.29 (polyconvexity)**.** *For $z \in \mathbb{R}^{N\times n}$, we denote by*

$$\boldsymbol{T}(z) = (T_1(z), \dots, T_{d(N,n)}(z))$$

the collection of all minors (or subdeterminants) of z in some fixed order (unimportant, but fixed). A function $F\colon \mathbb{R}^{N\times n} \to \overline{\mathbb{R}}$ is called polyconvex *if there exists a convex function $f\colon \mathbb{R}^{d(N,n)} \to \overline{\mathbb{R}}$ such that*

$$F(z) = f(\boldsymbol{T}(z))$$

for all $z \in \mathbb{R}^{N\times n}$. //

For example, we have

- $F(z) = f(z, \det z)$ if $N = n = 2$
- $F(z) = f(z, \operatorname{cof} z, \det z)$ if $N = n = 3$

There is yet another important notion of convexity in the calculus of variations, namely *rank-one convexity*. As oppose to polyconvexity, rank-one convexity is weaker than quasiconvexity.

Definition 1.30 (rank-one convexity)**.** *A function $F\colon \mathbb{R}^{N\times n} \to \overline{\mathbb{R}}$ is called* rank-one convex *if*

$$f((1-\lambda)z + \lambda w) \le (1-\lambda)f(z) + \lambda f(w)$$

for any $\lambda \in [0,1]$ and any $z, w \in \mathbb{R}^{N\times n}$ such that $\operatorname{rank}(z-w) \le 1$. //

The relationship between these convexity notions for a real-valued function $f\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ can be summarized as follows:

f convex $\implies$ f polyconvex $\implies$ f quasiconvex $\implies$ f rank-one convex.

In the scalar case, convexity and rank-one convexity coincide and hence all these notions of convexity collapse into the standard convexity. In the vectorial case, however, all these implications are strict provided that N, n are big enough. In some low dimension cases, that is, for $N = 2$ and $n \ge 2$, the question whether rank-one convexity implies quasiconvexity is still open. For more discussion, see [28, Theorem 5.3 and Remark 5.4].

Example 1.31.

1. Let $N = n = 2$. For simplicity, let us consider a functional whose integrand $F(x, u, z)$ is independent of x and u

$$J(u) = \int_\Omega \det(\nabla u)\, dx.$$

Here the integrand $F(z)$ is polyconvex (in fact it is *polyaffine*, meaning that both $\pm F$ are polyconvex) but not convex.

2. Let $N = n = 2$ as before and consider the integral functional

$$J(u) = \int_\Omega f_\gamma(\nabla u(x))\, dx,$$

where

$$f_\gamma(z) = |z|^4 - 2\gamma|z|^2 \det z.$$

Here we write $|z| = \sqrt{\operatorname{tr}(z^t z)}$.

From the results in [4] and [31] by Alibert, Dacorogna and Marcellini, it is known that f_γ is polyconvex if and only if $|\gamma| \leq 1$ whereas f_γ is quasiconvex if and only if $|\gamma| \leq 1+\varepsilon$ for some $\varepsilon > 0$ (see also [52] by Iwaniec and Kristensen). This shows that there exists some $\gamma > 1$ such that f_γ is quasiconvex but not polyconvex.

It is worth noting that f_γ is rank-one convex if and only if $|\gamma| \leq 2/\sqrt{3}$ and from the numerical results in [29] and [30], it is believed (though not proved) that $1 + \varepsilon = 2/\sqrt{3}$.

To end this section, we give an example of the application of the direct method to a minimization problem of a quasiconvex variational integral. Recall from Example 1.18 that the pointwise coercivity condition (1.7) for the integrand F implies the weak coercivity of the corresponding variational integral J in $\mathrm{W}^{1,p}(\Omega)$ (meaning that $S_c = \{u \in \mathrm{W}^{1,p}(\Omega) : J(u) \leq c\}$ is sequentially weakly precompact for all c) provided that $p > 1$, and observe that it also implies the lower bound (1.12). Thus if we assume that $F(x, u, \cdot)$ is quasiconvex and that the p-growth condition (1.11) holds, then we also get sequential weak lower semicontinuity and hence the direct method can be applied.

The above discussion can be summarized as follows.

Proposition 1.32. *Suppose that $p > 1$, Ω is a Lipschitz domain, J is an integral functional of the form (1.1) and $F\colon \Omega \times \mathbb{R}^N \times \mathbb{R}^{N\times n} \to \mathbb{R}$ is a Carathéodory integrand satisfying the p-coercivity assumption (1.7) and the p-growth assumption (1.11), i.e.*

$$-\alpha + \beta|z|^p \leq F(x, u, z) \leq \gamma(1 + |u|^p + |z|^p) \tag{1.13}$$

for some $\alpha, \beta, \gamma > 0$. *Then the minimization problem*

$$\text{minimize } J(u) \quad \text{over} \quad u \in \mathrm{W}^{1,p}_g(\Omega; \mathbb{R}^N)$$

has a solution for any Dirichlet boundary datum $g \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$.

Notice that this proposition is only stated for $p > 1$. For $p = 1$, the p-growth condition is sufficient for sequential weak lower semicontinuity in $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^d)$ but the pointwise coercivity condition (1.7) no longer guarantees that a minimizer exists in $\mathrm{W}^{1,1}(\Omega; \mathbb{R}^N)$. In fact, the space $\mathrm{BV}(\Omega; \mathbb{R}^N)$ of functions with bounded variation is a more suitable context in which to find a minimizer (in a generalized sense) of an integral functional whose integrand satisfies the 1-growth condition. It should be noted that with suitable modifications, much of the previous results that we have discussed can be extended to the case where the variational integral J has a normal integrand instead of a Carathéodory integrand.

While the coercivity conditions and growth conditions that have been introduced so far are all pointwise conditions on the integrand $F(x, u, z)$ of the variational integral J, different types of coercivity and growth conditions are also possible (see Chen and Kristensen [21], for example). The fact that sequential weakly lower semicontinuity holds under the quasiconvexity condition can be seen as an instance of the so-called 'compensated compactness' methods (see Murat and Tartar [100, 101]). For more information regarding these methods and other methods of solving variational problems in the calculus of variations, such as the 'concentration compactness' methods by Lions [68, 69, 70, 71]), Struwe's monograph [93] is a good place to look. Note that none of these other methods are suitable for the type of variational problems that we consider here in this chapter; for this we need the direct method.

1.3 Young Measure and Lower Semicontinuity

Over the recent years, Young measures have become an important tool for dealing with nonlinearity in calculus of variations and many related disciplines. There are several ways Young measures can be introduced, with various degrees of generality. The theory that is presented here is not the most general one possible but is sufficient for our purpose. Most of the material in this section is from the books [87] by Rindler, [45] by Fonseca and Leoni, [82] by Pedregal, and series of lecture notes and the paper [58] by Kristensen.

In the following, $C_0(\mathbb{R}^m)$ is the space of all continuous functions vanishing at infinity. Its dual is the space of all (signed) Radon measures, denoted by $\mathcal{M}(\mathbb{R}^m)$. We denote the subset of $\mathcal{M}(\mathbb{R}^m)$ consisting of all probability measures by $\mathcal{M}^1(\mathbb{R}^m)$. Recall that for a bounded open set $\Omega \subset \mathbb{R}^n$ (equipped with the Lebesgue measure), $\mathrm{L}^1(\Omega; C_0(\mathbb{R}^m))^* \simeq \mathrm{L}^\infty_w(\Omega; \mathcal{M}(\mathbb{R}^m))$ (see Proposition A.24 in the appendix).

Definition 1.33. *[Young measure] Let Ω be a bounded open subset of $\mathbb{R}^n$, equipped with the Lebesgue measure. A Young measure is an element $\nu \in \mathrm{L}^\infty_w(\Omega; \mathcal{M}(\mathbb{R}^m))$ such that $\nu(x) \in \mathcal{M}^1(\mathbb{R}^m)$ almost every $x \in \Omega$.* //

Remark 1.34.

1. For a Young measure ν, we often write $\nu_x := \nu(x)$.
2. The *barycenter* (or the *centre of mass*) of measure $\mu \in \mathcal{M}(\mathbb{R}^m)$ is defined, provided that $\int_{\mathbb{R}^m} |z| \, d\mu(z) < \infty$, by
$$\overline{\mu} := \int_{\mathbb{R}^m} z \, d\mu(z).$$
If ν is a Young measure, $\overline{\nu}$ is a function defined by $\overline{\nu}(x) := \overline{\nu}_x$.
3. If ν is a Young measure, then for any Carathéodory integrand $f \colon \Omega \times \mathbb{R}^m \to \mathbb{R}$, the function (in x variable)
$$f(x, \nu_x) := \int_{\mathbb{R}^m} f(x, z) \, d\nu_x(z)$$
is Lebesgue-measurable. (see [87, Theorem 4.1-4.5 and Problem 4.1]). //

Proposition 1.35 (Fundamental Theorem of Young Measure)**.** *For $p \in [1, \infty]$, let $V_j \in \mathrm{L}^p(\Omega; \mathbb{R}^m)$ be a bounded sequence. Then there is a subsequence (not relabelled) and a Young measure ν such that the following hold:*

1. *If $p < \infty$, then*
$$\langle\!\langle \nu, |\cdot|^p \rangle\!\rangle := \int_\Omega \int_{\mathbb{R}^m} |z|^p \, d\nu_x(z) \, dx < \infty. \tag{1.14}$$
If $p = \infty$, then there exists a compact set $K \subset \mathbb{R}^m$ such that
$$\operatorname{supp} \nu_x \subset K \quad \text{for almost all} \quad x \in \Omega. \tag{1.15}$$
2. *For any Carathéodory integrand $f \colon \Omega \times \mathbb{R}^m \to \mathbb{R}$ such that $(f(x, V_j))_{j \in \mathbb{N}}$ is uniformly integrable, we have*
$$f(x, V_j(x)) \overset{w}{\rightharpoonup} f(x, \nu_x) \text{ in } \mathrm{L}^1. \tag{1.16}$$
In particular, we have
$$\int_\Omega f(x, V_j(x)) \, dx \to \int_\Omega f(x, \nu_x) \, dx = \int_\Omega \int_{\mathbb{R}^m} f(x, z) \, d\nu_x(z) \, dx.$$

Proof. See [87, Theorem 4.1-4.5]) or [45, Theorem 8.6]. □

Definition 1.36. *For $p \in [1, \infty]$,*

- *The set of all Young measures that satisfies condition (1) in Proposition 1.35 is denoted by* $\mathrm{Y}^p(\Omega;\mathbb{R}^m)$.
- *If a sequence of measurable functions* $V_j\colon \Omega \to \mathbb{R}^m$ *satisfies the condition (2) in Proposition 1.35, then we say that the sequence* V_j *generates the Young measure* ν. *Symbolically, we write*

$$V_j \xrightarrow{\mathrm{Y}} \nu.$$

//

Remark 1.37. A sufficient condition for a sequence of measurable functions $V_j\colon \Omega \to \mathbb{R}^m$ to generate a Young measure ν is the *tightness condition*

$$\lim_{t\to\infty} \limsup_{j\to\infty} \mathcal{L}^n(\{|V_j| > t\}) = 0,$$

which is implied by boundedness in the L^p norm. //

Next, we shall state some of the most useful results regarding Young measure that is used in this work.

Proposition 1.38. *For* $p \in (1,\infty]$, *let* V_j *be a sequence in* $\mathrm{L}^p(\Omega;\mathbb{R}^m)$ *such that* $V_j \xrightarrow{\mathrm{Y}} \nu \in \mathrm{Y}^p(\Omega;\mathbb{R}^m)$. *Setting* $V(x) := \overline{\nu}_x$, *then for* $p \in (1,\infty)$ *we have*

$$V_j \overset{w}{\rightharpoonup} V \ \text{ in } \ \mathrm{L}^p,$$

or, for $p = \infty$,

$$V_j \overset{w^*}{\rightharpoonup} V \ \text{ in } \ \mathrm{L}^\infty.$$

Proof. See [87, Lemma 4.11]). □

Proposition 1.39. *For* $p \in [1,\infty]$, *let* V_j *be a bounded sequence in* $\mathrm{L}^p(\Omega;\mathbb{R}^m)$ *such that* $V_j \xrightarrow{\mathrm{Y}} \nu \in \mathrm{Y}^p(\Omega;\mathbb{R}^m)$. *Then for* $V \in \mathrm{L}^p(\Omega;\mathbb{R}^m)$,

$$V_j \to V \text{ in measure} \qquad \text{if and only if} \qquad \nu_x = \delta_{V(x)} \text{ a.e. } x \in \Omega.$$

Proof. See [87, Lemma 4.12]) or [45, Corollary 8.9]. □

The following proposition allows us to 'freeze' one variable in the integrand, which will prove to be useful when we deal with integrands with several variables.

Proposition 1.40. *Let* $p \in (1,\infty)$, u_j *be a bounded sequence in* $\mathrm{L}^p(\Omega;\mathbb{R}^N)$ *and* V_j *be a sequence in* $\mathrm{L}^p(\Omega;\mathbb{R}^M)$. *Suppose that there exist* $u \in \mathrm{L}^p(\Omega;\mathbb{R}^N)$ *and* $\nu \in \mathrm{Y}^p(\Omega;\mathbb{R}^M)$ *such that*

$$u_j \to u \text{ pointwise a.e.} \qquad \text{and} \qquad V_j \xrightarrow{\mathrm{Y}} \nu,$$

then $(u_j, V_j) \xrightarrow{\mathrm{Y}} \mu \in \mathrm{Y}^p(\Omega;\mathbb{R}^{N+M})$ *where*

$$\mu_x = \delta_{u(x)} \otimes \nu_x \text{ a.e. } x \in \Omega.$$

Proof. See [87, Lemma 5.19] or [45, Corollary 8.10]. □

There is a certain subclass of Young measure that is especially interesting to us, namely the class of *gradient Young measures*, whose elements are generated by a sequence of gradients. The reason why it is so important in calculus of variations is because it interacts well with quasiconvexity.

Definition 1.41 (Gradient Young Measure)**.** *Let Ω be a bounded open set in $\mathbb{R}^n$ and $p \in [1,\infty]$. For a Young measure $\nu \in \mathrm{Y}^p(\Omega;\mathbb{R}^{N\times n})$, we say that ν is a* $\mathrm{W}^{1,p}$-gradient Young measure *if there exists a bounded sequence u_j in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ such that*

$$\nabla u_j \xrightarrow{\mathrm{Y}} \nu.$$

- *The set of all $\mathrm{W}^{1,p}$-gradient Young measures is denoted by $\mathrm{GY}^p(\Omega;\mathbb{R}^{N\times n})$.*
- *If $\nu \in \mathrm{GY}^p(\Omega;\mathbb{R}^{N\times n})$ is constant for a.e. $x \in \Omega$, then we say that ν is a homogeneous gradient Young measure. In this case, we sometimes also write ν instead of ν_x (and likewise for $\overline{\nu}$).* //

Remark 1.42. For a gradient Young measure $\nu \in \mathrm{GY}^p(\Omega;\mathbb{R}^{N\times n})$, suppose that ∇u_j is a bounded sequence in $\mathrm{L}^p(\Omega;\mathbb{R}^{N\times n})$ that generates ν. Suppose that Ω is a bounded Lipschitz domain (so that we can apply the Poincaré–Wirtinger inequality), by subtracting $\overline{u_j}_\Omega$ from u_j we can assume that u_j is bounded in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ and that $u_j \overset{w}{\rightharpoonup} u$ in $\mathrm{W}^{1,p}$. From Proposition 1.38, it follows that $\overline{\nu} = \nabla u$ a.e. $x \in \Omega$. Such a u will be called an *underlying deformation* of ν. //

Proposition 1.43. *Let $p \in (1,\infty]$ and $\nu \in \mathrm{GY}^p(B_1(0);\mathbb{R}^{N\times n})$ be a gradient Young measure (where $B_1(0)$ is the unit ball in $\mathbb{R}^n$). Suppose that $h\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ is a quasiconvex function with p-growth, i.e. $|h(z)| \le C(1+|z|^p)$ for some $C \ge 0$, then*

$$h(\overline{\nu}) \le \langle \nu, h\rangle = \int_{\mathbb{R}^{N\times n}} h(z)\,d\nu(z).$$

For $p = \infty$, there is no restriction on the growth condition.

Proof. See [87, Lemma 5.11]). □

Proposition 1.44. *Let $p \in [1,\infty)$ and $\nu \in \mathrm{GY}^p(\Omega;\mathbb{R}^{N\times n})$. Then for almost every $x_0 \in \Omega$, the probability measure ν_{x_0} is a homogeneous gradient Young measure in its own right, i.e. $\nu_{x_0} \in \mathrm{GY}^p(B_1(0);\mathbb{R}^{N\times n})$.*

Proof. See [87, Lemma 5.14]). □

Proposition 1.45. *Let $p \in (1,\infty)$ and $\nu \in \mathrm{GY}^p(\Omega;\mathbb{R}^{N\times n})$. Suppose that $F\colon \Omega \times \mathbb{R}^N \times \mathbb{R}^{N\times n} \to \mathbb{R}$ is quasiconvex in the last variable with p-growth in the sense that*

$$|h(x,u,z)| \le C_{x,u}(1+|z|^p) \qquad \text{for a.e. } x \in \Omega$$

for some constant $C_{x,u} > 0$. Then we have

$$h(x,u,\overline{\nu}_x) \le h(x,u,\nu_x) = \int_{\mathbb{R}^{N\times n}} h(x,u,z)\,d\nu(z) \qquad \text{for a.e. } x \in \Omega.$$

Proof. This clearly follows from Proposition 1.44 and Proposition 1.43. □

Recall that $F^+ = \max\{F, 0\}$ and $F^- = \max\{-F, 0\}$ are the positive and negative parts of F, respectively. With all of the above, we have the following lower semicontinuity result.

Proposition 1.46. *Let $p \in (1, \infty)$, $\mathcal{F}\colon \Omega \times \mathbb{R}^N \times \mathbb{R}^{N\times n}$ be a Carathéodory integrand such that*

$$|F(x, u, z)| \leq C_{x,u}(1 + |z|^p) \qquad \text{for a.e. } x \text{ and } u$$

and that $F(x, u, \cdot)$ is quasiconvex for almost all x and u. Suppose $(u_j)_{j\in\mathbb{N}}$ be a sequence in $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$ such that $u_j \overset{w}{\rightharpoonup} u$ in $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$ and $(F^-(x, u_j, \nabla u_j))_{j\in\mathbb{N}}$ is uniformly integrable. Then we have

$$\int_\Omega F(x, u, \nabla u)\, dx \leq \liminf_{i\to\infty} \int_\Omega F(x, u_j, \nabla u_j)\, dx.$$

Proof. Without loss of generality, it is enough to prove the inequality for any subsequence of u_j; hence we can assume that $\nabla u_j \overset{Y}{\to} \nu$ for some $\nu \in \mathrm{GY}^p(\Omega; \mathbb{R}^{N\times n})$ and that $u_j \to u$ a.e. $x \in \Omega$. By Proposition 1.40, we get that $(u_j, \nabla u_j) \overset{Y}{\to} \mu$, where $\mu_x = \delta_{u(x)} \otimes \nu_x$ for a.e. $x \in \Omega$.

Since $(F^-(x, u_j, \nabla u_j))_{j\in\mathbb{N}}$ is uniformly integrable, Proposition 1.35 implies that

$$\int_\Omega F^-(x, u(x), \nu_x)\, dx = \lim_{j\to\infty} \int_\Omega F^-(x, u_j(x), \nabla u_j(x))\, dx.$$

On the other hand, we can define $F^M(x, u, z) := \max\{F^+(x, u, z), M\}$ and observe that $(F^M(x, u_j, \nabla u_j))_{j\in\mathbb{N}}$ is (trivially) uniformly integrable. Therefore we also have

$$\begin{aligned}\int_\Omega F^M(x, u(x), \nu_x)\, dx &= \lim_{j\to\infty} \int_\Omega F^M(x, u_j(x), \nabla u_j(x))\, dx \\ &\leq \liminf_{j\to\infty} \int_\Omega F^+(x, u_j(x), \nabla u_j(x))\, dx\end{aligned}$$

for any $M > 0$. By the monotone convergence theorem, we can take the limit as $M \to \infty$ and conclude that

$$\int_\Omega F^+(x, u(x), \nu_x)\, dx \leq \liminf_{j\to\infty} \int_\Omega F^+(x, u_j(x), \nabla u_j(x))\, dx$$

Using the fact that $F = F^+ - F^-$ and that $\liminf_j(a_j - b_j) = \liminf_j a_j - \lim_j b_j$, we have

$$\int_\Omega F(x, u(x), \nu_x)\, dx \leq \liminf_{j\to\infty} \int_\Omega F(x, u_j(x), \nabla u_j(x))\, dx.$$

Lastly, since $\nabla u_j \overset{w}{\rightharpoonup} \nabla u$, Propositions 1.38 and 1.45 imply that

$$\begin{aligned}\int_\Omega F(x, u(x), \nabla u(x))\, dx &= \int_\Omega F(x, u(x), \overline{\nu}_x)\, dx \\ &\leq \int_\Omega F(x, u(x), \nu_x)\, dx \\ &\leq \liminf_{j\to\infty} \int_\Omega F(x, u_j(x), \nabla u_j(x))\, dx. \qquad \square\end{aligned}$$

A more general version of the above proposition, stated in the framework of BV functions, can be found in the paper [58] by Kristensen. However, the above formulation is sufficient for our usage in chapter 3.

2 Generic Well-posedness with Respect to Linear Perturbations

In this chapter, we study when the act of perturbing a variational integral $J(u) = \int_\Omega F(x, u(x), \nabla u(x))\,dx$, viewed as an $\overline{\mathbb{R}}$-valued nonlinear functional on a Sobolev space $J\colon \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N) \to \overline{\mathbb{R}}$, by a linear functional $\xi \in \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)^*$ affects the variational problem in such a way that the minimization problem $\min(\mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N); J - \xi)$, that is

$$\text{minimize } (J-\xi)(u) \quad \text{over} \quad u \in \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N),$$

associated with the perturbation ξ behaves better than the unperturbed problem. More specifically, we want to show that the complement of the set $\mathcal{A} \subset \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)^*$ of all ξ such that the problem $\min(\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N); J - \xi)$ fails to be Tykhonov well-posed is larger than a residual set (in Baire's sense). In fact, we will see that the set $\mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)^*\backslash\mathcal{A}$ satisfies a much stronger condition than being residual (at least when $p > 1$) and that the problem $\min(\mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N); J - \xi)$ is Hadamard well-posed for $\xi \in \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)^*\backslash\mathcal{A}$. The results extend naturally to the problems $\min(\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N); J - \xi)$ with arbitrary Dirichlet boundary datum $g \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ and constrained minimization problems.

Regarding the Tykhonov well-posedness of the unperturbed variational problem $\min(\mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N); J)$, there are some related results by Sychev in this direction that partially inspired our work in this chapter. In [95] (based on his earlier work in [94]), Sychev studied a slightly weaker property than Tykhonov well-posedness that he called the 'stability' of a variational integral $J\colon \mathrm{W}^{1,1}(\Omega;\mathbb{R}^N) \to \overline{\mathbb{R}}$, and proved some interesting results under relatively mild assumptions. Some of these assumptions also motivated our assumptions in Theorem 2.3. A point $u \in \mathrm{W}^{1,1}(\Omega;\mathbb{R}^N)$ is called a *point of stability* in Sychev's sense if for any sequence u_j in $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^N)$

$$u_j \overset{w}{\rightharpoonup} u \quad \text{and} \quad J(u_j) \to J(u) \implies \nabla u_j \to \nabla u \ \text{ in } \mathrm{L}^1.$$

He proved that under suitable assumptions, the set of points of stability in $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^N)$ is dense in $\{u : J(u) < \infty\}$. While our results in this chapter do not imply his result, it is worth noting that a strong minimizer of the perturbed problem $\min(\mathrm{W}^{1,1}(\Omega;\mathbb{R}^N); J - \xi)$ is a point of stability in Sychev's sense.

Our motivation for studying the generic Tykhonov well-posedness of minimization problems $\min(\mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N); J - \xi)$ for $\xi \in \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)^*$ is a theorem by Stegall [91], now commonly known as the *Stegall's variational principle*. A prevalent form of the statement of the Stegall's variational principle as commonly seen, e.g. in [13] and [83], is formulated as a perturbed minimization theorem in a form similar to the (more well-known) Ekeland variational principle. This version can be stated as a well-posedness theorem as follows.

Proposition 2.1 (Stegall's Variational Principle)**.** *Let X be a Banach space and $C \subset X$ be a nonempty, closed and bounded convex set with the Radon–Nikodym Property. Let $f\colon C \to \overline{\mathbb{R}}$ be a lsc function that is bounded from below. Then for any $\varepsilon > 0$, there exists $x^* \in X^*$ such that $\|x^*\| < \varepsilon$ and the problem of minimizing $f + x^*$ is Tykhonov well-posed.*

The Radon–Nikodym Property is defined later in this chapter, but for now it is helpful to note that every nonempty, closed and bounded convex set in a reflexive space X has this property.

While a direct proof of this version of Stegall's variational principle, although very technical, is available (see the aforementioned books, for examples), in fact an even stronger theorem can be obtained naturally as a consequence of results in convex analysis regarding the generic Fréchet differentiability of the Fenchel conjugates of a particular class of functions. This approach relies on the continuity properties (in the multivalued functions sense) of subdifferentials of convex functions in a nontrivial way. The goal of this chapter is to lay the foundations needed for the 'improved' version of the Stegall's variational principle, and then apply it to minimization problems in the calculus of variations. Note that the groundwork about genericity and convex analysis is a collection of known results in the literature, the original work in this chapter is the application of those techniques to variational integrals.

The main generic well-posedness results in this section concerns the variational integrals $J : \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \to \overline{\mathbb{R}}$ with Caratheodory integrand $F(x,u,z)$, where F satisfies a certain pointwise lower bound and J is *strongly coercive* in the sense that

$$\lim_{\|u\|_{\mathrm{W}^{1,p}}\to\infty} \frac{J(u)}{\|u\|_{\mathrm{W}^{1,p}}} = \infty.$$

These assumptions are relatively weak and they follow from several more familiar assumptions such as the p-coercivity assumption for $F(x,u,z)$ (see [21] for a discussion of coercivity in the calculus of variations). The case $p \in (1,\infty)$ is simpler than the case $p = 1$ and also allows us to draw a stronger conclusion regarding the generic well-posedness in this case.

Let us first recall some terminology and notation regarding genericity and the duals of Sobolev spaces that are used in the main theorems. A G_δ-set is a set that can be written as an intersection of countably many open sets, and its complement is an F_σ-set, a set that can be written as a union of countably many closed sets. A property is *generic* in a topological space X if it holds on a dense G_δ-subset of X, or, equivalently, fails to hold on an F_σ-subset of X with empty interior. An *α-angle small* set is a particular case of an F_σ-set with empty interior (α-angle smallness and other types of smallness will be discussed more thoroughly later in this chapter).

The dual of the Sobolev space $\mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^N)$ is denoted by $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)^*$ and can, for $p \in [1,\infty)$, be identified (isomorphically) with

$$\mathrm{W}^{-1,p'}(\Omega;\mathbb{R}^N) = \left\{ \sum_{|\alpha|\leq 1} \partial^\alpha f_\alpha : f_\alpha \in \mathrm{L}^{p'}(\Omega;\mathbb{R}^N) \right\}$$

normed by

$$\|v\|_{\mathrm{W}^{-1,p'}} := \inf \left\{ \sum_{|\alpha|\leq 1} \|\partial^\alpha f_\alpha\|_{\mathrm{L}^{p'}} : v = \sum_{|\alpha|\leq 1} \partial^\alpha f_\alpha \right\},$$

where p' is the conjugate exponent of p (i.e. $\frac{1}{p} + \frac{1}{p'} = 1$ and $p' = \infty$ if $p = 1$). Our main result regarding generic well-posedness of linearly perturbed problems with a given Dirichlet boundary datum is as follows.

Theorem 2.2. *Let $p \in (1,\infty)$, let $\Omega \subset \mathbb{R}^n$ be a bounded open domain, let $J\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \to \overline{\mathbb{R}}$ be a variational integral of the form $J(u) = \int_\Omega F(x,u,\nabla u)\,dx$, where the functional J is strongly coercive as defined above, and let $g \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$.*

Suppose that the integrand $F\colon \Omega \times \mathbb{R}^N \times \mathbb{R}^{N\times n} \to \overline{\mathbb{R}}$ of J is normal and satisfies the pointwise lower bound (1.8)

$$F(x,u,z) \geq -f(x) - \gamma|u|^t - \beta|z|^p,$$

for some $\gamma,\beta \geq 0$, $1 \leq t \leq p^*$ and $f \in \mathrm{L}^1(\Omega)$, then there exists a set $\mathcal{A} \subset \mathrm{W}^{-1,p'}(\Omega;\mathbb{R}^N)$ such that $\mathcal{A}$ is α-angle small for all $\alpha \in (0,1)$ and the problem

$$\text{minimize } (J-\xi)(u) \quad \text{over} \quad u \in \mathrm{W}_g^{1,p}(\Omega;\mathbb{R}^N)$$

is Tykhonov well-posed at each $\xi \in \mathrm{W}^{-1,p'}(\Omega;\mathbb{R}^N)\backslash\mathcal{A}$. Moreover, the problem $\min(\mathrm{W}_g^{1,p}(\Omega;\mathbb{R}^N); J-\xi)$ is Hadamard well-posed for $\xi \in \mathrm{W}^{-1,p'}(\Omega;\mathbb{R}^N)\backslash\mathcal{A}$, where $\mathrm{W}^{-1,p'}(\Omega;\mathbb{R}^N)$ is endowed with its usual norm topology.

In particular, if Ω is a domain with C^1 boundary, then there exists a set $\widetilde{\mathcal{A}} \subset \mathrm{W}_0^{1,p'}(\Omega;\mathbb{R}^N)$ such that $\widetilde{\mathcal{A}}$ is α-angle small for some $\alpha \in (0,1)$ and the problem

$$\text{minimize } \int_\Omega F(x,u,\nabla u)\,dx - \int_\Omega \nabla h : \nabla u\,dx \quad \text{over} \quad u \in \mathrm{W}_g^{1,p}(\Omega;\mathbb{R}^N)$$

is Tykhonov well-posed at each $h \in \mathrm{W}_0^{1,p'}(\Omega;\mathbb{R}^N)\backslash\widetilde{\mathcal{A}}$. Moreover, the problem $\min(\mathrm{W}_g^{1,p}(\Omega;\mathbb{R}^N); J-\langle\nabla h,\cdot\rangle)$ is Hadamard well-posed for $h \in \mathrm{W}_0^{1,p'}(\Omega;\mathbb{R}^N)\backslash\widetilde{\mathcal{A}}$, where $\mathrm{W}_0^{1,p'}(\Omega;\mathbb{R}^N)$ is endowed with the usual norm topology.

The expression $\xi(u)$ for a general $u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ should be understood as $\widetilde{\xi}(u)$ where $\widetilde{\xi}$ is an arbitrary extension of $\xi \in \mathrm{W}^{-1,p'}(\Omega;\mathbb{R}^N)$ to an element of $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)^*$. The set $\mathcal{A}$ (and $\widetilde{\mathcal{A}}$) in Theorem 2.2 is very small, smaller than a meagre set (in Baire's sense), so its complement in $\mathrm{W}^{-1,p'}(\Omega;\mathbb{R}^N)$ is very large.

For the case $p = 1$, the fact that the space $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^N)$ does not have the RNP complicates the analysis of generic Tykhonov well-posedness significantly. Nevertheless, a slightly weaker result than Theorem 2.2 is available if we assume a further pointwise strong coercivity condition for the integrand F, which yields the following result.

Theorem 2.3. *Let $\Omega \subset \mathbb{R}^n$ be an open domain and $J\colon \mathrm{W}^{1,1}(\Omega;\mathbb{R}^N) \to \overline{\mathbb{R}}$ be a variational integral of the form $J(u) = \int_\Omega F(x,u,\nabla u)\,dx$.*

Suppose that the integrand $F\colon \Omega \times \mathbb{R}^N \times \mathbb{R}^{N\times n} \to \overline{\mathbb{R}}$ of J is normal and satisfies the pointwise strong coercivity condition

$$F(x,u,z) \geq h(x) + \theta(|z|) \tag{2.1}$$

for some $h \in \mathrm{L}^1(\Omega)$ and some convex, increasing function $\theta\colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $\frac{\theta(t)}{t} \to \infty$ as $t \to \infty$. For any $g \in \mathrm{W}^{1,1}(\Omega;\mathbb{R}^N)$ such that $\theta(|\nabla g|) \in \mathrm{L}^1(\Omega)$, there exists a dense G_δ set $D \subset \mathrm{W}^{-1,\infty}(\Omega;\mathbb{R}^N)$ such that the problem

$$\text{minimize } (J-\xi)(u) \quad \text{over} \quad u \in \mathrm{W}^{1,1}_g(\Omega;\mathbb{R}^N)$$

is Tykhonov well-posed for $\xi \in D$. Moreover, the problem $\min(\mathrm{W}^{1,1}_g(\Omega;\mathbb{R}^N); J-\xi)$ is Hadamard well-posed for $\xi \in D$, where $\mathrm{W}^{-1,\infty}(\Omega;\mathbb{R}^N) \supset D$ is endowed with its usual norm topology.

This chapter begins with a section about genericity and its complementary notion: the σ-smallness, where we give a short review of Baire's theory and a quick introduction to porous sets and α-angle small sets. Then the background results in convex analysis and its relation to the geometries of Banach spaces are recalled. Lastly, we give the proofs of the main results stated above, together with a couple of related results.

2.1 Baire's Category Theorem and σ-Small Sets

In this section, we will introduce a few notions of "small" and "σ-small" sets which are used to quantify how large a subset of a Banach space is.

An equivalent way to say that a subset of a Banach space X is large is to say that its complement in X is small, so it would be beneficial to quantify smallness of various types. Based on this, we can then define a σ-ideal of subsets of X generated by sets that are considered small. Since the measure theoretic approach to the problem is often not feasible when X is an infinite dimensional Banach space, it is more natural to quantify smallness in X using its topological or metric structure.

The first type of σ-small sets we shall discuss is Baire's notion of *meagre sets* (originally called "sets of first category"), which is a purely topological concept. Nowadays, spaces that satisfies the conclusion of the classical Baire's category theorem are called *Baire spaces*.

Definition 2.4 (nowhere dense and meagre sets)**.** *Let X be a topological space.*

- *A subset S is said to be* nowhere dense *if the interior of its closure is empty: $(\overline{S})^\circ = \emptyset$.*
- *Any countable union of nowhere dense sets is said to be a* meagre set *(or a set of* first category*).*
- *A complement of a meagre set is called a* residual set*.*
- *If a set is not meagre, then we say that it is a set of* second category*.* //

Definition 2.5 (Baire Space)**.** *A topological space X is called a* Baire space *if any of the following five equivalent conditions is satisfied:*

1. *Every nonempty open set $U \subset X$ is of* second category*.*
2. *If $M \subset X$ is a meagre set, then $M^\circ = \emptyset$.*
3. *If $R \subset X$ is a residual set, then $\overline{R} = X$.*
4. *Suppose that each F_i is a closed nowhere dense set, then $(\bigcup_{i=1}^{\infty} F_i)^\circ = \emptyset$.*
5. *Suppose that each U_i is a open dense set, then $\overline{\bigcap_{i=1}^{\infty} U_i} = X$.* //

To see that all of these conditions are equivalent, see [24]. The class of Baire spaces contains two important classes of spaces that are commonly used in analysis.

Proposition 2.6 (Baire's Category Theorem)**.** *A topological space X is a Baire space if it satisfies any of the following properties.*

1. *X is completely metrizable.*
2. *X is a locally compact Hausdorff space.*

The proof of this can also be found in [24].

As we can see from the definition of Baire spaces, a space X is a Baire space if each meagre set is 'small' (in the sense that is made precise above), so the σ-ideal generated by the family of nowhere dense sets is appropriate for quantifying smallness in a Banach space (which is, of course, a complete metric space). Recall that a σ-ideal on X is a family $\mathcal{F}$ of subsets of X such that

- $\emptyset \in \mathcal{F}$.
- For all $A, B \subset X$ such that $A \in \mathcal{F}$ and $B \subset A$, we have $B \in \mathcal{F}$.
- If $A_1, A_2, \cdots \in \mathcal{F}$, then $\bigcup_{j=1}^{\infty} A_j \in \mathcal{F}$.

A σ-ideal $\mathcal{F}$ is generated by a family $\mathcal{B}$ of subsets of X if $\mathcal{F}$ is the smallest σ-ideal containing $\mathcal{B}$, in which case we call $\mathcal{B}$ the *generator* of $\mathcal{F}$. Intuitively, a generator $\mathcal{B}$ of a σ-ideal consists of sets that are so small that a countable union of sets from $\mathcal{B}$ is still a small set.

The word *generic* that was introduced in the early part of this thesis means that a certain property holds on a dense G_δ set. It is equivalent to stating that said property fails to hold on a meagre set.

The σ-ideal of all meagre sets, while sharing many similarities with the family of all Lebesgue null sets, does not coincide with it on $\mathbb{R}^n$. Indeed, there are meagre sets of full Lebesgue measure and, conversely, Lebesgue null sets of second category. However, there is a notion of σ-smallness on $\mathbb{R}^n$ which lies in the intersection between the family of all meagre sets and the family of all Lebesgue null sets. We call this notion *σ-porosity* and its definition easily generalizes to an arbitrary metric space. The following presentation of porosity and σ-porosity is based on the survey paper [109] by Zajíček.

Definition 2.7 (porosity)**.** *Let X be a metric space and let $M \subset X$ be a given subset. For any $x \in X$ and $R > 0$, we define*

$$\gamma(x, R, M) := \sup\big\{r \geq 0 : \textit{there exists } z \in X \textit{ such that } B(z, r) \subset B(x, R) \backslash M\big\}.$$

- *The* lower porosity *of M is defined as*
$$\underline{p}(M, x) := 2 \liminf_{R \searrow 0} \frac{\gamma(x, R, M)}{R}.$$
- *M is* c-lower porous *if there exists $c > 0$ such that $\underline{p}(M, x) > c$ for all $x \in M$.*
- *A set M is called* σ-lower porous *if it is a union of countably many σ-lower porous sets.* //

For a metric space X and $M \subset X$, if M is σ-lower porous then it is meagre. If X is a Euclidean space $\mathbb{R}^d$, then M is also a Lebesgue null set. It should be noted that, unlike nowhere dense sets whose closure are still nowhere dense, the closure of a porous set need not be porous. Such distinction is discussed in detail in [86].

The last notion of σ-ideal of subsets that we want to discuss is the family of *α-angle small* sets, which is even stronger that σ-porosity. An α-angle small set consists of a countable union of *α-cone meagre* sets, where each α-cone meagre set S has the property that each point in it can be approached by a cone that does not intersect S. The precise definition is as follows.

Definition 2.8 (α-cone meagre and α-angle small sets)**.** *Let X be a Banach space, $x^* \in X^*$ be a linear functional such that $x^* \neq 0$, and let $0 < \alpha < 1$.*

- *We define a closed convex cone $K(x^*, \alpha)$ in X by*

$$K(x^*, \alpha) := \{x \in X : \alpha \|x^*\| \|x\| \leq \langle x^*, x\rangle\}.$$

  Note that $K(x^*, \alpha)$ has nonempty interior. In fact $K(x^*, \alpha)^\circ = \{x \in X : \alpha \|x^*\| \|x\| < \langle x^*, x\rangle\}$.

- *A subset $A \subset X$ is called* α-cone meagre *if for each $x \in A$ and $\varepsilon > 0$, there exist $z \in B(x, \varepsilon)$ and $x^* \in X^* \backslash \{0\}$ such that*

$$A \cap (z + K(x^*, \alpha)^\circ) = \emptyset.$$

- *A subset $M \subset X$ is called* α-angle small *if it is a countable union of α-cone meagre sets.* //

Observe that $K(x^*, \alpha) \supset K(x^*, \beta)$ whenever $0 < \alpha \leq \beta < 1$. As α gets closer to 0, $K(x^*, \alpha)$ approaches a half-space, whereas $K(x^*, \alpha)$ shrinks into a thinner cone as α tends to 1. Intuitively, it is not hard to imagine that an α-cone meagre set must be porous. Indeed, an α-cone meagre set is c-lower porous for some $c > 0$ depending on α.

Proposition 2.9. *Let X be a Banach space and let $M \subset X$ be an α-cone meagre set. Then M is c-lower porous for some constant $c > 0$ depending on α. In particular, an α-angle small set is σ-lower porous.*

Proof. For each $x \in M$ and $\varepsilon > 0$, we can find $z \in B(x, \varepsilon)$ and $x^* \in X^*$ (depending on z) such that $M \cap (z + K(x^*, \alpha)^\circ) = \emptyset$. For some $\gamma \in (0, 1)$ to be determined later, we let $x' \in S_X$ be such that $\langle x^*, x'\rangle > \gamma \|x^*\|$ and define $\bar{x} := \gamma x'$.

For each $y \in B(\bar{x}, \delta)$ we have

$$\begin{aligned}\langle x^*, y\rangle &\geq \langle x^*, \bar{x}\rangle - |\langle x^*, \bar{x} - y\rangle| \\ &> \gamma^2 \|x^*\| - \delta \|x^*\|.\end{aligned}$$

By choosing $\delta = \min\{(\frac{1-\alpha}{2})\gamma^2, 1 - \gamma\}$, we get $B(\bar{x}, \delta) \subset B(0, 1)$ and that

$$\langle x^*, y\rangle > \left(\frac{1+\alpha}{2}\right)\gamma^2 \|x^*\|$$

for all $y \in B(\bar{x}, \delta)$. Now, we may take $\gamma > \sqrt{\frac{2\alpha}{1+\alpha}}$ and hence getting

$$\langle x^*, y\rangle > \alpha \|x^*\| > \alpha \|x^*\| \|y\|$$

for all $y \in B(\bar{x}, \delta)$. This shows that $B(\bar{x}, \delta) \subset K(x^*, \alpha)^\circ$ and since $K(x^*, \alpha)^\circ$ is a cone, we in fact have $B(r\bar{x}, r\delta) \subset K(x^*, \alpha)^\circ$ for all $r > 0$.

For each $r > 0$, we may choose ε in the above to be $\varepsilon = r/2$ and the above discussion shows that we can find z_r, x_r^* and $\bar{x}_r$ such that $z_r + B(r\bar{x}_r, r\delta) \subset B(x, 3r/2)\backslash M$ since $B(r\bar{x}_r, \delta) \subset B(0, r)$, $z_r \in B(0, r/2)$ and $M \cap (z_r + K(x_r^*, \alpha)^\circ) = \emptyset$ by the assumption. This shows that $\underline{p}(M, 0) \geq 2(\frac{r\delta}{3r/2}) = 4\delta/3$. Since δ depends only on α and not on x, z_r or x_r^*, the proof is complete. □

2.2 Convex Analysis and Geometries of Banach Spaces

There is a strong tie between convex analysis and variational problems in calculus of variations, and in this case we are especially interested in its relationship to the well-posedness of linearly perturbed problems. Even though the variational integral $J\colon \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N) \to \overline{\mathbb{R}}$ of the form $J(u) = \int_\Omega F(x, u, \nabla u)\, dx$, considered in this chapter is not a convex functional on $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$, many tools from convex analysis turn out to be useful for studying the well-posedness of the problem $\min(\mathrm{W}_g^{1,p}(\Omega; \mathbb{R}^N); J - \xi)$, where $\xi \in \mathrm{W}^{-1,p'}(\Omega; \mathbb{R}^N)$. This section will go through some background materials in convex analysis, the geometries of Banach spaces and how they can be applied to the study of generic well-posedness in the calculus of variations.

First, let us recall some facts about convexity and lower semicontinuity of a function whose domain is a Banach space with its range in $\overline{\mathbb{R}} := (-\infty, \infty]$

Definition 2.10. *Let X be a Banach space and let $f\colon X \to \overline{\mathbb{R}}$ be a function*

- *The effective domain of f is denoted by*

$$\operatorname{dom} f := \{x \in X : f(x) < \infty\}$$

- *The set of all points of continuity of f is denoted by* $\operatorname{cont} f$ *(by convention,* $\operatorname{cont} f \subset \operatorname{dom} f$ *always).*

- *We say that f is proper if* $\operatorname{dom} f \neq \emptyset$.

- *The family of all proper, convex, lsc functions on X is denoted by*

$$\Gamma(X) := \big\{f\colon X \to \overline{\mathbb{R}} : f \text{ is proper, convex and lsc}\big\}.$$

//

Next, we introduce the *subdifferentiability* and *Fenchel conjugate* and their important properties. The subdifferential and ε-subdifferential of f encode its global behaviour on X regarding how its epigraph can be touched (or nearly touched) from below by affine functions. For this reason, the subdifferential of $J\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \to \overline{\mathbb{R}}$ is intimately connected to the minimization problem $\min(\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N); J - \xi)$ for $\xi \in X^*$.

The *subdifferential* ∂f of f can be defined for any proper function f, but it has particularly nice properties when f satisfies further assumptions such as convexity and continuity.

Definition 2.11 (subdifferential)**.** *Let X be a normed space, let $f\colon X \to \overline{\mathbb{R}}$ be a proper function and $\varepsilon \geq 0$.*

- *The* ε-subdifferential *of f is the multifunction $\partial_\varepsilon f\colon X \rightrightarrows X^*$ defined by*

$$\partial_\varepsilon f(x) := \big\{x^* \in X^* : \langle x^*, y - x\rangle \leq f(y) - f(x) + \varepsilon \text{ for all } y \in X\big\}$$

  for each $x \in \operatorname{dom} f$.

- *For the case $\varepsilon = 0$, we omit ε from the notation and merely write $\partial f := \partial_0 f$ to denote the* subdifferential *of f:*

$$\partial f(x) := \big\{x^* \in X^* : \langle x^*, y - x\rangle \leq f(y) - f(x) \text{ for all } y \in X\big\}.$$

- *Each $x^* \in \partial f(x)$ is called a* subderivative *of f at x.* //

For a proper convex function f, lower semicontinuity (resp. continuity) is strong enough to guarantee that $\partial_\varepsilon f$ (resp. ∂f) does not map to $\emptyset$.

A careful consideration of the definition of the subdifferential ∂f reveals that its inverse is directly related to the solutions of the perturbed minimization problem of f by linear functionals. Indeed, for any given $x^* \in X^*$, we can see that

$$\begin{aligned} \partial f^{-1}(x^*) &= \{x \in X : x^* \in \partial f(x)\} \\ &= \operatorname*{argmin}_X (f - x^*) \end{aligned} \tag{2.2}$$

since $x^* \in \partial f(x)$ if and only if $f(x) - \langle x^*, x\rangle \leq f(y) - \langle x^*, y\rangle$ for all $y \in X$, i.e. x is a minimizer of $f - x^*$. Here $\partial f^{-1} := (\partial f)^{-1}$ is the inverse of the multifunction ∂f.

Unlike differentiation of a real-valued function, taking the subdifferential of f is not a local operation. In the one-dimensional case, a subderivative $x^* \in \partial f(x)$ is the slope of a line tangent to the graph of f at the point $(x, f(x))$ such that the whole line lies below the epigraph of f, so it takes into account the value of f on the entire domain. For a convex function $f\colon \mathbb{R} \to \overline{\mathbb{R}}$, however, a tangent line to any

point $(x, f(x))$ necessarily lies below the entire epigraph of f, so it should not be surprising that ∂f behaves better if f is convex.

The structure of the set $\partial f(x)$ is relatively simple at each $x \in \operatorname{dom} f$ and has an Upper semicontinuity property (the proof of this can be found in [83]).

Proposition 2.12. *Let X be a normed space, let $f\colon X \to \overline{\mathbb{R}}$ be proper and $x \in \operatorname{dom} f$. The set $\partial f(x)$ is convex and w^*-closed. If we also assume that f is convex and $x \in \operatorname{cont} f$, then $\partial f(x)$ is w^*-compact and nonempty. Moreover, if f is locally Lipschitz with constant K on $U \in \mathcal{N}(x)$, then $\partial f(U) \subset KB_{X^*}$.*

It is known that for a convex function f, if $\operatorname{cont} f \neq \emptyset$, then f is locally Lipschitz on $(\operatorname{dom} f)^\circ$, hence the hypothesis for the last part of Proposition 2.12 can always be satisfied by choosing an appropriate $U \in \mathcal{N}(x)$ (recall that $\operatorname{cont} f \subset (\operatorname{dom} f)^\circ$). In fact, the following is true (see [12]).

Proposition 2.13. *Let $f\colon X \to \overline{\mathbb{R}}$ be convex and $x \in \operatorname{cont} f$, then $\partial f\colon X \rightrightarrows X^*$ is $\|\cdot\|$-w^* USC at x.*

Suppose that $f\colon \mathbb{R} \to \mathbb{R}$ is a convex function; we can recover f if we know, for each $c \in \mathbb{R}$, where the tangent line with slope c intersects the y-axis. This idea generalizes to any $f \in \Gamma(X)$, where X is a Banach space but in place of a 'slope of a tangent line', which is a real number $c \in \mathbb{R}$, we now consider a functional $x^* \in X^*$ associated with the affine functions that is tangent to f. The *Fenchel conjugate* f^* of f encodes the information regarding the behaviour of the affine functions that minorize f.

Definition 2.14 (Fenchel conjugate)**.** *Let X be a Banach space and let $f\colon X \to \overline{\mathbb{R}}$ be proper.*

- *The* Fenchel conjugate *of f is the function $f^*\colon X^* \to \overline{\mathbb{R}}$ defined by*

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

- *For a proper function $g\colon X^* \to \overline{\mathbb{R}}$, its Fenchel conjugate is the function g^* defined on X^{**}. When $g = f^*$, g^* will be called the* second conjugate *and is denoted by*

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

//

Remark 2.15.

- The above definition is well-defined since we assumed f to be proper, hence f^* cannot take the value $-\infty$.

- f^* need not be proper if f is not bounded below (e.g. the real-valued function $f(x) = -x^2$ has the Fenchel dual $f^* \equiv +\infty$). For f^* to be proper, it is necessary and sufficient that there exists an affine function minorizing f.
- Since f^* is the supremum of a family of affine functions (which are convex and continuous), f^* is convex and lower semicontinuous, i.e. $f^* \in \Gamma(X^*)$. In fact, f^* is even w^*-lsc.
- In this thesis f^{**} is the function $f^{**}\colon X^{**} \to \overline{\mathbb{R}}$ whose domain is X^{**}. Some authors use the symbol f^{**} to denote the restriction of our f^{**} to X, i.e. $f^{**}|_X$. //

The subdifferential ∂f^* of f^* plays many important roles in the minimization problem of $f - x^*$. This fact is hinted already in the definition of f^* since it can be rewritten as

$$f^*(x^*) = -\inf_X (f - x^*).$$

The role of f^* will become even more prominent when we discuss the relationship between subdifferential and Fréchet differentiability. The following proposition shows that the single-valuedness of ∂f^* implies that the minimizer of $f - x^*$, if it exists, must be unique.

Proposition 2.16. *Let X be a Banach space and let $f\colon X \to \overline{\mathbb{R}}$ be proper. Then*

$$\operatorname*{argmin}_X (f - x^*) \subseteq \partial f^*(x^*).$$

The two sets are equal if X is reflexive. //

The proof of this statement is not hard and almost follows from the definition of subdifferential (see [105] or [12]) This results can be strengthened significantly whenever f^* is Fréchet (or Gâteaux) differentiable at x^*.

Given a convex function $f\colon X \to \overline{\mathbb{R}}$, the Fréchet differentiability of f forces the values of ∂f at these points of differentiability to be singletons, and also give ∂f some desirable continuity properties (in the sense of multivalued functions). As was hinted in Proposition 2.16, it is reasonable to expect that the Fréchet differentiability of f^* is ultimately connected to the Tykhonov well-posedness.

To distinguish it from the subdifferential, we shall denote the Fréchet (or Gâteaux) derivative of f by f' or ∇f. For a convex function $f\colon X \to \overline{\mathbb{R}}$, there is an immediate characterization of Gâteaux differentiability in term of its subdifferential. The proof of this can be found in [12].

Proposition 2.17. *Let X be a Banach space and $f\colon X \to \overline{\mathbb{R}}$ be a convex function. Suppose that $x_0 \in \operatorname{cont} f$, then f is Gâteaux differentiable at x_0 if and only if $\partial f(x_0)$ is a singleton.*

$\partial f(x_0)$ being a singleton alone is not enough to guarantee that f is Fréchet differentiable at x_0 but it is a necessary condition. The extra bit of information that we need is the upper semicontinuity property of ∂f at x_0. According to Proposition 2.13, for a convex function f, its subdifferential ∂f is $\|\cdot\|$-w^* USC at the points of continuity. If we can show that a stronger continuity property holds, i.e. that ∂f is $\|\cdot\|$-$\|\cdot\|$ USC at x_0, then it is Fréchet differentiable there (the proof can be found in [12] as well).

Proposition 2.18 (Šmulian's Theorem)**.** *Let X be a Banach space and $f\colon X \to \overline{\mathbb{R}}$ be a convex function. Suppose that $x_0 \in \operatorname{cont} f$, then the following are equivalent.*

1. *f is Gâteaux (resp. Fréchet) differentiable at x_0.*
2. *$\partial f(x_0)$ is a singleton and ∂f is $\|\cdot\|$-w^* USC (resp. $\|\cdot\|$-$\|\cdot\|$ USC) at x_0.*

One of the most important results that will be applied to the problems in the calculus of variations is the following proposition which shows a strong connection between Fréchet differentiability of f^* and the Tykhonov well-posedness of the linearly perturbed minimization problem for f (see [63] by Lassonde).

Proposition 2.19. *Let X be a Banach space and let $f\colon X \to \overline{\mathbb{R}}$ be proper and lsc. Let $x^* \in (\operatorname{dom} f^*)^\circ$, then the following are equivalent:*

1. *f^* is Fréchet differentiable at x^*.*
2. *$f - x^*$ attains a strong minimum on X.*

Recall that $f - x^*$ attaining a strong minimum means that the problem $\min(X; f - x^*)$ is Tykhonov well-posed.

Proposition 2.19 says that we can study the Tykhonov well-posedness of the minimization problem of the functional $f - x^*$ over X by studying the differentiability of f^* at x^* instead, provided that all the necessary assumptions are satisfied. Hence, the generic well-posedness of our minimization problem is intimately connected to the generic differentiability of a convex and lsc function, which is a topic that is relatively thoroughly studied.

The following result is based on a theorem from a paper [84] by Preiss and Zajíček on the single-valuedness of monotone operators (the subdifferential ∂f is an example of a monotone operator).

Theorem 2.20 (Preiss–Zajíček)**.** *Let X be a Banach space such that X^* is separable and let $T\colon X \rightrightarrows X^*$ be a monotone operator. Then there exists a subset $\mathcal{A} \subset \operatorname{Dom}(T)$ such that A is α-angle small for all $\alpha \in (0, 1)$ and T is single-valued and $\|\cdot\|$-$\|\cdot\|$ USC on $\operatorname{Dom}(T)\backslash\mathcal{A}$.*

This theorem, together with the characterization of Fréchet differentiability, leads immediately to generic differentiability.

Proposition 2.21. *Let X be a Banach space such that X^* is separable and $f\colon X \to \overline{\mathbb{R}}$ be a proper convex lsc function. Then there exists $\mathcal{A} \subset (\operatorname{dom} f)^\circ$ such that $\mathcal{A}$ is α-angle small for all $\alpha \in (0,1)$ and f is Fréchet differentiable on $(\operatorname{dom} f)^\circ \backslash \mathcal{A}$. In particular, f is generically Fréchet differentiable on $(\operatorname{dom} f)^\circ$.*

Proof. Recall that $(\operatorname{dom} f)^\circ = \operatorname{cont} f \subset \operatorname{Dom} \partial f$ whenever f is a proper convex lsc function. Since ∂f is monotone, Theorem 2.20 says that there exists an α-angle small set $\widetilde{\mathcal{A}} \subset \operatorname{Dom} \partial f$ such that ∂f is single-valued and $\|\cdot\|$-$\|\cdot\|$ USC on $(\operatorname{Dom} \partial f) \backslash \widetilde{\mathcal{A}}$.

Let $\mathcal{A} := (\operatorname{dom} f)^\circ \cap \widetilde{A}$, then it is easy to verify that $\mathcal{A}$ is also α-angle small for all $\alpha \in (0,1)$ and ∂f is single-valued and $\|\cdot\|$-$\|\cdot\|$ USC on $(\operatorname{dom} f)^\circ \backslash \mathcal{A}$. By Šmulian's theorem (Proposition 2.18), f is Fréchet differentiable at x if and only of $\partial f(x)$ is a singleton and ∂f is $\|\cdot\|$-$\|\cdot\|$ USC at x. Since $(\operatorname{dom} f)^\circ = \operatorname{cont} f$, we conclude that f is Fréchet differentiable on $(\operatorname{dom} f)^\circ \backslash \mathcal{A}$. □

The previous proposition shows that a proper convex lsc function $f\colon X \to \overline{\mathbb{R}}$ is generically differentiable on a very large subset of $(\operatorname{dom} f)^\circ$, provided that X has a separable dual, but this assumption proves to be too restrictive in some cases, e.g. when X is the dual of $\mathrm{W}_0^{1,1}(\Omega)$. For this reason, the *Radon–Nikodym Property* (RNP) which is a more refined geometric property of a normed space X is needed. The definition and background materials related to the Radon–Nikodym Property can be found in the appendix. For a thorough discussion about generic Fréchet differentiability, see [67] by Lindenstrauss, Preiss, and Tišer.

One of the reasons for studying spaces (or sets) with the RNP is that, roughly speaking, the RNP is dual to the statement "every continuous convex functions are generically differentiable on its domain". The exact formulation of this duality requires the notion of an *Asplund space*.

Definition 2.22 (Asplund space)**.** *Let X be a Banach space, $U \subset X$ and $V \subset X^*$ be nonempty open sets. Then*

1. *X is an* Asplund space *if any lsc convex function $f\colon U \to \mathbb{R}$ is generically Fréchet differentiable.*

2. *X is a* weak Asplund space *if any lsc convex function $f\colon U \to \mathbb{R}$ is generically Gâteaux differentiable.*

3. *X^* is a w^*-*Asplund space *if any w^*-lsc convex function $f\colon V \to \mathbb{R}$ is generically Fréchet differentiable.* //

According to Theorem 2.20, X is an Asplund space if X^* is separable. In fact, a separable space is Asplund if and only if its dual is separable. A detailed treatment of Asplund and weak Asplund spaces and their characterizations can be found in [83].

The partial duality between spaces with the RNP and Asplund spaces was studied by Asplund [8]; Namioka and Phelps [81]; and Collier [26].

Proposition 2.23 (Asplund-Namioka-Phelps-Collier)**.** *Let X be a Banach space, then*

1. *X is Asplund if and only if X^* has the w^*-RNP.*
2. *X has the RNP if and only if X^* is w^*-Asplund.*

Moreover, on X^* the w^*-RNP and the RNP are equivalent.

Since the Fenchel conjugate f^* of a proper function $f\colon X \to \overline{\mathbb{R}}$ is w^*-lsc and convex, the above theorem and Proposition 2.19 shows that X having the RNP is enough to guarantee the generic Tykhonov well-posedness of the linearly perturbed minimization problem associated with f. It is well-known that reflexive spaces have the RNP but spaces like $C(K)$ or $\mathrm{L}^1(\Omega)$ lack the RNP, so Theorem 2.23 would apply when $X = \mathrm{W}^{1,p}(\Omega)$ with $p > 1$ but not when $X = \mathrm{W}^{1,1}(\Omega)$. For this reason, we need to consider the RNP of subsets of X instead of the RNP on the whole space. A thorough treatment of convex sets with the RNP can be found in Bourgin's monograph [14]. A very useful theorem that we are going to present can be dated back to Lindenstrauss' work [66] on weakly compact convex sets and was later improved upon by Troyanski [104], the proof of it can be found in [14].

Proposition 2.24 (Lindenstrauss-Troyanski)**.** *Let X be a Banach space and let $K \subset X$ be a weakly compact convex set, then K has the RNP. Moreover, K is the closure of the convex hull of its strongly exposed points.*

A *strongly exposed point* of a closed and convex set $C \subset X$ is a point $x \in C$ such that there exists $\xi \in X^*$ that verifies

$$\xi(x_j) \to \inf_C \xi \implies x_j \to x$$

for any sequence $(x_j) \subset C$. The final part of Proposition 2.24 is not needed in this thesis, however.

Similar to the characterization given in Proposition 2.23, the RNP on a convex set C can be characterized by the Fréchet differentiability of the Fenchel conjugates of the family of lsc functions on C. The proof of the following proposition can be found in [12, Theorem 6.6.12].

Proposition 2.25. *Suppose that X is a Banach space and C is a closed and bounded convex subset of X. Then C has the RNP if and only if for each proper lsc function $f\colon C \to \overline{\mathbb{R}}$ that is bounded from below, its Fenchel conjugate f^* is Fréchet differentiable on a dense G_δ-subset of X^*.*

2.3 Applications of Convex Analysis to Linearly Perturbed Variational Integrals

In this last section of the chapter, we apply the results from the previous section to prove the generic well-posedness of minimization problems of the form $\min(\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N); J - \xi)$, where $\xi \in \mathrm{W}^{-1,p'}(\Omega;\mathbb{R}^N)$, together with a few related problems. Apart from the pointwise lower bound (1.8) for the integrand $F(x,u,z)$ of the variational integral $J(u) = \int_\Omega F(x,u,\nabla u)\,dx$, which guarantees strong lower semicontinuity, we need to assume a coercivity condition on the functional J to ensure that its Fenchel conjugate J^* has finite values.

Definition 2.26 (strong coercivity)**.** *Let X be a Banach space. A functional $J\colon X \to \overline{\mathbb{R}}$ is said to be* strongly coercive *if*

$$\lim_{\|u\|\to\infty} \frac{J(u)}{\|u\|} = \infty. \tag{2.3}$$

//

This terminology is commonly used in the field of convex analysis (for example, in [12]). For the case $p > 1$, the strong coercivity of J follows from many other common assumptions such as the pointwise *p-coercivity* assumption (1.7) on the integrand.

Lemma 2.27. *Let $X = \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ and let $J\colon X \to \overline{\mathbb{R}}$ be a strongly coercive variational integral of the form (1.1) whose integrand $F(x,u,z)$ is normal and satisfies pointwise the lower bound (1.8). Then its Fenchel conjugate J^* is finite on X^*, i.e.* $\operatorname{Dom} J^* = \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)^*$.

Proof. Let $\xi \in X^*$ be given and suppose that $\|\xi\| = M$. By the superlinear growth of J, there exists an $r > 0$ such that $J(u) \geq M\|u\|$ on $X\backslash B(0,r)$.

From the lower bound assumption, we may integrate (1.8) over Ω and get

$$J(u) \geq \|h\|_{\mathrm{L}^1} - C'\|u\|^p_{\mathrm{W}^{1,p}}$$

for some $C' > 0$. Hence for $u \in B(0,r)$, we have $\|g\|_{\mathrm{L}^1} - rC' \geq -C$ for some constant $C > 0$. Thus

$$J(u) \geq -C \geq \langle \xi, u\rangle - (C + rM)$$

for all $u \in B(0,r)$.

For $u \in X \backslash B(0,r)$, we have

$$J(u) \geq M \|u\|_{\mathrm{W}^{1,p}} \geq |\langle \xi, u \rangle| \geq \langle \xi, u \rangle - (C + rM).$$

Therefore the functional J is minorized by the affine function $\xi - (C + rM)$ on the whole space X. This implies that $J^*(\xi) < \infty$ and since ξ is arbitrary, we thus have $\operatorname{Dom} J^* = X^*$. □

We are now ready to state and prove our first generic Tykhonov well-posedness theorem.

Theorem 2.28. *Let $p > 1$, $\Omega \subset \mathbb{R}^n$ be an open domain, $X = \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$ or $\mathrm{W}^{1,p}_0(\Omega; \mathbb{R}^N)$ and let $J \colon X \to \overline{\mathbb{R}}$ be a strongly coercive variational integral of the form (1.1). Suppose that the integrand $F(z,u,z)$ of J is normal and satisfies the lower bound (1.8). Then there exists a set $\mathcal{A} \subset X^*$ such that $\mathcal{A}$ is α-angle small for all $\alpha \in (0,1)$ and the problem $\min(X; J - \xi)$ is Tykhonov well-posed at each point $\xi \in X^* \backslash \mathcal{A}$. Moreover, the problem $\min(X; J - \xi)$ is Hadamard well-posed for $\xi \in X^* \backslash \mathcal{A}$, where X^* is endowed with its norm topology.*

Proof. From Proposition 1.21 and Lemma 2.27, our integral functional J is lsc in the strong topology of X and $\operatorname{Dom} J^* = X^*$. Since $\operatorname{cont} J^* = (\operatorname{Dom} J^*)^\circ$, it follows that J^* is continuous on X^*.

The function $J^* \colon X^* \to \overline{\mathbb{R}}$ is a continuous convex function and since X is a separable reflexive space, $X^{**} = X$ is separable. Hence Proposition 2.21 guarantees the existence of a set $\mathcal{A} \subset X^*$ such that $\mathcal{A}$ is α-angle small for all $\alpha \in (0,1)$ and J^* is Fréchet differentiable on $X \backslash \mathcal{A}$. By Proposition 2.19, $J - \xi$ attains a strong minimum for each $\xi \in X^* \backslash \mathcal{A}$, which is equivalent to Tykhonov well-posedness of the problem $\min(X; J - \xi)$. Since $\mathcal{A}$ is porous and hence meagre, this proves the generic Tykhonov well-posedness of $\min(X; J - \xi)$.

To prove Hadamard well-posedness, recall from Proposition 2.16 that

$$\operatorname{argmin}(X; J - \xi) = \partial J^*(\xi)$$

since our space X is reflexive. At each point $\xi \in X^* \backslash \mathcal{A}$, J^* is Fréchet differentiable hence ∂J^* is $\|\cdot\|$-$\|\cdot\|$ USC by Šmulian theorem (Proposition 2.18). Hence $\min(X; J - \xi)$ is Hadamard well-posed in the generalized sense at all such points.

Lastly, if we consider only $\xi \in X^* \backslash \mathcal{A}$, $\operatorname{argmin}(X, J - \xi)$ is single-valued for all such ξ and hence it is continuous in the usual sense. This proves solution Hadamard well-posedness for $\xi \in X^* \backslash \mathcal{A}$. □

Note that Proposition 2.19 which is used here only assumes that J is lsc without requiring J to be convex. There is a weaker relative of this proposition, which is

perhaps more well-known, that assumes J to be convex as well to get the same conclusion. This weaker version of Proposition 2.19 is inadequate for our purpose but can be found in [35], for example.

A few generalizations of Theorem 2.28 is readily available with some modifications. The main result regarding minimization problems with fixed Dirichlet boundary data g also follows from the above theorem.

Proof of Theorem 2.2. It is not hard to verify that for $g \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$, the functional $J_g(u) := J(u-g)$ has superlinear growth and its integrand

$$F_g(x,u,z) := F(x,u-g(x),z-\nabla g(x))$$

satisfies the same lower bound (1.8) as $F(x,u,z)$ (but with different constants). This shows that J_0 satisfies the assumptions from Theorem 2.28 with $X = \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)$ and $X^* = \mathrm{W}^{-1,p'}(\Omega;\mathbb{R}^N)$.

From Theorem 2.28, there exists an α-angle small subset $\mathcal{A} \subset \mathrm{W}^{-1,p'}(\Omega;\mathbb{R}^N)$ such that the problem $\min(\mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N); J_g - \xi)$ is Tykhonov well-posed for all $\xi \in \mathcal{A}$. We may extend ξ (by the Hahn-Banach theorem, for example) to a functional in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)^*$ and add the constant $\xi(g)$ to the functional $J_g - \xi$ to get an equivalent minimization problem

$$\text{minimize } J(u-g) - \xi(u-g) \quad \text{over} \quad u \in \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N).$$

This problem is Tykhonov well-posed since adding a constant does not change the minimizer or the behaviour of minimizing sequences, and it is clearly equivalent to the problem $\min(\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N); J - \xi)$. The statement about Hadamard well-posedness is also similarly transferred.

When Ω is a C^1 domain, there is an isomorphism between $\mathrm{W}^{-1,p'}(\Omega;\mathbb{R}^N)$ and $\mathrm{W}^{1,p'}_0(\Omega;\mathbb{R}^N)$ for $p > 1$ (see [89] or [62]), which means that for any $\xi \in \mathrm{W}^{-1,p'}(\Omega;\mathbb{R}^N)$, there exists $h \in \mathrm{W}^{1,p'}_0(\Omega;\mathbb{R}^N)$ such that

$$\xi(u) = \int_\Omega \nabla h : \nabla u \, dx$$

for all $u \in \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)$ (this formula gives an extension of ξ to a functional on $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ in an obvious way). Since the property of being an α-angle small set is invariant under equivalent norm (provided that we replace α with a suitable α'), the conclusion from the previous part of this theorem continues to hold under this equivalent representation of ξ. □

Incorporating other kinds of constraints besides the boundary constraint are also possible. The kind of constraints we shall consider here are integral side constraints.

Theorem 2.29. *Under the same assumptions and notations as in Theorem 2.28, let $K \subset X$ be either*

1. $K = \{u \in X : H_2(u) \leq c\}$, *where* H_2 *is an integral functional whose integrand satisfies* (1.8).

2. $K = \{u \in X : H_1(u) = c\}$, *where* H_1 *is an integral functional whose integrand satisfies* (1.9).

Then the problem $\min(K; J - \xi)$ *is both Tykhonov well-posed at each point* $\xi \in X^* \backslash \mathcal{A}$ *and is Hadamard well-posed for* $\xi \in X^* \backslash \mathcal{A}$, *where* X^* *is endowed with its norm topology. Moreover, the same modifications work for Theorem 2.2 if we replace the problem* $\min(X; J - \xi)$ *by* $\min(K; J - \xi)$.

Proof. By Corollary 1.22 (resp. Proposition 1.21), H_1 (resp. H_2) is continuous (resp. lsc) in $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$, hence K is a closed set. Let us define the function $\delta_K \colon \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N) \to \overline{\mathbb{R}}$ to be

$$\delta_K(u) = \begin{cases} 0 & ; u \in K \\ \infty & ; u \notin K. \end{cases} \tag{2.4}$$

Since K is closed, δ_K is lsc.

Now, we define $\widetilde{J} := J + \delta_K$. $\widetilde{J}$ is lsc since it is the sum of two lsc functions. Clearly, $\widetilde{J}$ has superlinear growth, hence the proof of Theorem 2.28 works for $\widetilde{J}$. It is not hard to verify that the problems $\min(X; \widetilde{J} - \xi)$ and $\min(K; J - \xi)$ are equivalent, hence we are done.

The last part of this theorem can be proved the same way. □

Theorem 2.28 and its subsequent relatives rely heavily on reflexivity of $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$ with $p > 1$, so the same approach will not work if we want to prove well-posedness for the problem $\min(\mathrm{W}_0^{1,1}(\Omega; \mathbb{R}^N); J - \xi)$. In this case, the RNP will replace the role of reflexivity. However, the conclusion of the theorem will be slightly weaker as we can no longer show that the set where the Tykhonov well-posedness fails is an α-angle small set, we can only show that it is meagre.

In the proof of well-posedness for the problem $\min(\mathrm{W}_0^{1,1}(\Omega; \mathbb{R}^N); J - \xi)$, the following result by Krein and Smulian is crucial.

Proposition 2.30 (Krein–Šmulian Theorem)**.** *Let* X *be a Banach space and* $K \subset X$ *be a weakly compact subset. Then* $\overline{\operatorname{conv}} K$, *the closed convex hull of* K, *is also weakly compact.*

Proof. See [6, Theorem 3.42]. □

Before proving the next theorem, let us recall de la Vallée–Poussin's characterization of weak compactness in L^1.

Proposition 2.31 (de la Vallée–Poussin criterion)**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded open domain. Suppose that $\mathcal{F} \subset \mathrm{L}^1(\Omega;\mathbb{R}^N)$ is a given family of functions. Then $\mathcal{F}$ is relatively weakly compact in $\mathrm{L}^1(\Omega;\mathbb{R}^N)$ if and only if there exists an increasing function $\theta\colon [0,\infty) \to [0,\infty)$ such that*

1. $\frac{\theta(t)}{t} \to \infty$ *as* $t \to \infty$.
2. $\sup_{\mathcal{F}} \int_\Omega \theta(|f(x)|)\,dx < \infty$.

Additionally, if such a function exists we may take it to be convex.

Now we are ready to prove the generic well-posedness theorem for the problem $\min(\mathrm{W}^{1,1}_0(\Omega;\mathbb{R}^N); J - \xi)$.

Theorem 2.32. *Let $\Omega \subset \mathbb{R}^n$ be a bounded open domain, $X = \mathrm{W}^{1,1}(\Omega;\mathbb{R}^N)$ or $\mathrm{W}^{1,1}_0(\Omega;\mathbb{R}^N)$ and $J\colon X \to \overline{\mathbb{R}}$ be an integral functional of the form (1.1). Suppose that the integrand F of J is normal and satisfies the* pointwise strong coercivity *condition in the last variable, i.e. we have*

$$F(x,v,z) \geq h(x) + \theta(|z|) \tag{2.5}$$

for some $h \in \mathrm{L}^1(\Omega)$ and some increasing function $\theta\colon [0,\infty) \to [0,\infty)$ such that $\frac{\theta(t)}{t} \to \infty$ as $t \to \infty$. Then there exists a G_δ-subset of X^* such that the problem $\min(X; J - \xi)$ is Tykhonov well-posed at each $\xi \in D$. Moreover, the problem $\min(X; J - \xi)$ is Hadamard well-posed for $\xi \in D$, where $X^* \supset D$ is endowed with its norm topology.

Proof. First we observe that (2.5) implies the lower bound (1.8) for $p = 1$, hence J is strongly lsc in $\mathrm{W}^{1,1}(\Omega;\mathbb{R}^N)$. Moreover, the pointwise strong coercivity in the last variable of $F(x,u,z)$ also implies that the functional J is strongly coercive in the sense of (2.3). Since we can add a constant to the functional J without altering its generic Tykhonov well-posedness property, we may assume without loss of generality that there exists $\rho > 0$ and $u_0 \in B(0,\rho)$ such that $J(u_0) = 0$.

Suppose that $M > 0$ is given, since J is strongly coercive we can find $R > 0$ such that

$$J(u) > M \|u\|_{\mathrm{W}^{1,1}}$$

whenever $u \in X \backslash B_X(0,R)$. We can choose this R so that $R > \rho$ also. Now, let us define the set $S = S_M$ to be

$$S_M := \{u \in X : J(u) \leq RM\}.$$

It is easy to check that $S \subset B_X(0,R)$. The condition (2.5) and de la Vallée–Poussin criterion implies that S is a relatively weakly compact set, so $\overline{S}^w$ is a

weakly compact set. By the Krein–Šmulian theorem, the set $K := \overline{\text{conv}}\,(\overline{S}^w)$ is also a weakly compact set.

Let $J|_K := J + \delta_K$, where δ_K is defined as in (2.4). We will first show that J^* and $J|_K^*$ coincide on $B_{X^*}(0, M/2)$. Recall that for a function $I\colon X \to \overline{\mathbb{R}}$, its Fenchel conjugate is defined by

$$I^*(\xi) = \sup_{u\in X}(\langle \xi, u\rangle - I(u)).$$

Since $J|_K \geq J$ we automatically have $J|_K^* \leq J^*$, so it is enough to prove that $J|_K^*(\xi) \geq J^*(\xi)$ for all $\xi \in B_{X^*}(0, M/2)$.

Suppose, for contradiction, that there exists $\xi \in B_{X^*}(0, M/2)$ such that $J|_K^*(\xi) < J^*(\xi)$, which implies that

$$\begin{aligned}\langle \xi, u_0\rangle &\leq \sup_{u\in S}(\langle \xi, u\rangle - J(u))\\ &\leq \sup_{u\in K}(\langle \xi, u\rangle - J(u))\\ &< \sup_{u\in X}(\langle \xi, u\rangle - J(u)).\end{aligned}$$

Thus we can find $u \notin S$ such that $\langle \xi, u_0\rangle < \langle \xi, u\rangle - J(u)$, i.e.

$$RM < J(u) < \langle \xi, u - u_0\rangle\,. \tag{2.6}$$

There are two possible cases, i.e. $u \in B_X(0, R)\backslash S$ or $u \notin B_X(0, R)$.

For the first case, since $u \in B(0, R)$ and $u_0 \in B(0, \rho)$ we have $\|u - u_0\| \leq \rho + R$, hence

$$|\langle \xi, u - u_0\rangle| \leq (M/2)(\rho + R) < RM.$$

This contradicts (2.6). For the second case, since $\|u\| \geq R$ we have $|J(u)| \geq M\|u\|$. However, (2.6) would then imply that

$$M\|u\| \leq |J(u)| < |\langle \xi, u - u_0\rangle| \leq (M/2)(\rho + R) \leq M\|u\|\,,$$

which is again a contradiction. This shows that we indeed have $J^*(\xi) = J|_K^*(\xi)$ for all $\xi \in B_{X^*}(0, M/2)$.

For the second part, K is constructed for each given $M > 0$ so that it is a bounded closed convex set which is weakly compact. By Lindenstrauss–Troyanski's theorem (Theorem 2.24), K has the RNP. By Proposition 2.25, $J|_K^*$ is Fréchet differentiable on a dense G_δ subset of X^*. Since differentiability is a local property, the set of differentiability of J^* is completely determined by that of $J|_K^*$ by letting $M \to \infty$.

More precisely, for each $m > 0$ we can construct $K_m = \overline{\text{conv}}\,(\overline{S}_m^w)$ as before, and let D_m be the dense G_δ set determined by $J|_{K_m}^*$. One can verify that J^* is differentiable on

$$D = \bigcap_{m=1}^{\infty}(D_m \cup U_m)\,,$$

where $U_m = \{\xi \in X^* : \|\xi\| > m/4\}$. Since each $D_m \cup U_m$ is a dense G_δ set, D is also a dense G_δ set. This shows that J^* is generically differentiable. By Proposition 2.19, $J - \xi$ attains a strong minimum for each $\xi \in D$, hence the problem $\min(X; J - \xi)$ is generically Tykhonov well-posed.

The statement regarding Hadamard well-posedness follows directly from Fréchet differentiability (see the proof of Theorem 2.28). □

It is worth noting that if we impose the condition that the problem $\min(X; J - \xi)$ is Tykhonov well-posed for all $\xi \in X^*$ (or $\xi \in \mathrm{Dom}\,(\partial J^*)$), then this forces the functional J to be convex. This result follows from the famous Brøndsted–Rockafellar theorem and the sketch of proof can be found in [12].

The proof of Theorem 2.3 is a direct consequence of Theorem 2.32.

Proof of 2.3. Similarly to Theorem 2.2, we may consider the functional $J_g(u) := J(u - g)$ for $g \in \mathrm{W}^{1,1}(\Omega; \mathbb{R}^N)$ and observe that the growth condition (2.5) still holds for $F_g(x, u, z) := F(x, u - g(x), z - \nabla g(x))$. Indeed, since θ is convex and increasing, we have $\theta(|z|/2) \le \theta(|\nabla g(x)|/2 + |z - \nabla g(x)|/2) \le \frac{1}{2}(\theta(|z - \nabla g(x)|) + \theta(|z|))$, hence

$$\begin{aligned} F_g(x, u, z) &\ge h(x) + \theta(|z - \nabla g(x)|) \\ &\ge h(x) - \theta(|\nabla g(x)|) + 2\theta(|z|/2). \end{aligned}$$

The function $h - \theta(|\nabla g|)$ is integrable and $\widetilde{\theta}(t) := 2\theta(t/2)$ is convex and increasing. We can then apply the result from Theorem 2.32 to J_g. The rest of the proof is identical to the proof of Theorem 2.2 and so is left to the interested reader. □

3 Well-Posedness of Strictly Quasiconvex Variational Problems with Varying Dirichlet Data

The main goal of this chapter is to prove the generalized Hadamard well-posedness of the variational problem

$$\text{minimize } J(u) \quad \text{over} \quad u \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$$

with varying Dirichlet boundary data $g \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$. In this chapter we assume throughout that $\Omega \subset \mathbb{R}^n$ where $n \geq 2$ and $p \geq 2$. The fixed variational integral J is of the form

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

where the integrand $F\colon \Lambda \times \mathbb{R}^N \times \mathbb{R}^{N\times n} \to \mathbb{R}$ is *uniformly strictly p-quasiconvex* in the last argument ($F(x,u,\cdot) - \nu|\cdot|^p$ is quasiconvex for all x, u) and satisfies certain growth and coercivity conditions. To simplify some proofs in this chapter, we assume that $F(\cdot, u, z)$ is defined on a slightly larger (open, bounded) domain $\Lambda \subset \mathbb{R}^n$, where $\Omega \subset\subset \Lambda$. The generalized Hadamard well-posedness is established by showing that the solution map $\mathcal{S}\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \rightrightarrows \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$, which is a particular case of the multivalued functions

$$\mathcal{S}^\varepsilon(g) = \{u \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N) : J(u) \leq \varepsilon + \inf_{\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)} J\}$$

when $\varepsilon = 0$, is Upper semicontinuous (in fact, we will show that it is an USCO map). The result relies on a uniform integrability estimate for the family

$$\Big\{F(\cdot, u_j, \nabla u_j) : u_j \text{ minimizes } J \text{ in } \mathrm{W}^{1,p}_{g_j}(\Omega;\mathbb{R}^N)\Big\},$$

where g_j is a $\mathrm{W}^{1,p}$-convergent sequence, and the majority of the lemmas and propositions in this chapter is devoted to establishing this.

It is well-known that for a general quasiconvex variational problem, even for the simplest case where the integrand F of J only depends on the gradient the map, i.e. $J(u) = \int_\Omega F(\nabla u)\,dx$, the uniqueness of extremals and minimizers for J cannot be expected, so it is not possible to establish the classical Hadamard well-posedness for varying Dirichlet data. Indeed, for this simplified variational integral, Müller and Šverák [80] showed that the strong quasiconvexity of F (meaning $F(\cdot) - \nu|\cdot|^2$ is quasiconvex for some $\nu > 0$) is not sufficient to imply that the extremals of J are unique or even partially regular. Not even strong polyconvexity is enough to guarantee the uniqueness of extremals for J, as shown by Székelyhidi [96] (see also the work by Kristensen and Taheri [60]). The non-uniqueness of minimizers for strongly polyconvex variational integrals can be seen in the work [90] by Spadaro,

which is based on results of Lawson and Osserman [65] on minimal surfaces. This problem of non-uniqueness of Dirichlet problems will be addressed again in chapter 5.

While the conclusion of our main theorem relies on the uniform strict quasiconvexity of the integrand, most results in this chapter can be proved under significantly weaker assumptions. A function $G\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ is said to be *uniformly strictly p-quasiconvex* at $z \in \mathbb{R}^{N\times n}$ if

$$G(\cdot) - \nu|\cdot|^p \quad \text{is quasicovex at } z,$$

and G is uniformly strictly p-quasiconvex if $G(\cdot) - \nu|\cdot|^p$ is quasiconvex for all $z \in \mathbb{R}^{N\times n}$. This is a stronger condition than merely being quasiconvex, but in general it is weaker than the *strong quasiconvexity* condition that will be introduced in chapter 4 (uniform strict quasiconvexity and strong quasiconvexity coincides for $p = 2$, however). A variational integral $J\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ is *coercive* if

$$J(u) \to \infty \quad \text{if} \quad \|u\|_{\mathrm{W}^{1,p}} \to \infty. \tag{J1}$$

The main result regarding the semicontinuity properties of $\mathcal{S}$ and $\mathcal{S}^\varepsilon$ in this chapter is as follows.

Theorem 3.1. *Suppose that J is a coercive variational integral whose integrand F is Carathéodory and satisfies the growth conditions*

$$|F(x,u,z)| \le C\left(a(x) + b(x)|u|^\gamma + |z|^p\right) \tag{F1}$$

and that there exists a function $G\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ that is uniformly strictly p-quasiconvex at $z = 0$ such that

$$F(x,u,z) \ge G(z) - C\left(a(x) + b(x)|u|^\gamma\right), \tag{F2}$$

for some $C > 0$, $a \in \mathrm{L}^1(\Lambda)$, $b \in \mathrm{L}^\rho(\Lambda)$ where γ, ρ satisfies the conditions (E). Then for each $\varepsilon \ge 0$, the multivalued function $\mathcal{S}^\varepsilon\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \rightrightarrows \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ is $\|\cdot\|$-w-USCO.

Moreover, assuming that $F(x,u,\cdot)$ is uniformly strictly p-quasiconvex for all $(x,u) \in \Omega \times \mathbb{R}^N$, then $\mathcal{S}\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \rightrightarrows \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ is is $\|\cdot\|$-$\|\cdot\|$-USCO.

Theorem 3.1 is more or less a consequence of its sequential version.

Theorem 3.2. *Suppose that J is a coercive variational integral whose integrand F is Carathéodory, quasiconvex in the last variable, and satisfies the growth conditions (F1) and condition (F2) for some $C > 0$, $a \in \mathrm{L}^1(\Omega)$, $b \in \mathrm{L}^\rho(\Omega)$ where γ, ρ satisfies the conditions (E).*

Let $\varepsilon \ge 0$ be fixed, $g_j \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ be a sequence such that $g_j \to g$ in norm, and let $u_j \in \mathcal{S}^\varepsilon(g_j)$. Then there exists a subsequence g_{j_k} of g_j and $u \in \mathcal{S}^\varepsilon(g)$ such

that $u_{j_k} \overset{w}{\rightharpoonup} u$ *in* $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ *as* $k \to \infty$*, where* u_{j_k} *is the corresponding subsequence of* u_j*.*

Moreover, for $\varepsilon = 0$ *(i.e.* $\mathcal{S}^\varepsilon = \mathcal{S}$*), if we assume further that* $F(x,u,\cdot)$ *is uniformly strictly* p*-quasiconvex for all* $(x,u) \in \Omega \times \mathbb{R}^N$*, then we may conclude that* $u_{j_k} \to u$ *strongly in* $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ *as* $k \to \infty$*.*

A key ingredient for proving the uniform integrability result, which is essential in the proof of Theorem 3.2, is a Caccioppoli-type inequality that holds for both cubes in the interior of our domain Ω and the cubes near the boundary $\partial\Omega$. The technique used here is a variant of the so-called hole-filling argument by Widman [107], but with some extra difficulties that come from the fact that we need our Caccioppoli inequalities to hold uniformly for minimizers in different (but nearby) Dirichlet classes. The Caccioppoli inequalities state that for any minimizer $u \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$ of J, if the cube Q is sufficiently small then

$$\fint_{\alpha Q} |\nabla u|^p + |u|^q \, dx \le C \left(\fint_{\beta Q} (|\nabla u|^p + |u|^q)^m \, dx \right)^{1/m} + C \fint_{\beta Q} a(x) + b(x)^\rho + |\nabla g|^p + |g|^q \, dx.$$

Here α, β are some suitable numbers such that $1 \le \alpha < \beta \le 4$, $m = \frac{n}{n+p} < 1$, $q = \frac{\gamma\rho}{\rho-1} \in [p, p^*)$ and C blows up when $\beta - \alpha \to 0$.

Using a Gehring-type lemma, the above Caccioppoli inequalities are then used to prove a uniform integrability estimate for the family

$$\left\{ |\nabla u_j|^p : u_j \text{ minimizes } J \text{ in } \mathrm{W}^{1,p}_{g_j}(\Omega;\mathbb{R}^N) \right\}$$

whenever g_j is strongly $\mathrm{W}^{1,p}$ convergent. The proof is inspired by an estimate due to D'Onofrio and Iwaniec [34] on their work on the so-called p-harmonic transform. Under our growth conditions, this then implies the uniform integrability of $\{F(\cdot, u_j, \nabla u_j) : u_j \text{ minimizes } J \text{ in } \mathrm{W}^{1,p}_{g_j}(\Omega;\mathbb{R}^N)\}$ which allows us to utilize Young measures to deal with the issue of semicontinuity of J.

Theorem 3.3. *Let* $\Omega \subset \mathbb{R}^n$ *be a bounded Lipschitz domain,* J *be a coercive variational integral and assume that the growth condition* (F1)*, the coercivity condition* (F2) *hold for the integrand* $F\colon \Lambda \times \mathbb{R}^N \times \mathbb{R}^{N\times n} \to \mathbb{R}$*, where* Λ *is an open set such that* $\Omega \subset\subset \Lambda$*. For a sequence* $f_j \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$*, let* u_j *be a minimizer of the functional*

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

in $\mathrm{W}^{1,p}_{f_j}(\Omega;\mathbb{R}^N)$*, i.e.* $J(u_j) \le J(u)$ *for all* $u \in \mathrm{W}^{1,p}_{f_j}(\Omega;\mathbb{R}^N)$*.*

Suppose that $f_j \to f$ *in* $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$*, then the sequence* $(|\nabla u_j|^p)_{j\in\mathbb{N}}$ *is uniformly integrable in* Ω*. Moreover, the sequence* $(F(\cdot, u_j, \nabla u_j))_{j\in\mathbb{N}}$ *is also uniformly integrable in* Ω*.*

3.1 Caccioppoli Inequalities and Uniform Integrability

Before we go on to prove the Caccioppoli inequalities, we will need some results regarding dyadic cubes in $\mathbb{R}^n$. The definition of dyadic cubes that we will use is as follows.

Definition 3.4 (Dyadic cubes)**.** *For a given $k \in \mathbb{Z}$, define the* dyadic mesh of level *k on $\mathbb{R}^n$ to be*

$$\mathcal{M}_k := \{Q_{\mathbf{m}}^{(k)} : \mathbf{m} = (m_1, \ldots, m_n) \in \mathbb{Z}^n\},$$

where $Q_{\mathbf{m}}^{(k)}$ denotes the half-open cube

$$\begin{aligned} Q_{\mathbf{m}}^{(k)} &:= [0, 2^k)^n + 2^k \mathbf{m} \\ &= \{x \in \mathbb{R}^n : x_i \in [2^k m_i, 2^k(m_i + 1)) \text{ for each } i = 1, \ldots, n\}. \end{aligned}$$

The union of all such meshes $\mathcal{M} = \bigcup_{k\in\mathbb{Z}} \mathcal{M}_k$ we be called the dyadic mesh *on $\mathbb{R}^n$ and each of its element $Q \in \mathcal{M}$ will be called a* dyadic cube. //

The dyadic cubes are very useful because they satisfy the following properties.

1. For any $Q, Q' \in \mathcal{M}$, either $Q \cap Q' = \emptyset$ or one cube contains the other.
2. Any family $\mathcal{F} \subset \mathcal{M}$ has a subfamily $\mathcal{F}' \subset \mathcal{F}$ such that the cubes in $\mathcal{F}'$ are pairwise disjoint and
$$\bigcup_{Q\in\mathcal{F}'} Q = \bigcup_{Q\in\mathcal{F}} Q.$$
3. Each point $x \in \mathbb{R}^n$ has a unique sequence $Q^{(k)} \in \mathcal{M}_k$ such that
$$\cdots \subset Q^{(-2)} \subset Q^{(-1)} \subset Q^{(0)} \subset Q^{(1)} \subset Q^{(2)} \subset \ldots$$
and $\{x\} = \bigcap_{k\in\mathbb{Z}} Q^{(k)}$.

It is possible to select a family of disjoint dyadic cubes in a way that the average of a given function $f \in \mathrm{L}^1(\mathbb{R}^n)$ behaves nicely on each cube. This process is known as the Calderón–Zygmund decomposition lemma. Here we will need a variant of it, which can be proved very similarly. Denote the average of f on a measurable set E, with $0 < |E| = \mathcal{L}^n(E) < \infty$, by

$$\overline{f}_E := \fint_E f\,dx = \frac{1}{|E|}\int_E f\,dx.$$

Lemma 3.5 (A variant of Calderón–Zygmund decomposition lemma)**.** *Let $f \in \mathrm{L}^1(\mathbb{R}^n)$ be a positive ($f \geq 0$) function. Let $\rho > 0$ be given; then for any $t > \frac{1}{\rho^n}\|f\|_{\mathrm{L}^1(\mathbb{R}^n)}$, we can find a countable family $\mathcal{F}$ of disjoint dyadic cubes $Q \in \mathcal{M}$ such that*

1. *Each $Q \in \mathcal{F}$ has radius less than or equal to ρ.*

2. *$t < \overline{f}_Q \leq 2^n t$ for all $Q \in \mathcal{F}$.*

3. *$f(x) \leq t$ almost every $x \in \mathbb{R}^n \backslash \left(\bigcup_{Q \in \mathcal{F}} Q \right)$.*

(Sketch of proof). Choose $k \in \mathbb{Z}$ such that $r = 2^k \in (\rho/2, \rho]$. Then

$$\fint_{Q_r} f = \frac{1}{|Q_r|} \int_{Q_r} f \leq \frac{1}{(2r)^n} \|f\|_{\mathrm{L}^1} \leq \frac{1}{\rho^n} \|f\|_{\mathrm{L}^1} < t,$$

hence we have $\fint_{Q_r} f < t$ for any dyadic cube of radius $r < \rho$. We can then proceed to find a family of dyadic subcubes of $\{Q \in \mathcal{M} : Q \text{ has radius equal to } r\}$ that covers the set $\{f > t\}$ as in the proof of the standard Calderón–Zygmund lemma. □

In order to do some estimate near the boundary of Ω, we require that the boundary of the domain is sufficiently regular. The following condition, which holds for any Lipschitz domain, is essential in applying certain forms of the Poincaré inequality near the boundary of Ω (in fact a weaker capacity-type condition also suffices, see [75]).

Definition 3.6 (No Internal Cusp)**.** *Let $\Omega \subset \mathbb{R}^n$ be an open set. We say that Ω has no internal cusp if there exists $\rho > 0$ and $A \in (0, 1)$ such that for any cube $Q(x; r)$ centred at some $x \in \partial\Omega$, we have*

$$|Q(x; r) \backslash \Omega| \geq A|Q(x; r)| \tag{3.1}$$

for all $r \in (0, \rho)$. //

The next lemma allows us to get a similar lower bound on the size of $Q(x; r) \backslash \Omega$ even when x is not directly centred on $\partial\Omega$, provided that x is not too far away from $\partial\Omega$.

Lemma 3.7. *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with no internal cusp. Then there exists a constant $\beta > 0$ such that for any cube $Q(x; r)$ in $\mathbb{R}^n$ satisfying* dist $(x, \partial\Omega) \leq 2r$ *(here $d(x, y) = \max_{i=1,\dots,n} |x_i - y_i|$), we have*

$$|Q(x; \alpha r) \backslash \Omega| \geq \beta |Q(x; \alpha r)|$$

for all $\alpha \in [3, 4]$.

Proof. Let $Q(x; r)$ be a cube such that dist $(x, \partial\Omega) \leq 2r$ and let A, ρ be the constants as in the statement of Definition (3.6. Since Ω is a bounded domain, there exists a (closed) cube $\widehat{Q}$ such that $\Omega + Q(0; \rho) \subset \widehat{Q}$. There are three cases we need to consider:

(1) $Q(x; r) \cap \widehat{Q} = \emptyset$:

In this case, it is not hard to see that

$$|Q(x;\alpha r)\backslash\Omega| \geq |Q(x;r)| = \frac{1}{\alpha^n}|Q(x;ar)| \geq \frac{1}{4^n}|Q(x;ar)|$$

for all $\alpha \in [3,4]$.

(2) $Q(x;r) \subsetneq \widehat{Q}$:

By the assumption, there exists some i such that either $x + re_i$ or $x - re_i$ is not in $\widehat{Q}$ (e_i is the i^{th} standard basis vector). Without loss of generality, we assume that $x + re_1 \notin \widehat{Q}$. We now claim that, for any $\alpha \in [3,4]$,

$$Q(x + (2/3)\alpha r e_1; (1/4)\alpha r) \subset Q(x;\alpha r)\backslash\widehat{Q}.$$

Indeed, it is clear that $d(x,y) \leq (2/3)\alpha r + (1/4)\alpha r < \alpha r$ for all $y \in Q(x + (2/3)\alpha r e_1; (1/4)\alpha r)$, hence y is in $Q(x;\alpha r)$. On the other hand, we have

$$y_1 \geq x_1 + (2/3)\alpha r - (1/4)\alpha r > x_1 + (1/3)\alpha r > x_1 + r$$

hence $y \notin \widehat{Q}$. This implies that

$$|Q(x;\alpha r)\backslash\Omega| \geq |Q(x + (2/3)\alpha r e_1; (1/4)\alpha r)| = \frac{1}{4^n}|Q(x;\alpha r)|$$

whenever $\alpha \in [3,4]$.

(3) $Q(x;r) \subset \widehat{Q}$:

Let $\gamma = \frac{\rho}{4R} < \frac{1}{4}$. Let $y \in \partial\Omega$ be chosen so that $d(x,y) \leq 2r$. For any $\alpha \in [3,4]$, we have

$$Q(y;\alpha\gamma r) \subset Q(x;\alpha r)$$

since $\alpha\gamma < 1$ and thus for any $z \in Q(y;\alpha\gamma r)$,

$$d(x,z) \leq d(x,y) + d(y,z) \leq 2r + \alpha\gamma r < 3r \leq \alpha r.$$

Now, the cube $Q(y;\alpha\gamma r)$ is centred on $\partial\Omega$ and $\alpha\gamma r < \rho$ so we can use the 'no internal cusp' assumption to conclude that

$$|Q(x;\alpha r)\backslash\Omega| \geq |Q(y;\alpha\gamma r)\backslash\Omega| \geq A|Q(y;\alpha\gamma r)| = A\gamma^n|Q(x;\alpha r)|.$$

Lastly, we choose $a = \min\{1/4^n, A\gamma^n\}$ to conclude that for any $\alpha \in [3,4]$,

$$|Q(x;\alpha r)\backslash\Omega| \geq \beta|Q(x;\alpha r)|$$

in all three cases. □

Next, we recall the famous Vitali's covering lemma and the Poincaré inequality. The proofs can be found in the books mentioned below. Note that it is very important that in the Poincaré inequality, the constant does not depend on the size of the cube.

Lemma 3.8 (Vitali's covering lemma)**.** *Let $\mathcal{F}$ be a family of (open/closed/half-open) non-degenerated cubes in $\mathbb{R}^n$ such that $\sup_{Q\in\mathcal{F}} \operatorname{diam}(Q) < \infty$. There exists a subfamily $\mathcal{F}' \subset \mathcal{F}$ such that all the cubes in $\mathcal{F}'$ are pairwise-disjoint and*

$$\bigcup_{Q\in\mathcal{F}} Q \subset \bigcup_{Q\in\mathcal{F}'} 4Q.$$

Proof. See [53, Theorem 14.1.1]. □

Lemma 3.9 (Poincaré inequality)**.** *Let Q be a cube in $\mathbb{R}^n$. For every function $u \in \mathrm{W}^{1,p}(Q)$, it holds that*

$$\|u - \overline{u}_Q\|_{\mathrm{L}^{p^*}(Q)} \leq c(n,p) \|\nabla u\|_{\mathrm{L}^p(Q)}.$$

Moreover, if $E \subset Q$ is a measurable set such that $|E| > 0$ and $u \equiv 0$ on E, then we have

$$\|u\|_{\mathrm{L}^{p^*}(Q)} \leq c(n,p) \left(\frac{|Q|}{|E|}\right)^{1/p^*} \|\nabla u\|_{\mathrm{L}^p(Q)}.$$

Note that the constant $c(n,p)$ does not depend on the diameter of Q.

Proof. See [48, Theorem 3.17]. □

This Poincaré inequality is enough for our purpose. For more general inequalities of this type, see the monograph [75] by Maz'ya.

Lastly, we will need the following iteration lemma whose proofs can be found in the book [48] by Giusti.

Lemma 3.10 (Iteration lemma)**.** *Let $E\colon [r,R] \to [0,\infty)$ be a bounded function, $\delta \in [0,1)$; $A, B \geq 0$ and $\alpha > 0$. Suppose that*

$$E(t) \leq \delta E(s) + \left(\frac{A}{(s-t)^\alpha} + B\right)$$

for all s,t such that $r \leq t \leq s \leq R$. Then there exists a constant $C = C(\alpha,\delta)$ such that

$$E(t) \leq C\left(\frac{A}{(s-t)^\alpha} + B\right)$$

Proof. See [48, Lemma 6.1]. □

Now that we are done with stating preliminary results, the assumptions regarding the integrand $F(x,u,z)$ and the integral functional J are as follows.

Assumptions: Now, let us state the assumptions that shall be used throughout the rest of this section. Let $n \geq 2$ and $p \geq 2$, $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain (which, of course, implies that it has no internal cusp) and let Λ be an open set such that $\Omega \subset\subset \Lambda$. Let $F\colon \Lambda \times \mathbb{R}^N \times \mathbb{R}^{N\times n} \to \mathbb{R}$ be a Carathéodory integrand of the functional

$$J(u; E) = \int_E F(x, u(x), \nabla u(x))\, dx,$$

where E is a measurable subset of Λ. The following are the assumptions regarding F and J.

1. F satisfies the following growth condition:

$$|F(x, u, z)| \leq C(a(x) + b(x)|u|^\gamma + |z|^p) \tag{F1}$$

2. There exists a function $G\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ that is uniformly strictly quasiconvex at $z = 0$ such that

$$F(x, u, z) \geq G(z) - C(a(x) + b(x)|u|^\gamma) \tag{F2}$$

   for some $C \geq 0$. By *uniform strict quasiconvexity* at z_0, we mean that there exists $\nu > 0$ such that

$$G - \nu|\cdot|^p \quad \text{is } p\text{-quasiconvex at } z_0.$$

3. For each $x_0 \in \Omega$ and each $u_0 \in \mathbb{R}^N$, $F(x_0, u_0, \cdot)$ is quasiconvex, i.e.

$$\int_B F(x_0, u_0, z + \nabla\varphi)\, dx \geq \int_B F(x_0, u_0, z)\, dx \tag{F3}$$

   for all $z \in \mathbb{R}^{N\times n}$ and all $\varphi \in \mathrm{W}^{1,p}_0(B; \mathbb{R}^N)$.

4. J is coercive in the sense that

$$J(u; \Omega) \to \infty \quad \text{as} \quad \|u\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)} \to \infty. \tag{J1}$$

Here we assume $p \in [2, n)$, $\gamma \in [1, p^*)$, $a \in \mathrm{L}^1(\Lambda)$, $b \in \mathrm{L}^\rho(\Lambda)$, where ρ satisfies the conditions (with $q := \frac{\gamma\rho}{\rho-1}$)

$$\left.\begin{array}{l} 1.\ \rho > \frac{p^*}{p^*-\gamma}.\ \text{(so that } q < p^*\text{)} \\ 2.\ \rho \leq \frac{p}{p-\gamma} \text{ if } \gamma < p.\ \text{(so that } q \geq p\text{)} \\ 3.\ \rho \geq \frac{n-1}{\gamma n-\gamma-n} \text{ if } \gamma \leq \frac{n}{n-1}.\ \text{(so that } q_* \geq 1\text{)} \\ 4.\ \rho < \frac{n+p}{n+p-\gamma n} \text{ if } \gamma < \frac{n+p}{n}.\ \text{(so that } q \cdot \frac{n}{n+p} > 1\text{)} \end{array}\right\} \tag{E}$$

Remark 3.11. Here are some facts regarding the exponents p, γ, ρ, q that will be used later.

- Recall that for any $r \in [1, n)$ we write $r^* = \frac{nr}{n-r}$ and for any $r \in \left[\frac{n}{n-1}, \infty\right)$ we write $r_* = \frac{nr}{n+r}$. Observe that $(r_*)^* = r$ and $(r^*)_* = r$.
- By Hölder–Young inequality, we have $|ba^\gamma| \leq C(\varepsilon, \rho, \gamma)b^\rho + \varepsilon a^q$.
- We can discard condition 2 and 4 if we let $\gamma \in [p, p^*)$. In this case we can also discard condition 3 as well if we assume that $n \geq 3$ or $\gamma > 2$.
- All the conditions are satisfied if we let $\gamma \in [p, p^*)$ and $\rho = \infty$, e.g. $b(x)$ is a constant function.
- All the conditions are satisfied if we let $\gamma = p$ and $\rho > \frac{p^*}{p^*-p}$. //

Theorem 3.12 (Caccioppoli inequality (interior estimate))**.** *Assume that the conditions* (F1) *and* (F2) *are satisfied. Then there exists a constant* $C = C(p, n, \gamma, \rho, \Omega, \nu)$ *and a decreasing function* $\omega\colon [0, \infty) \to [0, \infty]$ *such that for any local minimizer* $u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$ *in* Ω *of the functional*

$$J(u; E) = \int_E F(x, u(x), \nabla u(x))\, dx,$$

i.e. $J(u; Q) \leq J(u + \varphi; Q)$ *for any* $\varphi \in \mathrm{W}^{1,p}_0(Q)$*, where* Q *is a cube compactly contained in* Ω*, and any cube* Q_R *such that* $Q_{2R} \subset\subset \Omega$ *and* $R \leq \omega(\|\nabla u\|_{\mathrm{L}^p(\Omega)})$*, we have*

$$\fint_{Q_R} |\nabla u(x)|^p + |u(x)|^q\, dx \leq C\left(\fint_{Q_{2R}} \left(|\nabla u(x)|^p + |u(x)|^q\right)^m\, dx\right)^{1/m} + C\fint_{Q_{2R}} a(x) + b(x)^\rho\, dx. \tag{3.2}$$

Here $m = \frac{n}{n+p} < 1$ *and* $q = \frac{\gamma\rho}{\rho-1}$ *is an exponent which satisfies* $q \in [p, p^*)$.

Proof. Let $u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$ be a local minimizer of $J(\cdot; \Omega)$, Q be a cube in Ω such that $2Q \subset\subset \Omega$. Let s, t be numbers such that $R \leq t < s \leq 2R$. Choose a smooth cut-off function $\eta \in C^\infty_c(Q_s)$ such that

$$\chi_{Q_t}(x) \leq \eta(x) \leq \chi_{Q_s}(x), \quad |\nabla \eta| \leq \frac{2}{s-t}$$

for all $x \in Q_s$. We shall write $\overline{u}_s = \overline{u}_{Q_s} = \fint_{Q_s} u\, dx$. Define

- $\widetilde{u}_t = \eta(u - \overline{u}_s)$.
- $v = u - \widetilde{u}_t = \overline{u}_s + (1 - \eta)(u - \overline{u}_s)$.

Without loss of generality, we may assume that $G(0) = 0$. By the uniform strict quasiconvexity of $G(z)$ at $z = 0$, we have

$$\begin{aligned}
\nu \int_{Q_s} |\nabla \widetilde{u}_t|^p \, dx &\leq \int_{Q_s} G(\nabla \widetilde{u}_t) \, dx \\
&\leq \int_{Q_s} F(x, u, \nabla \widetilde{u}_t) + C(a(x) + b(x)|u|^\gamma) \, dx \\
&= \underbrace{\int_{Q_s} F(x, u, \nabla u) \, dx}_{\text{(I)}} + \underbrace{\int_{Q_s} (F(x, u, \nabla \widetilde{u}_t) - F(x, u, \nabla u)) \, dx}_{\text{(II)}} \\
&\quad + \underbrace{C \int_{Q_s} a(x) + b(x)|u|^\gamma \, dx}_{\text{(III)}}.
\end{aligned}$$

Each term can be estimated as follows. For (I), since $J(u; Q_s) \leq J(u + \varphi; Q_s)$ for any $\varphi \in \mathrm{W}_0^{1,p}(Q_s)$, we can take $\varphi = -\widetilde{u}_t$ to get

$$\begin{aligned}
\text{(I)} = \int_{Q_s} F(x, u, \nabla u) \, dx &\leq \int_{Q_s} F(x, v, \nabla v) \, dx \\
&\leq C \int_{Q_s} a(x) + b(x)|v|^\gamma + |\nabla v|^p \, dx.
\end{aligned}$$

On the other hand, we observe that $\nabla \widetilde{u}_t = \nabla u$ on Q_t, thus

$$\begin{aligned}
\text{(II)} &\leq \left| \int_{Q_s \setminus Q_t} (F(x, u, \nabla \widetilde{u}_t) - F(x, u, \nabla u)) \, dx \right| \\
&\leq C \int_{Q_s \setminus Q_t} a(x) + b(x)|u|^\gamma + |\nabla \widetilde{u}_t|^p + |\nabla u|^p \, dx.
\end{aligned}$$

Putting all the above estimates together and the fact that $|\nabla \widetilde{u}_t|^p \lesssim |\nabla u|^p + |\nabla v|^p$, we have

$$\begin{aligned}
\int_{Q_t} |\nabla u|^p \, dx &\leq \frac{1}{\nu} (\text{(I)} + \text{(II)} + \text{(III)}) \\
&\lesssim \int_{Q_s} a(x) + b(x)|v|^\gamma + |\nabla v|^p \, dx \\
&\quad + \int_{Q_s \setminus Q_t} a(x) + b(x)|u|^\gamma + |\nabla \widetilde{u}|^p \, dx + |\nabla u|^p \, dx \\
&\quad + \int_{Q_s} a(x) + b(x)|u|^\gamma \, dx \\
&\lesssim \int_{Q_s \setminus Q_t} |\nabla u|^p \, dx + \int_{Q_s} a(x) + b(x)(|u|^\gamma + |v|^\gamma) + |\nabla v|^p \, dx.
\end{aligned}$$

Now, we observe that

- $|\nabla v|^p = |(1 - \eta)\nabla u - (u - \overline{u}_s)\nabla \eta|^p \lesssim |\nabla u|^p \chi_{Q_s \setminus Q_t} + \frac{1}{(s-t)^p} |u - \overline{u}_s|^p$.

- $|u| \le |u - \overline{u}_s| + |\overline{u}_s|$ and $|v| \le |u - \overline{u}_s| + |\overline{u}_s|$
- $b(x)(|u|^\gamma + |v|^\gamma) \lesssim b(x)^\rho + |u|^q + |v|^q$ by Hölder–Young inequality, where $q = \frac{\gamma\rho}{\rho - 1}$.

Thus

$$\int_{Q_t} |\nabla u|^p \, dx \lesssim \int_{Q_s \setminus Q_t} |\nabla u|^p \, dx + \frac{1}{(s-t)^p} \int_{Q_s} |u - \overline{u}_s|^p \, dx + \int_{Q_s} a(x) + b(x)^\rho + |\overline{u}_s|^q + |u - \overline{u}_s|^q \, dx.$$

We can add $\int_{Q_t} |u|^q$ to both sides (where it gets absorbed into $|\overline{u}_s|^q$ and $|u - \overline{u}_s|^q$ on the right-hand side) to get

$$\int_{Q_t} |\nabla u|^p + |u|^q \, dx \le C\Big(\int_{Q_s \setminus Q_t} |\nabla u|^p \, dx + \frac{1}{(s-t)^p} \underbrace{\int_{Q_s} |u - \overline{u}_s|^p \, dx}_{\text{(IV)}} + \underbrace{|Q_s||\overline{u}_s|^q}_{\text{(V)}} + \int_{Q_s} a(x) + b(x)^\rho \, dx + \underbrace{\int_{Q_s} |u - \overline{u}_s|^q \, dx}_{\text{(VI)}} \Big). \quad (3.3)$$

Until now, the constant C only depends on p, n, γ, ρ and ν.

For the term (IV), we have

$$\begin{aligned} \int_{Q_s} |u - \overline{u}_s|^p \, dx &\lesssim \int_{Q_s} |u - \overline{u}_{2R}|^p \, dx + \int_{Q_s} |\overline{u}_{2R} - \overline{u}_s|^p \, dx \\ &\le \int_{Q_{2R}} |u - \overline{u}_{2R}|^p \, dx + |Q_s| \left| \fint_{Q_s} \overline{u}_{2R} - u \, dx \right|^p \\ &\le \int_{Q_{2R}} |u - \overline{u}_{2R}|^p \, dx + |Q_s| \fint_{Q_s} |\overline{u}_{2R} - u|^p \, dx. \end{aligned}$$

Since $s \in [R, 2R]$, it follows that $|Q_s| \le |Q_{2R}|$ and $\fint_{Q_s} |\overline{u}_{2R} - u|^p \, dx \le 2^n \fint_{Q_{2R}} |\overline{u}_{2R} - u|^p \, dx$, hence we have

$$\int_{Q_s} |u - \overline{u}_s|^p \, dx \lesssim \int_{Q_{2R}} |u - \overline{u}_{2R}|^p \, dx.$$

The term (V) is easy to estimate since $|\overline{u}_s| \le 2^n |\overline{u}_{2R}|$, thus

$$|Q_s||\overline{u}_s|^q \le 2^n |Q_{2R}||\overline{u}_{2R}|^q.$$

By the Poincaré inequality (Lemma 3.9), we can estimate (VI) by

$$\begin{aligned}\int_{Q_s} |u-\overline{u}_s|^q\,dx &\le c(n,q)\left(\int_{Q_s}|\nabla u|^{q_*}\,dx\right)^{q/q_*}\\ &\le c|Q_s|^{\frac{q(p-q_*)}{pq_*}}\left(\int_{Q_s}|\nabla u|^p\,dx\right)^{\frac{q}{p}}\\ &\le cR^{\frac{q(p-q_*)}{pq_*}}\left(\int_{Q_s}|\nabla u|^p\,dx\right)^{\frac{q-p}{p}}\int_{Q_s}|\nabla u|^p\,dx\\ &\le cR^{\frac{q(p-q_*)}{pq_*}}\left(\|\nabla u\|_{\mathrm{L}^p(\Omega)}\right)^{q-p}\int_{Q_s}|\nabla u|^p\,dx, \qquad (3.4)\end{aligned}$$

where $q_* = \frac{nq}{n+q}$. Hence for any $\varepsilon > 0$, there exists $c_0(n,p,q,\varepsilon)$ such that

$$cR^{\frac{q(p-q_*)}{pq_*}}\left(\|\nabla u\|_{\mathrm{L}^p(\Omega)}\right)^{q-p} \le \varepsilon$$

if $R \le \omega(\|\nabla u\|_{\mathrm{L}^p(\Omega)})$, where

$$\omega(M) := c_0(n,p,q,\varepsilon)M^{-\frac{(q-p)pq_*}{q(p-q_*)}}. \qquad (3.5)$$

Fixed $\varepsilon = \frac{1}{2C}$, where C is the constant in (3.3). For $R \le \omega(\|\nabla u\|_{\mathrm{L}^p(\Omega)})$, we have

$$\begin{aligned}\int_{Q_t}|\nabla u|^p + |u|^q\,dx \le C\Big(&\int_{Q_s\setminus Q_t}|\nabla u|^p\,dx + \frac{1}{(s-t)^p}\int_{Q_{2R}}|u-\overline{u}_{2R}|^p\,dx\\ &+ |Q_{2R}|\overline{u}_{2R}|^q + \int_{Q_s}a(x)+b(x)^\rho\,dx + \varepsilon\int_{Q_s}|\nabla u|^p\,dx\Big)\end{aligned}$$

Since $\int_{Q_s}|\nabla u|^p\,dx = \int_{Q_s\setminus Q_t}|\nabla u|^p\,dx + \int_{Q_t}|\nabla u|^p\,dx$, the term $\varepsilon\int_{Q_t}|\nabla u|^p\,dx$ can be absorbed into the left-hand side, which leads to

$$\begin{aligned}\int_{Q_t}|\nabla u|^p + |u|^q\,dx \le C\Big(&\int_{Q_s\setminus Q_t}|\nabla u|^p + |u|^q\,dx + \frac{1}{(s-t)^p}\int_{Q_{2R}}|u-\overline{u}_{2R}|^p\,dx\\ &+ |Q_{2R}||\overline{u}_{2R}|^q + \int_{Q_{2R}}a(x)+b(x)^\rho\,dx\Big)\end{aligned}$$

Next we shall use the famous 'hole-filling trick', i.e. adding $C\int_{Q_t}|\nabla u|^p + |u|^q$ to both sides and divide by $(1+C)$ to get

$$\begin{aligned}\int_{Q_t}|\nabla u|^p + |u|^q\,dx \le \delta\int_{Q_s}|\nabla u|^p + |u|^q\,dx + \frac{1}{(s-t)^p}\int_{Q_{2R}}|u-\overline{u}_{2R}|^p\,dx\\ + |Q_{2R}||\overline{u}_{2R}|^q + \int_{Q_{2R}}a(x)+b(x)^\rho\,dx\end{aligned}$$

with $\delta = \frac{C}{1+C} < 1$. We can thus conclude, using the iteration lemma (Lemma 3.10) with $E(t) := \int_{Q_t} |\nabla u|^p + |u|^q \, dx$, that

$$\int_{Q_R} |\nabla u|^p + |u|^q \, dx \leq C\Big(\frac{1}{R^p} \underbrace{\int_{Q_{2R}} |u - \overline{u}_{2R}|^p \, dx}_{\text{(VII)}} + |Q_{2R}| \underbrace{|\overline{u}_{2R}|^q}_{\text{(VIII)}} + \int_{Q_{2R}} a(x) + b(x)^\rho \, dx \Big). \tag{3.6}$$

As before, observe that under our assumptions $p_* = \frac{np}{n+p} > 1$, hence we can use Poincaré inequality to estimate the term (VII) and get

$$\int_{2Q} |u - \overline{u}_{2R}|^p \, dx \leq c(n,p) \left(\int_{2Q} |\nabla u|^{p_*} \, dx \right)^{p/p_*} = c(n,p) \left(\int_{2Q} |\nabla u|^{pm} \, dx \right)^{1/m},$$

where $m := \frac{n}{n+p} < 1$. On the other hand, since $qm > 1$ we can estimate (VIII) by

$$|\overline{u}_{2R}|^q \leq \left(\fint_{Q_{2R}} |u| \, dx \right)^q \leq \left(\fint_{Q_{2R}} |u|^{qm} \, dx \right)^{1/m}.$$

Dividing both sides of (3.6) by $|Q_R|$ and substitute the above estimates gives

$$\begin{aligned} \fint_{Q_R} |\nabla u|^p + |u|^q \, dx &\lesssim \frac{1}{R^{n+p}} \left(\int_{2Q} |\nabla u|^{pm} \, dx \right)^{1/m} + \left(\fint_{Q_{2R}} |u|^{qm} \, dx \right)^{1/m} \\ &\quad + \fint_{Q_{2R}} a(x) + b(x)^\rho \, dx \\ &\leq C \left(\fint_{Q_{2R}} (|\nabla u|^p + |u|^q)^m \, dx \right)^{1/m} + C \fint_{Q_{2R}} a(x) + b(x)^\rho \, dx \end{aligned}$$

which holds for all $R \leq \omega(\|\nabla u\|_{\mathrm{L}^p(\Omega)})$ and the latest constant C only depends on $p, n, \gamma, \rho, \Omega, \nu$ but not on u or R. □

We have dealt with the estimate on cubes that lie sufficiently deep inside the domain Ω, i.e. all cubes Q such that $2Q \subset\subset \Omega$. Now we shall consider what happens near the boundary $\partial\Omega$. Recall the assumption that Ω is a Lipschitz domain, hence an extension domain, and Λ is an open set such that $\Omega \subset \Lambda$, thus there exists an extension operation $E : \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N) \to \mathrm{W}^{1,p}(\mathbb{R}^n; \mathbb{R}^N)$ that is bounded linear and $\operatorname{supp} Eu \subset \Lambda$ for all $u \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$.

In the next theorem, we shall sometimes use the same symbol to denote both an element of $\mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$ and its extension (with respect to this fixed extension operator E) in $\mathrm{W}^{1,p}(\mathbb{R}^n; \mathbb{R}^N)$ for simplicity. There is a lemma that we will be using to prove our theorem.

Lemma 3.13. *Suppose that $F\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ is rank-one convex and there exists $p \in [1,\infty)$ and $\lambda \geq 0$ such that*

$$|F(z)| \leq C\left(\lambda + |z|\right)^p,$$

then we have

$$|F(z) - F(w)| \leq C\left(\lambda + |z| + |w|\right)^{p-1}|z - w|$$

for all $z, w \in \mathbb{R}^{N\times n}$.

Proof. See [48, Lemma 5.2]. □

Theorem 3.14 (Caccioppoli inequality (boundary estimates))**.** *Assume that the conditions (F1) and (F2) are satisfied. Then there exists a constant $C = C(p, n, \gamma, \rho, \Omega, \nu)$ and a decreasing function $\omega\colon [0,\infty) \to [0,\infty]$ such that for any global minimizer $u \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$ of the functional*

$$J(u;\Omega) = \int_\Omega F(x, u(x), \nabla u(x))\,dx,$$

where $g \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^n)$, and any cube $Q_R = Q(x;R)$ such that $\operatorname{dist}(x,\partial\Omega) \leq 2R$ (in the l^∞ norm) and R small enough so that $Q_{4R} \subset \Lambda$ and $R \leq \omega\left(\|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^n)} + \|\nabla g\|_{\mathrm{L}^p(\mathbb{R}^n)}\right)$, we have

$$\begin{aligned}\fint_{Q_{3R}} |\nabla(u(x) - g(x))|^p + |u(x)|^q\,dx \leq C\left(\fint_{Q_{4R}} \left(|\nabla(u(x) - g(x))|^p + |u(x)|^q\right)^m dx\right)^{1/m} \\ + C\fint_{Q_{4R}} a(x) + b(x)^\rho + |\nabla g|^p + |g|^q\,dx.\end{aligned} \tag{3.7}$$

Here $m = \frac{n}{n+p} < 1$, $q = \frac{\gamma\rho}{\rho-1}$ is an exponent which satisfies $q \in [p, p^*)$, the function g above should be interpreted as its extension $g \in \mathrm{W}^{1,p}(\mathbb{R}^n;\mathbb{R}^N)$, and u is extended by defining $u \equiv g$ on $\Lambda\backslash\Omega$ (not by the extension operator).

In particular, we have

$$\begin{aligned}\fint_{Q_{3R}} |\nabla(u(x)|^p + |u(x)|^q\,dx \leq C\left(\fint_{Q_{4R}} \left(|\nabla u(x)|^p + |u(x)|^q\right)^m dx\right)^{1/m} \\ + C\fint_{Q_{4R}} a(x) + b(x)^\rho + |\nabla g|^p + |g|^q\,dx.\end{aligned} \tag{3.8}$$

Proof. The proof of this theorem is similar to Theorem 3.12 in many places. Let $u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ be a minimizer of $J(\cdot;\Omega)$, Q be a cube in Ω such that $\operatorname{dist}(x,\partial\Omega) \leq 2$ (in the l^∞ norm). Let R be small enough so that $Q_{4R} \subset \Lambda$ and let s, t be numbers satisfying $3R \leq t < s \leq 4R$. Choose a smooth cut-off function $\eta \in C^\infty_c(Q_s)$ such that

$$\chi_{Q_t}(x) \leq \eta(x) \leq \chi_{Q_s}(x), \quad |\nabla\eta| \leq \frac{2}{s-t}$$

for all $x \in Q_s$. Extend u outside Ω by letting $u := g$ on $\Lambda\backslash\Omega$ (so that $u - g \equiv 0$ on $\mathbb{R}^n\backslash\Omega$). Define

- $\widetilde{u}_t = \eta(u-g)$.
- $v = u - \widetilde{u}_t = g + (1-\eta)(u-g)$.

Without loss of generality, we may assume that $G(0) = 0$ as before. By the uniform strict quasiconvexity of $G(z)$ at $z = 0$, we have

$$\begin{aligned}
\nu \int_{Q_s} |\nabla \widetilde{u}_t|^p \, dx &\le \int_{Q_s} G(\nabla \widetilde{u}_t) \, dx \\
&\le \int_{Q_s} F(x,u,\nabla \widetilde{u}_t) + a(x) + b(x)|u|^\gamma \, dx \\
&= \underbrace{\int_{Q_s} F(x,u,\nabla u) \, dx}_{\text{(I)}} + \underbrace{\int_{Q_s} (F(x,u,\nabla \widetilde{u}_t) - F(x,u,\nabla u)) \, dx}_{\text{(II)}} \\
&\quad + \underbrace{\int_{Q_s} a(x) + b(x)|u|^\gamma \, dx}_{\text{(III)}}. \qquad (3.9)
\end{aligned}$$

For the term (I), observe that $\widetilde{u}_t \in \mathrm{W}^{1,p}_0(Q_s \cap \Omega; \mathbb{R}^n)$, so

$$\begin{aligned}
\int_{Q_s} F(x,u,\nabla u) \, dx &\le \int_{Q_s \cap \Omega} F(x, u - \widetilde{u}_t, \nabla(u - \widetilde{u}_t)) \, dx + \int_{Q_s \setminus \Omega} F(x, v - 0, \nabla(v-0)) \, dx \\
&= \int_{Q_s} F(x,v,\nabla v) \, dx \\
&\lesssim \int_{Q_s} a(x) + b(x)|v|^\gamma + |\nabla v|^p \, dx.
\end{aligned}$$

As for (II), we apply Lemma 3.13 and the Hölder–Young inequality to get

$$\begin{aligned}
\text{(II)} &\lesssim \int_{Q_s} \left((a(x) + b(x)|u|^\gamma)^{\frac{1}{p}} + |\nabla \widetilde{u}_t| + |\nabla u| \right)^{p-1} |\nabla \widetilde{u}_t - \nabla u| \, dx \\
&\lesssim \int_{Q_s} \left((a(x) + b(x)|u|^\gamma)^{\frac{p-1}{p}} \right) |\nabla v| + (|\nabla \widetilde{u}_t| + |\nabla u|)^{p-1} |\nabla v| \, dx \\
&\lesssim \int_{Q_s} a(x) + b(x)^\rho + |u|^q + |\nabla v|^p + \varepsilon |\nabla \widetilde{u}_t|^p + \varepsilon |\nabla u|^p + C(\varepsilon)|\nabla v|^p \, dx,
\end{aligned}$$

where $q = \frac{\gamma\rho}{\rho - 1}$. (III) and parts of (I) can be easily absorbed into (II).

For ε small enough, we can subtract $\varepsilon |\nabla \widetilde{u}_t|^p$ from the left-hand side of (3.9) to get

$$\int_{Q_t} |\nabla \widetilde{u}_t|^p \, dx \le \int_{Q_s} |\nabla \widetilde{u}_t|^p \, dx \le C \int_{Q_s} a(x) + b(x)^\rho + |u|^q + |\nabla v|^p + \varepsilon |\nabla u|^p \, dx.$$

Now, $|\nabla \widetilde{u}_t|^p = |\nabla(u-g)|^p$ on Q_t and $\varepsilon |\nabla u|^p \lesssim \varepsilon |\nabla(u-g)|^p + \varepsilon |\nabla g|^p$ so for small enough ε, we can absorb $\varepsilon |\nabla(u-g)|^p$ (on Q_t) into the left-hand side. Adding $|u|^q$

to both sides gives

$$\int_{Q_t} |\nabla(u-g)|^p + |u|^q \, dx \le C \int_{Q_s \setminus Q_t} |\nabla(u-g)|^p \, dx + \int_{Q_s} \underbrace{|u-g|^q}_{\text{(IV)}} + \underbrace{|\nabla v|^p}_{\text{(V)}} \, dx$$
$$+ C \int_{Q_s} a(x) + b(x)^\rho + |\nabla g|^p + |g|^q \, dx. \tag{3.10}$$

Since $s \in [3R, 4R]$, Lemma 3.7 says that the term (V) can be estimated by Poincaré's inequality (Lemma 3.9) as follows.

$$\begin{aligned}
\int_{Q_s} |u-g|^q \, dx &\le c(n,q) \left(\int_{Q_s} |\nabla(u-g)|^{q_*} \, dx \right)^{q/q_*} \\
&\le |Q_s|^{\frac{q(p-q_*)}{pq_*}} \left(\int_{Q_s} |\nabla(u-g)|^p \, dx \right)^{\frac{q}{p}} \\
&\le R^{\frac{q(p-q_*)}{pq_*}} \left(\int_{Q_s} |\nabla(u-g)|^p \, dx \right)^{\frac{q-p}{p}} \int_{Q_s} |\nabla(u-g)|^p \, dx \\
&\le R^{\frac{q(p-q_*)}{pq_*}} \left(\|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^n)} + \|\nabla g\|_{\mathrm{L}^p(\mathbb{R}^n)} \right)^{q-p} \int_{Q_s} |\nabla(u-g)|^p \, dx,
\end{aligned} \tag{3.11}$$

where $q_* = \frac{nq}{n+q}$. Hence for any $\varepsilon > 0$ (not the same ε we have used above in this proof), there exists $c_0(n,p,q,\varepsilon)$ such that

$$cR^{\frac{q(p-q_*)}{pq_*}} \left(\|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^n)} + \|\nabla g\|_{\mathrm{L}^p(\mathbb{R}^n)} \right)^{q-p} \le \varepsilon$$

if $R \le \omega \left(\|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^n)} + \|\nabla g\|_{\mathrm{L}^p(\mathbb{R}^n)} \right)$, where

$$\omega(M) = c_0(n,p,q,\varepsilon) M^{-\frac{(q-p)pq_*}{q(p-q_*)}} \tag{3.12}$$

is the same function as in Theorem 3.12 (but perhaps with a different constant $c_0(n,p,q,\varepsilon)$). As for (V), we have

$$\begin{aligned}
|\nabla v|^p &= |\nabla g + (1-\eta)\nabla(u-g) - (u-g)\nabla\eta|^p \\
&\lesssim |\nabla g|^p + |\nabla(u-g)|^p \chi_{Q_s \setminus Q_t} + \frac{1}{(s-t)^p} |u-g|^p.
\end{aligned}$$

Putting the estimates for (IV) and (V) together and choosing $\varepsilon > 0$ small enough so that (part of) (IV) can be absorbed into the left-hand side of (3.10); we get

$$\int_{Q_t} |\nabla(u-g)|^p + |u|^q \, dx \le C \int_{Q_s \setminus Q_t} |\nabla(u-g)|^p + |u|^q \, dx + \frac{1}{(s-t)^p} \int_{Q_s} |u-g|^p \, dx$$
$$+ C \int_{Q_s} a(x) + b(x)^\rho + |\nabla g|^p + |g|^q \, dx. \tag{3.13}$$

As in the proof of Theorem 3.12, we now use the hole-filling trick and the iteration lemma (Lemma 3.10), which gives

$$\int_{Q_{3R}} |\nabla(u-g)|^p + |u|^q \, dx \leq \frac{1}{R^p} \underbrace{\int_{Q_{4R}} |u-g|^p \, dx}_{\text{(VI)}} + C \int_{Q_{4R}} a(x) + b(x)^\rho + |\nabla g|^p + |g|^q \, dx. \tag{3.14}$$

We can use Poincaré's inequality to estimate the term (VI) and get

$$\int_{4Q} |u-g|^p \, dx \leq c(n,p) \left(\int_{4Q} |\nabla(u-g)|^{p_*} \, dx \right)^{p/p_*} = c(n,p) \left(\int_{4Q} |\nabla(u-g)|^{pm} \, dx \right)^{1/m},$$

where $m := \frac{n}{n+p} < 1$. We can artificially add $|u|^q$ to the term $|\nabla(u-g)|^p$ in the above and divide both sides of (3.14) after substituting the estimate for (VI) to get

$$\fint_{Q_{3R}} |\nabla(u-g)|^p + |u|^q \, dx \leq \left(\fint_{4Q} (|\nabla(u-g)|^p + |u|^q)^m \, dx \right)^{1/m} + C \fint_{Q_{4R}} a(x) + b(x)^\rho + |\nabla g|^p + |g|^q \, dx, \tag{3.15}$$

which concludes the proof. □

From Theorem 3.12 and Theorem 3.14, we can see that (assuming all the assumptions from both theorems) for all cubes $Q = Q_R$ with sufficiently small radius R, we have the estimate of the form

$$\fint_{\alpha Q} |\nabla u|^p + |u|^q \, dx \leq C \left(\fint_{\beta Q} (|\nabla u|^p + |u|^q)^m \, dx \right)^{1/m} + C \fint_{\beta Q} a(x) + b(x)^\rho + |\nabla f|^p + |f|^q \, dx, \tag{3.16}$$

for every $u \in \mathrm{W}^{1,p}_f(\Omega; \mathbb{R}^N)$ that minimize $J(v) = \int_\Omega F(x, v, \nabla v) \, dx$, where $\alpha = 1, \beta = 2$ if $2Q \subset\subset \Omega$ and $\alpha = 3, \beta = 4$ for Q near $\partial\Omega$. The estimates also holds trivially, with $\alpha = \beta = 1$, if $Q \cap \Omega = \emptyset$ (since $u = f$ on $\Lambda \backslash \Omega$). It is very important to note that the size of R depends only on $\|\nabla u\|_{\mathrm{L}^p}$ and $\|\nabla f\|_{\mathrm{L}^p}$, not on u itself.

Now, we may set

- $\lambda := \frac{1}{m} = \frac{n+p}{n} > 1$,
- $g := (|\nabla u|^p + |u|^q)^{\frac{1}{\lambda}}$,
- $h := (a(x) + b(x)^\rho + |\nabla f|^p + |f|^q)^{\frac{1}{\lambda}}$,

so that the above estimate can we written as

$$⨍_{\alpha Q} g^\lambda \, dx \leq C \left(⨍_{\beta Q} g \, dx \right)^\lambda + C ⨍_{\beta Q} h^\lambda \, dx$$

where $g, h \in \mathrm{L}^\lambda(\Omega)$ and $g, h \geq 0$. Alternatively, we can write

$$\left(⨍_{\alpha Q} g^\lambda \, dx \right)^{1/\lambda} \leq C ⨍_{\beta Q} g \, dx + C \left(⨍_{\beta Q} h^\lambda \, dx \right)^{1/\lambda}. \tag{3.17}$$

Here the constant C does not depend on g or h (or the radius R of Q).

Under the above assumptions, we can prove some form of equi-integrability of g. We shall need some simple inequality first.

Lemma 3.15. *Suppose that $\lambda > 1$. For any $x \geq 0$ and $t > 0$, we have*

$$x^{1/\lambda} \leq t + \frac{x}{t^{\lambda - 1}}. \tag{3.18}$$

Proof. Write $s = x^{1/\lambda}$, then the inequality we want to prove can be rewritten as

$$s \leq t + \frac{s^\lambda}{t^{\lambda-1}},$$

which is equivalent to

$$t^{\lambda-1} \leq \frac{t}{s} t^{\lambda-1} + s^{\lambda-1},$$

but this last inequality follows immediately from the fact that $f(x) = x^{\lambda-1}$ is monotonic on $[0, \infty)$. □

Proposition 3.16. *Let $\lambda > 1$ and assume that $g, h \in \mathrm{L}^\lambda(\mathbb{R}^n)$ are positive functions, i.e. $g, h \geq 0$. Suppose that there exists $\rho > 0$ such that for any cube Q with radius less that ρ, Q belongs to one of the families $\mathcal{F}_1, \mathcal{F}_2$ or $\mathcal{F}_3$; where each $Q^{(i)} \in \mathcal{F}_i$ satisfies an estimate of the form*

$$\left(⨍_{\alpha^{(i)} Q^{(i)}} g^\lambda \, dx \right)^{1/\lambda} \leq A ⨍_{\beta^{(i)} Q^{(i)}} g \, dx + \left(⨍_{\beta^{(i)} Q^{(i)}} h^\lambda \, dx \right)^{1/\lambda}, \tag{3.19}$$

with $1 \leq \alpha_i \leq \beta_i \leq 4$ and $A \geq 1$. Then

$$\int_{\{g > t\}} g^\lambda \, dx \leq C t^{\lambda-1} \int_{\{g>t\}} g \, dx + C \int_{\{h>t\}} h^\lambda \, dx \tag{3.20}$$

holds for all $t > \frac{1}{(2A+2)(4\rho)^{n/\lambda}} \|g\|_{\mathrm{L}^\lambda(\mathbb{R}^n)}$, where C is a constant depending only on n, λ and A.

Proof. Suppose that $s^\lambda > \frac{1}{\rho^n} \int_{\mathbb{R}^n} g^\lambda \, dx$, by the Calderón–Zygmund decomposition lemma (Lemma 3.5), we can find a countable family $\mathcal{F}$ of disjoint cubes such that for any $Q \in \mathcal{F}$

- Q has radius less that ρ.
- $s^\lambda \le \fint_Q g^\lambda \le 2^n s^\lambda$.
- $g^\lambda \le s^\lambda$ for almost every $x \in \mathbb{R}^n \setminus \bigcup_{Q\in\mathcal{F}} Q$.

Set $t := \frac{s}{4^{n/\lambda}(2A+2)}$, note that $s \le 4^{n/\lambda} 4At$.

Recall that according to our assumption, we must have $\mathcal{F} \subset \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$, hence

$$\{g > s\} \subset \bigcup_{Q\in\mathcal{F}} Q = \bigcup_{Q^{(1)}\in\mathcal{F}_1} Q^{(1)} \cup \bigcup_{Q^{(2)}\in\mathcal{F}_2} Q^{(2)} \cup \bigcup_{Q^{(3)}\in\mathcal{F}_3} Q^{(3)}.$$

Since the cubes are pairwise disjoint, we have

$$\int_{\{g>s\}} g^\lambda\,dx \le \int_{\bigcup_{Q\in\mathcal{F}} Q} g^\lambda\,dx = \sum_{i=1}^{\infty} \int_{Q_i} g^\lambda\,dx \le \sum_{i=1}^{\infty} \int_{Q_i} 2^n s^\lambda\,dx$$

i.e.

$$\int_{\{g>s\}} g^\lambda\,dx \le 2^n s^\lambda \left(\sum_{i=1}^{\infty} |Q_i|\right). \tag{3.21}$$

In order to estimate $|Q_i|$, suppose that $|Q_i|$ belongs to $\mathcal{F}_k$ and let α_i, β_i to be the constants $\alpha^{(k)}$ and $\beta^{(k)}$, respectively. We have

$$\begin{aligned}
s = (s^\lambda)^{1/\lambda} &\le \left(\fint_{Q_i} g^\lambda\,dx\right)^{1/\lambda} \\
&\le \left(\alpha_i^n \fint_{\alpha_i Q_i} g^\lambda\,dx\right)^{1/\lambda} \\
&\le \alpha_i^{n/\lambda} A \fint_{\beta_i Q_i} g\,dx + \alpha_i^{n/\lambda} \left(\fint_{\beta_i Q_i} h^\lambda\,dx\right)^{1/\lambda} \\
&\le \underbrace{4^{n/\lambda} A \fint_{\beta_i Q_i} g\,dx}_{\text{(I)}} + \underbrace{4^{n/\lambda} \left(\fint_{\beta_i Q_i} h^\lambda\,dx\right)^{1/\lambda}}_{\text{(II)}}
\end{aligned}$$

since $1 \le \alpha_i \le 4$. The term (I) can be estimated by

$$\begin{aligned}
\fint_{\beta_i Q_i} g\,dx &= \frac{1}{|\beta_i Q_i|}\left(\int_{\beta_i Q_i \cap \{g\le t\}} g\,dx + \int_{\beta_i Q_i \cap \{g>t\}} g\,dx\right) \\
&\le \frac{|\beta_i Q_i \cap \{g \le t\}|}{|\beta_i Q_i|} t + \frac{1}{|\beta_i Q_i|} \int_{\beta_i Q_i \cap \{g>t\}} g\,dx \\
&\le t + \frac{1}{|\beta_i Q_i|} \int_{\beta_i Q_i \cap \{g>t\}} g\,dx.
\end{aligned}$$

As for (II), we have

$$\begin{aligned}\left(\fint_{\beta_i Q_i} h^\lambda\,dx\right)^{1/\lambda} &= \left(\frac{1}{|\beta_i Q_i|}\int_{\beta_i Q_i\cap\{h\le t\}} h^\lambda\,dx + \frac{1}{|\beta_i Q_i|}\int_{\beta_i Q_i\cap\{h>t\}} h^\lambda\,dx\right)^{1/\lambda}\\ &\le \left(\frac{1}{|\beta_i Q_i\cap\{h\le t\}|}\int_{\beta_i Q_i\cap\{h\le t\}} h^\lambda\,dx + \frac{1}{|\beta_i Q_i|}\int_{\beta_i Q_i\cap\{h>t\}} h^\lambda\,dx\right)^{1/\lambda}\\ &\le t + \left(\frac{1}{|\beta_i Q_i|}\int_{\beta_i Q_i\cap\{h>t\}} h^\lambda\,dx\right)^{1/\lambda}\\ &\le t + \left(t + \frac{1}{t^{\lambda-1}}\frac{1}{|\beta_i Q_i|}\int_{\beta_i Q_i\cap\{h>t\}} h^\lambda\,dx\right),\end{aligned}$$

where the last line we used Lemma 3.15.

Substituting the bounds for (I) and (II) back, we get

$$\begin{aligned}4^{n/\lambda}(A+2)t = s \le \text{(I)} + \text{(II)} \le 4^{n/\lambda}At + \frac{4^{n/\lambda}A}{|\beta_i Q_i|}\int_{\beta_i Q_i\cap\{g>t\}} g\,dx\\ + 4^{n/\lambda}\left(2t + \frac{1}{t^{\lambda-1}}\frac{1}{|\beta_i Q_i|}\int_{\beta_i Q_i\cap\{h>t\}} h^\lambda\,dx\right),\end{aligned}$$

hence

$$At \le \frac{A}{|\beta_i Q_i|}\int_{\beta_i Q_i\cap\{g>t\}} g\,dx + \frac{1}{t^{\lambda-1}|\beta_i Q_i|}\int_{\beta_i Q_i\cap h>t} h^\lambda\,dx$$

which leads to the estimate for the size of $\beta_i Q_i$, i.e.

$$|\beta_i Q_i| \le \frac{1}{t}\int_{\beta_i Q_i\cap\{g>t\}} g\,dx + \frac{1}{At^\lambda}\int_{\beta_i Q_i\cap h>t} h^\lambda\,dx. \tag{3.22}$$

Now, the family $\mathcal{C} = \{\beta_i Q_i\}_{i\in\mathbb{N}}$ clearly covers the set $\bigcup_{i=1}^\infty Q_i$ that we started with since $\beta_i \ge 1$. By Vitali's covering lemma (Lemma 3.8), we can extract a subfamily $\mathcal{C}' = \{\beta_{i_k} Q_{i_k}\}_{k\in\mathbb{N}}$ of $\mathcal{C}$ such that

$$\bigcup_{i=1}^\infty \beta_i Q_i \subset \bigcup_{k=1}^\infty 4\beta_i Q_{i_k}$$

and the cubes in $\mathcal{C}'$ are pairwise disjoint. Thus

$$\begin{aligned}
\sum_{i=1}^{\infty} |Q_i| = \left|\bigcup_{i=1}^{\infty} Q_i\right| &\leq \left|\bigcup_{i=1}^{\infty} \beta_i Q_i\right| \\
&\leq \left|\bigcup_{k=1}^{\infty} 4\beta_{i_k} Q_{i_k}\right| \\
&= 4^n \left|\bigcup_{k=1}^{\infty} \beta_{i_k} Q_{i_k}\right| \\
&= 4^n \sum_{k=1}^{\infty} |\beta_{i_k} Q_{i_k}| \\
&\leq 4^n \sum_{k=1}^{\infty} \frac{1}{t} \int_{\beta_{i_k} Q_{i_k} \cap \{g>t\}} g\,dx + \frac{1}{At^\lambda} \int_{\beta_{i_k} Q_{i_k} \cap \{h>t\}} h^\lambda\,dx \\
&\leq \frac{4^n}{t} \int_{\{g>t\}} g\,dx + \frac{4^n}{At^\lambda} \int_{\{h>t\}} h^\lambda\,dx. \qquad (3.23)
\end{aligned}$$

We can now substitute the above estimate back into (3.21) (and use the fact that $s^\lambda \leq 4^{n+\lambda} A^\lambda t^\lambda$) to get

$$\begin{aligned}
\int_{\{g>s\}} g^\lambda &\leq 2^n s^\lambda \left(\sum_{i=1}^{\infty} |Q_i|\right) \\
&\leq \frac{8^n s^\lambda}{t} \int_{\{g>t\}} g\,dx + \frac{8^n s^\lambda}{At^\lambda} \int_{\{h>t\}} h^\lambda\,dx \\
&\leq 8^n 4^{n+\lambda} A^\lambda t^{\lambda-1} \int_{\{g>t\}} g\,dx + 8^n 4^{n+\lambda} A^{\lambda-1} \int_{\{h>t\}} h^\lambda\,dx.
\end{aligned}$$

On the other hand, we can easily see that

$$\int_{\{s\geq g>t\}} g^\lambda\,dx \leq s^{\lambda-1} \int_{\{s\geq g>t\}} g\,dx \leq 4^{n+\lambda-\frac{n}{\lambda}-1} A^{\lambda-1} t^{\lambda-1} \int_{\{g>t\}} g\,dx,$$

hence, by adding the two inequalities,

$$\int_{\{g>t\}} g^\lambda\,dx \leq C(n,\lambda,A) t^{\lambda-1} \int_{\{g>t\}} g\,dx + C(n,\lambda,A) \int_{\{h>t\}} h^\lambda\,dx.$$

This estimate holds for all $s > \frac{1}{\rho^{n/\lambda}} \|g\|_{\mathrm{L}^\lambda(\mathbb{R}^n)}$. Since $s = 4^{n/\lambda}(2A+2)t$, the estimate holds whenever $t > \frac{1}{(2A+2)(4\rho)^{n/\lambda}} \|g\|_{\mathrm{L}^\lambda(\mathbb{R}^n)}$. □

The estimate we get from this theorem will allow us to prove the uniform integrability result we want. Before proving the next proposition, let us recall some facts from measure theory.

For a measure space (Ω, m), a function $g\colon \Omega \to \Lambda$ induces a measure m_g on Λ called the *pushforward measure*, defined by $m_g(E) = g_{\#}m(E) := m(g^{-1}(E))$. The following *change of variables* formula holds for any $\Phi\colon \Lambda \to \mathbb{R}$ (assuming that Φ is integrable with respect to m_g or $\Phi \circ g$ is integrable with respect to m).

$$\int_\Lambda \Phi(s)\, dm_g(s) = \int_\Omega \Phi(g(x))\, dm(x).$$

Lemma 3.17. *Let μ, ν be measures on $\mathcal{B}([0,\infty))$, the Borel σ-algebra of $[0,\infty)$. Let (Ω, m) be a measure space where Ω is an open subset of $\mathbb{R}^n$.*

1. $\int_{(t,\infty)} \nu([0,s))\, d\mu(s) = \nu([0,t])\mu((t,\infty)) + \int_{(t,\infty)} \mu((s,\infty))\, d\nu(s)$.
2. $\int_{[0,t]} \nu((s,\infty))\, d\mu(s) = \nu((t,\infty))\mu([0,t]) + \int_{[0,t]} \mu([0,s))\, d\nu(s)$.
3. *Let $g\colon \Omega \to [0,\infty)$ be a measurable function and $\Phi\colon \mathbb{R} \to \mathbb{R}$ be a continuous function such that $\Phi \circ g \in \mathrm{L}^1(\Omega, m)$. Then for $\mu(E) = \int_{g^{-1}(E)} F(g)\, dm$, we have $F(s)\, dm_g(s) = d\mu(s)$, i.e. if we define $\tilde{\mu}(E) := \int_E F(s)\, dm_g(s)$, then $\tilde{\mu} = \mu$.*

Proof. 1. By Fubini–Tonelli theorem, we have

$$\begin{aligned}
\int_{(t,\infty)} \nu([0,s))\, d\mu(s) &= \int_{(t,\infty)} \nu([0,t]) + \nu((t,s))\, d\mu(s) \\
&= \nu([0,t])\mu((t,\infty)) + \int_{(t,\infty)} \left(\int_{(t,\infty)} \chi_{\{\tau<s\}}(s,\tau)\, d\nu(\tau) \right) d\mu(s) \\
&= \nu([0,t])\mu((t,\infty)) + \int_{(t,\infty)} \int_{(t,\infty)} \chi_{\{\tau<s\}}(s,\tau)\, d\mu(s)\, d\nu(\tau) \\
&= \nu([0,t])\mu((t,\infty)) + \int_{(t,\infty)} \mu((s,\infty))\, d\nu(s).
\end{aligned}$$

2. Similar proof as the above.

3. An application of the change of variables formula on $\Phi(s) = F(s)\chi_E(s)$ gives

$$\begin{aligned}
\tilde{\mu}(E) &= \int_{[0,\infty)} F(s)\chi_E(s)\, dm_g(s) \\
&= \int_\Omega F(g(x))\chi_E(g(x))\, dm(x) \\
&= \int_\Omega F(g)\chi_{g^{-1}(E)}\, dm \\
&= \int_{g^{-1}(E)} F(g)\, d\mathcal{L}^n = \mu(E),
\end{aligned}$$

hence $F(s)\, dm_g(s) = d\mu(s)$.

□

This next proposition was largely inspired by a result in [34] by D'Onofrio and Iwaniec. The proof is adapted from the proof of [34, Lemma 6.1].

Proposition 3.18. *Suppose that $g, h \in \mathrm{L}^\lambda(\mathbb{R}^n, \mathcal{L}^n)$ satisfy $g, h \geq 0$ and the relation*

$$\int_{\{g>t\}} g^\lambda \, d\mathcal{L}^n \leq C \left(t^{\lambda-1} \int_{\{g>t\}} g \, d\mathcal{L}^n + \int_{\{h>t\}} h^\lambda \, d\mathcal{L}^n \right) \tag{3.24}$$

for all $t > t_0 \geq 0$, $\lambda > 1$ and some $C > 1$. Then there exists some $C' = C'(n, \lambda, C), \alpha = \alpha(n, \lambda, C) > 0$ such that

$$\int_{\{g>t^2\}} g^\lambda \, d\mathcal{L}^n \leq C' \left(\int_{\{h>t\}} h^\lambda \, d\mathcal{L}^n + \frac{1}{t^\alpha} \int_\Omega h^\lambda \, d\mathcal{L}^n + \frac{t_0^\alpha}{t^{2\alpha}} \int_{\mathbb{R}^n} g^\lambda \, d\mathcal{L}^n \right) \tag{3.25}$$

for all $t > 0$.

Proof. Recall that $\mathcal{L}^n$ denotes the n-dimensional Lebesgue measure and ${\mathcal{L}^n}_g = g_\# \mathcal{L}^n$ is the pushforward of $\mathcal{L}^n$ by g. First we choose $\alpha > 0$ small enough so that

$$\frac{\alpha}{\alpha + \lambda - 1} \int_{\{g>t\}} g^\lambda \, d\mathcal{L}^n \leq t^{\lambda-1} \int_{\{g>t\}} g \, d\mathcal{L}^n + \int_{\{h>t\}} h^\lambda \, d\mathcal{L}^n \tag{3.26}$$

for all $t > t_0$. Define $G \colon (0, \infty) \to \mathbb{R}$ by

$$G(t) := t^{\lambda-1} \int_{\{g>t\}} g \, d\mathcal{L}^n.$$

For any $t > 0$, we have the relation

$$G(2t) \leq 2^\lambda \int_0^\infty \frac{G(s)}{s} \, ds. \tag{3.27}$$

Indeed, let $t > 0$ be fixed, then for any $s \in [t, 2t]$ we have

$$G(2t) = (2t)^{\lambda-1} \int_{\{g>2t\}} g \leq (2s)^{\lambda-1} \int_{\{g>s\}} g = 2^{\lambda-1} G(s),$$

hence

$$\int_t^{2t} \frac{G(2t)}{s} \, ds \leq 2^{\lambda-1} \int_t^{2t} \frac{G(s)}{s} \, ds \leq 2^{\lambda-1} \int_t^\infty \frac{G(s)}{s} \, ds.$$

However, the left-hand side is just $\ln(2) G(2t)$, thus

$$G(2t) \leq 2 \ln(2) G(2t) \leq 2^\lambda \int_t^\infty \frac{G(s)}{s} \, ds$$

as claimed.

Observe that

$$G(t) \leq \int_{\{g>t\}} g^{\lambda-1} g \, d\mathcal{L}^n = \int_{\{g>t\}} g^\lambda \, d\mathcal{L}^n \leq \|g\|_\lambda^\lambda,$$

hence $G(t)$ is a bounded function. Also, G is continuous at almost every $t > 0$ since the atomic part of $g^{-1}(\{c\})$ (with respect to the Lebesgue measure) is at most a countable set. Thus

$$A(t) := \int_t^\infty \frac{G(s)}{s}\, ds$$

is absolutely continuous and $A'(t) = -G(t)/t$ at almost every point $t > 0$.

From direct calculation,

$$\begin{aligned}\frac{d}{dt}(t^\alpha A(t)) &= \alpha t^{\alpha-1} A(t) - t^\alpha \frac{G(t)}{t} \\ &= t^{\alpha-1}\left(\alpha \int_t^\infty \frac{G(s)}{s}\, ds - G(t)\right).\end{aligned}$$

Now, we observe that

$$A(t) = \int_t^\infty \frac{G(s)}{s}\, ds = \int_t^\infty s^{\lambda-2}\left(\int_{\{g>s\}} g\, d\mathcal{L}^n\right) ds \le \frac{1}{\lambda - 1}\, \|g\|_\lambda^\lambda$$

for all $t > 0$ (last inequality is obtained by integrating by parts, i.e. Lemma 3.17 (1)), hence $t^\alpha A(t) \to 0$ as $t \to 0$. Since $t^\alpha A(t)$ is absolutely continuous on any closed interval in $(0, \infty)$, we have

$$\begin{aligned}T_1^\alpha A(T_1) - T_0^\alpha A(T_0) &= \int_{T_0}^{T_1} \frac{d}{dt}(t^\alpha A(t))\, dt \\ &= \int_{T_0}^{T_1} t^{\alpha-1}\left(\alpha \int_t^\infty \frac{G(s)}{s}\, ds - G(t)\right) dt.\end{aligned}$$

We may take $T_0 \to 0$ to get

$$T^\alpha A(T) = \int_0^T t^{\alpha-1}\left(\alpha \int_t^\infty \frac{G(s)}{s}\, ds - G(t)\right) dt, \tag{3.28}$$

i.e.

$$T^\alpha \int_T^\infty \frac{G(s)}{s}\, ds = \int_0^T t^{\alpha-1} \underbrace{\left(\alpha \int_t^\infty \frac{G(s)}{s}\, ds - G(t)\right)}_{(I)} dt. \tag{3.29}$$

The term (I) can be estimated using the assumption, as we will show below.

Define

$$\mu(E) = \int_{g^{-1}(E)} g\, d\mathcal{L}^n, \quad \nu(E) = \int_E (\lambda - 1) s^{\lambda-2} ds.$$

We can see that $\mu((t, \infty)) = \int_{\{g>t\}} g\, d\mathcal{L}^n$ and $\nu([0, s]) = s^{\lambda-1}$. By applying the Lemma 3.17 (1), we get

$$\int_t^\infty s^{\lambda-1}\, d\mu(s) = t^{\lambda-1} \int_{\{g>t\}} g\, d\mathcal{L}^n + \int_t^\infty (\lambda - 1) s^{\lambda-2}\left(\int_{\{g>s\}} g\, d\mathcal{L}^n\right) ds \tag{3.30}$$

which, together with the fact that $d\mu(s) = sd\mathcal{L}^n{}_g(s)$, yields

$$\begin{aligned}
\int_{\{g>t\}} g^\lambda \, d\mathcal{L}^n &= \int_{\{g>t\}} g \cdot g^{\lambda-1} \, d\mathcal{L}^n \\
&= \int_{(t,\infty)} s \cdot s^{\lambda-1} \, d\mathcal{L}^n{}_g(s) \\
&= \int_{(t,\infty)} s^{\lambda-1} \, d\mu(s) \\
&= t^{\lambda-1} \int_{\{g>t\}} g \, d\mathcal{L}^n + (\lambda-1) \int_t^\infty s^{\lambda-2} \left(\int_{\{g>s\}} g \, d\mathcal{L}^n \right) ds \\
&= G(t) + (\lambda-1) \int_t^\infty \frac{G(s)}{s} \, ds.
\end{aligned}$$

Now, substitute the above into (3.26), we get

$$\begin{aligned}
\frac{\alpha}{\alpha+\lambda-1} \left(G(t) + (\lambda-1) \int_t^\infty \frac{G(s)}{s} \, ds \right) &\leq G(t) + \int_{\{h>t\}} h^\lambda \, d\mathcal{L}^n \\
G(t) + (\lambda-1) \int_t^\infty \frac{G(s)}{s} \, ds &\leq \frac{\alpha+\lambda-1}{\alpha} G(t) + \frac{\alpha+\lambda-1}{\alpha} \int_{\{h>t\}} h^\lambda \, d\mathcal{L}^n \\
(\lambda-1) \int_t^\infty \frac{G(s)}{s} \, ds &\leq \frac{\lambda-1}{\alpha} G(t) + \frac{\alpha+\lambda-1}{\alpha} \int_{\{h>t\}} h^\lambda \, d\mathcal{L}^n \\
\alpha \int_t^\infty \frac{G(s)}{s} \, ds - G(t) &\leq \frac{\alpha+\lambda-1}{\lambda-1} \int_{\{h>t\}} h^\lambda \, d\mathcal{L}^n, \qquad (3.31)
\end{aligned}$$

which holds for all $t > t_0$. Using the estimate (3.31), the inequality (3.29) implies that

$$\begin{aligned}
T^\alpha \int_T^\infty \frac{G(s)}{s} \, ds &= \int_0^{t_0} t^{\alpha-1} \left(\alpha \int_t^\infty \frac{G(s)}{s} \, ds - G(t) \right) dt \\
&\quad + \int_{t_0}^T t^{\alpha-1} \left(\alpha \int_t^\infty \frac{G(s)}{s} \, ds - G(t) \right) dt \\
&\leq \int_0^{t_0} t^{\alpha-1} \left| \alpha \int_t^\infty \frac{G(s)}{s} \, ds - G(t) \right| dt \\
&\quad + \frac{\alpha+\lambda-1}{\lambda-1} \int_0^T t^{\alpha-1} \left(\int_{\{h>t\}} h^\lambda \right) dt
\end{aligned}$$

(which holds even for $T < t_0$).

Using the fact that $\left| \alpha \int_t^\infty \frac{G(s)}{s} \, ds \right| + |G(t)| \leq (\alpha/(\lambda-1)+1) \, \|g\|_\lambda^\lambda$ and since $(\alpha/(\lambda-1)+1) \int_0^t t^{\alpha-1} \, dt = c(\alpha,\lambda) t^\alpha$, we have

$$T^\alpha \int_T^\infty \frac{G(s)}{s} \, ds \leq c t_0^\alpha \, \|g\|_\lambda^\lambda + \frac{\alpha+\lambda-1}{\lambda-1} \int_0^T t^{\alpha-1} \left(\int_{\{h>t\}} h^\lambda \right) dt.$$

Setting $\nu(E) = \int_{\{h>t\}} h^\lambda \, d\mathcal{L}^n$ and $\mu(E) = \int_E t^{\alpha-1}\, dt$ (which gives $d\mu(t) = t^{\alpha-1}dt$), integration by parts (Lemma 3.17 (2)) gives

$$\begin{aligned} T^\alpha \int_T^\infty \frac{G(s)}{s}\, ds &\le ct_0^\alpha \|g\|_\lambda^\lambda + \frac{\alpha+\lambda-1}{\lambda-1}\left(\frac{T^\alpha}{\alpha}\int_{\{h>T\}} h^\lambda \, d\mathcal{L}^n + \int_{[0,T]} \frac{t^\alpha}{\alpha}\cdot t^\lambda \, d\mathcal{L}^n{}_h(t)\right) \\ &= ct_0^\alpha \|g\|_\lambda^\lambda + \frac{\alpha+\lambda-1}{\alpha(\lambda-1)}\left(T^\alpha \int_{\{h>T\}} h^\lambda \, d\mathcal{L}^n + \int_{\{h\le T\}} h^{\alpha+\lambda}\, d\mathcal{L}^n\right). \end{aligned}$$

Let $\varepsilon \in (0,1)$ be some fixed constant (to be determined later), the last term in the above inequality can be estimated by

$$\begin{aligned} \int_{\{h\le T\}} h^{\alpha+\lambda}\, d\mathcal{L}^n &= \int_{\{\varepsilon T<h\le T\}} h^{\alpha+\lambda}\, d\mathcal{L}^n + \int_{\{h\le \varepsilon T\}} h^{\alpha+\lambda}\, d\mathcal{L}^n \\ &\le T^\alpha \int_{\{h>\varepsilon T\}} h^\lambda \, d\mathcal{L}^n + (\varepsilon T)^\alpha \int_{\mathbb{R}^n} h^\lambda \, d\mathcal{L}^n. \end{aligned}$$

The term $\int_{\{h>T\}} h^\lambda$ can be absorbed into $\int_{\{h>\varepsilon T\}} h^\lambda$, hence

$$\int_T^\infty \frac{G(s)}{s}\, ds \le \frac{ct_0^\alpha}{T^\alpha}\|g\|_\lambda^\lambda + \frac{\alpha+\lambda-1}{\alpha(\lambda-1)}\left(2\int_{\{h>\varepsilon T\}} h^\lambda \, d\mathcal{L}^n + \varepsilon^\alpha \int_{\mathbb{R}^n} h^\lambda \, d\mathcal{L}^n\right). \quad (3.32)$$

Recall that the assumption (3.24) says that $\int_{\{g>t\}} g^\lambda \, d\mathcal{L}^n \le C(G(t)+\int_{\{h>t\}} h^\lambda \, d\mathcal{L}^n)$, so using 3.27 we get

$$\begin{aligned} \int_{\{g>2T\}} g^\lambda \, d\mathcal{L}^n &\le C\left(G(2T) + \int_{\{h>2T\}} h^\lambda \, d\mathcal{L}^n\right) \\ &\le C\left(\int_T^\infty \frac{G(s)}{s}\, ds + \int_{\{h>2T\}} h^\lambda \, d\mathcal{L}^n\right) \\ &\le C\left(\frac{t_0^\alpha}{T^\alpha}\|g\|_\lambda^\lambda + \int_{\{h>\varepsilon T\}} h^\lambda \, d\mathcal{L}^n + \varepsilon^\alpha \int_{\mathbb{R}^n} h^\lambda \, d\mathcal{L}^n\right), \end{aligned} \quad (3.33)$$

where we use (3.32) in the last line and absorb $\int_{\{h>2T\}} h^\lambda \, d\mathcal{L}^n$ into $\int_{\{h>\varepsilon T\}} h^\lambda \, d\mathcal{L}^n$.

For $t \ge 2$, $\varepsilon := 2/t < 1$ so we may take $T = t^2/2$ to get

$$\int_{\{g>t^2\}} g^\lambda \, d\mathcal{L}^n \le C\left(\int_{\{h>t\}} h^\lambda \, d\mathcal{L}^n + \frac{1}{t^\alpha}\int_{\mathbb{R}^n} h^\lambda \, d\mathcal{L}^n + \frac{t_0^\alpha}{t^{2\alpha}}\int_{\mathbb{R}^n} g^\lambda \, d\mathcal{L}^n\right),$$

as required.

Lastly, we observe that the above inequality holds for $t \le 2$ as well. Indeed, for $t \le 2$ the left-hand side is bounded above by $\|g\|_\lambda^\lambda$ whereas the right-hand side is bounded below by $\frac{C}{2^\alpha}\|h\|_\lambda^\lambda$. Thus by increasing the value of C, if necessary, the inequality will hold for $t \le 2$ as well. □

Putting together Proposition 3.16 and Proposition 3.18, we immediately get the following result.

Proposition 3.19. *Let $\lambda > 1$ and assume that $g, h \in \mathrm{L}^\lambda(\mathbb{R}^n)$ are positive functions. Suppose that there exists $\rho > 0$ such that for any cube Q with radius less that ρ, Q belongs to one of the families $\mathcal{F}_1, \mathcal{F}_2$ or $\mathcal{F}_3$, where each $Q^{(i)} \in \mathcal{F}_i$ satisfies an estimate of the form*

$$\left(\fint_{\alpha^{(i)}Q^{(i)}} g^\lambda \, dx\right)^{1/\lambda} \leq A \fint_{\beta^{(i)}Q^{(i)}} g \, dx + \left(\fint_{\beta^{(i)}Q^{(i)}} h^\lambda \, dx\right)^{1/\lambda}, \tag{3.34}$$

with $1 \leq \alpha_i \leq \beta_i \leq 4$ and $A \geq 1$. Then there exists some $C = C(n, \lambda, A), \alpha = \alpha(n, \lambda, A) > 0$ such that

$$\int_{\{g > t^2\}} g^\lambda \, dx \leq C \left(\int_{\{h > t\}} h^\lambda \, dx + \frac{1}{t^\alpha} \int_\Omega h^\lambda \, dx + \frac{t_0^\alpha}{t^{2\alpha}} \int_{\mathbb{R}^n} g^\lambda \, dx \right) \tag{3.35}$$

holds for all $t > 0$. Here $t_0 = \rho^{-n/\lambda} \|g\|_{\mathrm{L}^\lambda(\mathbb{R}^n)}$. Equivalently, we can also write

$$\int_{\{g > t^2\}} g^\lambda \, d\mathcal{L}^n \leq C \left(\int_{\{h > t\}} h^\lambda \, d\mathcal{L}^n + \frac{1}{t^\alpha} \|h\|_{\mathrm{L}^\lambda(\mathbb{R}^n)} + \frac{\rho^{-\alpha n/\lambda}}{t^{2\alpha}} \|g\|_{\mathrm{L}^\lambda(\mathbb{R}^n)}^{\alpha+\lambda} \right). \tag{3.36}$$

Proof. It is clear that, under the assumptions, Proposition 3.16 gives us the estimate

$$\int_{\{g > t\}} g^\lambda \, dx \leq C \left(t^{\lambda - 1} \int_{\{g > t\}} g \, dx + \int_{\{h > t\}} h^\lambda \, dx \right) \tag{3.37}$$

for all $t > t_0 = \frac{1}{(2A+2)(4\rho)^{n/\lambda}} \|g\|_{\mathrm{L}^\lambda(\mathbb{R}^n)}$. Proposition 3.19 can then be applied to get the final result. □

Using Proposition 3.19 and the Caccioppoli inequalities we proved earlier in Theorem 3.12 and Theorem 3.14, we arrive at our main uniform integrability result for minimizers of the integral functional J whose integrand satisfies assumptions (F1) and (F2).

Theorem 3.20. *Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain and assume that (F1) and (F2) hold for the integrand $F \colon \Lambda \times \mathbb{R}^N \times \mathbb{R}^{N \times n} \to \mathbb{R}$, where Λ is an open set such that $\Omega \subset\subset \Lambda$. Let $f \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$ be given and let $u \in \mathrm{W}_f^{1,p}(\Omega; \mathbb{R}^N)$ be a minimizer of the functional*

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

i.e. $J(u) \leq J(v)$ for any $v \in \mathrm{W}_f^{1,p}(\Omega; \mathbb{R}^N)$. Then there exists constants $C = C(p, n, \gamma, \rho, \Omega, \Lambda, \nu)$, $\beta = \beta(p, n, \gamma, \rho, \Omega, \Lambda, \nu)$ such that

$$\begin{aligned} \int_{\Omega \cap \{|\nabla u|^p(x) > t\}} |\nabla u|^p \, dx \leq C \int_{\{H_f > t^{1/2}\}} H_f(x) \, dx \\ + \frac{C}{t^\beta} \left(\|a + b^\rho\|_{\mathrm{L}^1(\Lambda)} + \mathcal{A}\left(\|f\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)} \right) \right) \\ + \frac{C}{t^{2\beta}} \mathcal{B}\left(\|u\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)} + \|f\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)} \right) \end{aligned}$$

holds for all $t > 0$. *Here* $\mathcal{A}, \mathcal{B}\colon [0,\infty) \to [0,\infty)$ *are two increasing functions such that* $\mathcal{A}(0) = \mathcal{B}(0) = 0$ *and* $H_f\colon \mathbb{R}^n \to [0,\infty)$ *is a function defined to be*

$$H_f(x) := a(x) + b(x)^\rho + |\nabla Ef(x)|^p + |Ef(x)|^q,$$

where $E\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \to \mathrm{W}^{1,p}(\mathbb{R}^n;\mathbb{R}^N)$ *is a Sobolev extension operator.*

Proof. Since u is a global minimizer, we can see that it must be a local minimizer in Ω in the sense that $J(u;Q) \le J(u+\varphi;Q)$ for any $\varphi \in \mathrm{W}^{1,p}_0(Q)$, where Q is a cube compactly contained in Ω. Hence u satisfies the all the assumptions of both Theorem 3.12 and Theorem 3.14. As before, we write

- $\lambda = \frac{n+p}{n} > 1$, $q = \frac{\gamma\rho}{\rho-1} \in [p, p^*)$,
- $g := (|\nabla u|^p + |u|^q)^{\frac{1}{\lambda}}$,
- $h := (a(x) + b(x)^\rho + |\nabla f|^p + |f|^q)^{\frac{1}{\lambda}}$,

where the functions g, h are considered as functions on $\mathbb{R}^n$. f is extended from Ω to $\mathbb{R}^n$ by an extension operator $E\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \to \mathrm{W}^{1,p}(\mathbb{R}^n;\mathbb{R}^N)$, and its support is contained in Λ. In fact, we can assume also that $f \equiv 0$ near $\partial\Lambda$. As for u, we define it to be $u = f$ on $\mathbb{R}^n\backslash\Omega$. We extend a, b by setting $a = b = 0$ on $\mathbb{R}^n\backslash\Lambda$.

Now, let us consider the estimate of the form

$$\left(\fint_{\alpha Q} g^\lambda\,dx\right)^{1/\lambda} \le C\fint_{\beta Q} g\,dx + C\left(\fint_{\beta Q} h^\lambda\,dx\right)^{1/\lambda} \tag{3.38}$$

for cubes in $\mathbb{R}^n$. For a given cube $Q = Q(x;R)$ such that $R \le \rho$ (with $\rho = \omega\left(\|\nabla u\|_{\mathrm{L}^p(\mathbb{R}^n)} + \|\nabla g\|_{\mathrm{L}^p(\mathbb{R}^n)}\right)$, ω be a function that works for both Theorem 3.12 and 3.14), there are 3 possible cases

1. $Q \cap \Omega = \emptyset$: It is clear that (3.38) holds with $\alpha = \beta = 1$ since $u = f$ on $\mathbb{R}^n\backslash\Omega$, hence $|\nabla u|^p + |u|^q = |\nabla f|^p + |f|^q$.
2. $Q \cap \Omega \ne \emptyset$ and $\operatorname{dist}(x, \partial\Omega) > 2R$: In this case, we have $2Q \subset\subset \Omega$, hence we can apply Theorem 3.12 to get the estimate (3.38) with $\alpha = 1, \beta = 2$.
3. $Q \cap \Omega \ne \emptyset$ but $\operatorname{dist}(x, \partial\Omega) \le 2R$: In this case, Q verifies the assumption of Theorem 3.14, so (3.38) holds with $\alpha = 3, \beta = 4$.

Any cube Q_R with $R \le \rho$ falls in one of the above categories. Denote the family of cubes that belong to each category by $\mathcal{F}_1, \mathcal{F}_2$ and $\mathcal{F}_3$, respectively. These families satisfy the conditions of Proposition 3.19, and hence we get the estimate

$$\int_{\{g>t^2\}} g^\lambda\,dx \le C\left(\int_{\{h>t\}} h^\lambda\,dx + \frac{1}{t^\alpha}\|h\|_{\mathrm{L}^\lambda(\mathbb{R}^n)} + \frac{\rho^{-\alpha n/\lambda}}{t^{2\alpha}}\|g\|^{\alpha+\lambda}_{\mathrm{L}^\lambda(\mathbb{R}^n)}\right). \tag{3.39}$$

The constant C here depends only on $n, \gamma, \rho, \Omega, \Lambda$. Expanding the above yields

$$\begin{aligned}\int_{\Omega\cap\{|\nabla u|^p>t^{2\lambda}\}} |\nabla u|^p\,dx &\le \int_{\{|\nabla u|^p+|u|^q>t^{2\lambda}\}} |\nabla u|^p + |u|^q\,dx \\ &\lesssim \int_{\{H_f>t^\lambda\}} H_f\,dx \\ &\quad + \frac{1}{t^\alpha}\left(\int_{\mathbb{R}^n} a(x) + b(x)^\rho + |\nabla f|^p + |f|^q\,dx\right)^{1/\lambda} \\ &\quad + \frac{\rho^{-\alpha n/\lambda}}{t^{2\alpha}}\left(\int_{\mathbb{R}^n} |\nabla u|^p + |u|^q\,dx\right)^{\alpha/\lambda+1}.\end{aligned}$$

where $H_f(x) := a(x) + b(x)^\rho + |\nabla f(x)|^p + |f(x)|^q$.

Since the Sobolev extension operator E is bounded, we have

$$\|f\|_{\mathrm{W}^{1,p}(\mathbb{R}^n)} \le C\,\|f\|_{\mathrm{W}^{1,p}(\Omega)}$$

and hence

$$\|u\|_{\mathrm{W}^{1,p}(\mathbb{R}^n)} \le \|u\|_{\mathrm{W}^{1,p}(\Omega)} + \|f\|_{\mathrm{W}^{1,p}(\mathbb{R}^n\setminus\Omega)} \le \|u\|_{\mathrm{W}^{1,p}(\Omega)} + C\,\|f\|_{\mathrm{W}^{1,p}(\Omega)}.$$

Using Sobolev (and Hölder) inequalities, it follows that

$$\begin{aligned}\int_{\Omega\cap\{|\nabla u|^p>t^{2\lambda}\}} |\nabla u|^p\,dx &\lesssim \int_{\{H_f>t^\lambda\}} H_f\,dx \\ &\quad + \frac{1}{t^\alpha}\left(\|a+b^\rho\|_{\mathrm{L}^1(\Lambda)} + \mathcal{A}\left(\|f\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)}\right)\right) \\ &\quad + \frac{1}{t^{2\alpha}}\mathcal{B}\left(\|u\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)} + \|f\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)}\right),\end{aligned}$$

where $\mathcal{A}(t) = C(t^p + t^q)^{\frac{1}{\lambda}}$ and $\mathcal{B}(t) = C\omega(t)^{-\alpha n/\lambda}(t^p + t^q)^{\frac{\alpha}{\lambda}+1}$ for some suitable constant C. It is clear that $\mathcal{A}(0) = \mathcal{B}(0) = 0$ and they are both increasing functions (since $\omega(t)$ is decreasing).

Lastly, we substitute $t = s^{\frac{1}{2\lambda}}$ and $\frac{\alpha}{2\lambda} = \beta$ to get

$$\begin{aligned}\int_{\Omega\cap\{|\nabla u|^p>s\}} |\nabla u|^p\,dx &\le C\int_{\{H_f>s^{1/2}\}} H_f\,dx \\ &\quad + \frac{C}{s^\beta}\left(\|a+b^\rho\|_{\mathrm{L}^1(\Lambda)} + \mathcal{A}\left(\|f\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)}\right)\right) \\ &\quad + \frac{C}{s^{2\beta}}\mathcal{B}\left(\|u\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)} + \|f\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)}\right),\end{aligned}$$

which holds for all $s > 0$. □

Finally, with Theorem 3.20, we can now prove the uniform integrability that we want.

Proof of Theorem 3.3. Suppose that each $u_j \in \mathrm{W}^{1,p}_{f_j}(\Omega;\mathbb{R}^N)$ is a minimizer of $J(u,\Omega)$ in $\mathrm{W}^{1,p}_{f_j}(\Omega;\mathbb{R}^N)$, where $f_j \to f$ in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$. Since u_j minimizes $J(u,\Omega)$ in $\mathrm{W}^{1,p}_{f_j}(\Omega;\mathbb{R}^N)$, we have

$$\begin{aligned}\int_\Omega F(x,u_j,\nabla u_j)\,dx &\le \int_\Omega F(x,f_j,\nabla f_j)\,dx\\ &\le C\int_\Omega a(x)+b(x)|f_j|^\gamma+|\nabla f_j|^p\,dx\\ &\le C\int_\Omega a(x)+b(x)^\rho+|f_j|^{\frac{\gamma\rho}{\rho-1}}+|\nabla f_j|^p\,dx\end{aligned}$$

which is bounded uniformly in $j\in\mathbb{N}$ since $\|f_j\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)}$ is bounded (and $\frac{\gamma\rho}{\rho-1}<p^*$). From assumption (J1), this shows that $\|u_j\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)}$ is also bounded, so we can find $M>0$ such that

$$\|u_j\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)}+\|f_j\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)}\le M\quad\text{for all}\quad j\in\mathbb{N}.$$

By Theorem 3.20, each u_j satisfies the estimate

$$\begin{aligned}\int_{\Omega\cap\{|\nabla u_j|^p(x)>t\}}|\nabla u_j|^p\,dx &\le C\int_{\{H_j>t^{1/2}\}}H_j(x)\,dx\\ &\quad+\frac{C}{t^\beta}\left(\|a+b^\rho\|_{\mathrm{L}^1(\Lambda)}+\mathcal{A}\left(\|f_j\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)}\right)\right)\\ &\quad+\frac{C}{t^{2\beta}}\mathcal{B}\left(\|u_j\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)}+\|f_j\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)}\right)\\ &\le C\int_{\{H_j>t^{1/2}\}}H_j(x)\,dx+\frac{C}{t^\beta}\left(\|a+b^\rho\|_{\mathrm{L}^1(\Lambda)}+\mathcal{A}(M)\right)\\ &\quad+\frac{C}{t^{2\beta}}\mathcal{B}(M)\end{aligned}\tag{3.40}$$

since $\mathcal{A}$ and $\mathcal{B}$ are both increasing. Here $H_j(x)=a(x)+b(x)^\rho+|\nabla Ef_j(x)|^p+|Ef_j(x)|^{\frac{\gamma\rho}{\rho-1}}$, where $E\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)\to\mathrm{W}^{1,p}(\mathbb{R}^n;\mathbb{R}^N)$ is a Sobolev extension operator.

Clearly, the last two terms in (3.40) tend to zero as $t\to\infty$ independent of j, so it suffices to show that

$$\limsup_{t\to\infty}\sup_{j\in\mathbb{N}}\int_{\{H_j>t^{1/2}\}}H_j(x)\,dx=0.\tag{3.41}$$

Indeed, since E is continuous, we have $Ef_j\to Ef$ in $\mathrm{W}^{1,p}(\mathbb{R}^n;\mathbb{R}^N)$ as well, hence $\nabla Ef_j\to\nabla Ef$ in $\mathrm{L}^p(\mathbb{R}^n;\mathbb{R}^N)$ and $Ef_j\to Ef$ in $\mathrm{L}^{\frac{\gamma\rho}{\rho-1}}(\mathbb{R}^n;\mathbb{R}^N)$ (by Sobolev inequality). According to Vitali's convergence theorem (Proposition A.11), this implies that $(|\nabla Ef_j(x)|^p)_{j\in\mathbb{N}}$ and $(|Ef_j(x)|^{\frac{\gamma\rho}{\rho-1}})_{j\in\mathbb{N}}$ are uniformly integrable, hence $(H_j)_{j\in\mathbb{N}}$ is uniformly integrable. Therefore (3.41) holds as was claimed. This implies that

$$\limsup_{t\to\infty}\sup_{j\in\mathbb{N}}\int_{\Omega\cap\{|\nabla u_j|^p(x)>t\}}|\nabla u_j|^p\,dx=0,\tag{3.42}$$

i.e. $(|\nabla u_j|^p)_{j\in\mathbb{N}}$ is uniformly integrable in Ω.

Lastly, since u_j is bounded in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$, the sequence $|u_j|^{\frac{\gamma\rho}{\rho-1}}$ is uniformly integrable by de la Vallée–Poussin's criterion (Proposition 2.31) because $\frac{\gamma\rho}{\rho-1} < p^*$ and u_j is bounded in $\mathrm{L}^{p^*}(\Omega;\mathbb{R}^N)$. Thanks to the inequality

$$|F(x,u_j,\nabla u_j)| \le C\left(a(x)+b(x)^\rho + |u_j|^{\frac{\gamma\rho}{\rho-1}} + |\nabla u_j|^p\right),$$

and because we have shown that the right-hand side is uniformly integrable, it follows that
$(F(x,u_j,\nabla u_j))_{j\in\mathbb{N}}$ is also uniformly integrable. □

3.2 Upper Semicontinuity of the Solution Map $\mathcal{S}$

Now we are ready to prove the continuity property of the solution map $\mathcal{S}\colon g \mapsto \inf_{\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)} J$.

Proposition 3.21. *Suppose that F is a Carathéodory integrand satisfying the conditions (F1) and (F2); then there exist $C = C(n,p,\rho,\gamma,\nu)$ and $\alpha = \alpha(n,p,\rho,\gamma,\nu) > 0$ such that*

$$\int_\Omega F(x,u,\nabla u)\,dx \ge \alpha\int_\Omega |\nabla u|^p\,dx - C\int_\Omega a(x)+b(x)^\rho + |u|^q + |\nabla g|^p\,dx \tag{3.43}$$

whenever $u \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$.

Moreover, $(F^-(x,u_j,\nabla u_j))_{j\in\mathbb{N}}$ is uniformly integrable whenever $u_j \in \mathrm{W}^{1,p}_{g_j}(\Omega;\mathbb{R}^N)$, $(|\nabla g_j|^p)_{j\in\mathbb{N}}$ is uniformly integrable, and $(u_j)_{j\in\mathbb{N}}$ is bounded in the $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ norm.

Proof. Let $g \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$, $u \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ and assume (without loss of generality) that the function G from (F2) satisfies $G(0) = 0$. Since $u - g \in \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^n)$, we have $\int_\Omega \nabla(u-g)\,dx = 0$ and hence

$$\begin{aligned}\nu\int_\Omega |\nabla(u-g)|^p\,dx &\le \int_\Omega G(\nabla(u-g))\\ &\le \int_\Omega F(x,u,\nabla(u-g)) + C(a(x)+b(x)|u|^\gamma)\,dx.\end{aligned}$$

Recall from Lemma 3.13 that

$$\begin{aligned}|F(x,u,\nabla(u-g)) - F(x,u,\nabla u)| &\le \left((a(x)+b(x)|u|^\gamma)^{1/p} + |\nabla(u-g)| + |\nabla u|\right)^{p-1}|\nabla g|\\ &\le C_\varepsilon(a(x)+b(x)^\rho + |u|^q + |\nabla g|^p + \varepsilon|\nabla u|^p),\end{aligned}$$

we have used the Hölder–Young inequality in the last line. Together with the fact that

$$\int_\Omega |\nabla u|^p\,dx \le 2^{p-1}\int_\Omega |\nabla(u-g)|^p\,dx + 2^{p-1}\int_\Omega |\nabla g|^p\,dx,$$

we can pick ε small enough so that we have

$$\frac{\nu}{2^p}\int_\Omega |\nabla u|^p\,dx \le \int_\Omega F(x,u,\nabla u)\,dx + C\int_\Omega a(x)+b(x)^\rho+|u|^q+|\nabla g|^p\,dx$$

This proves the estimate we want.

As for the last part, we observe that if $(u_j)_{j\in\mathbb{N}}$ is bounded in $\mathrm{W}^{1,p}$ norm, the $(|u_j|^q)_{j\in\mathbb{N}}$ is uniformly integrable by de la Vallée–Poussin criterion (Proposition 2.31) since $q<p^*$. If we assume that $(|\nabla g_j|^p)_{j\in\mathbb{N}}$ is uniformly integrable as well, then the sequence

$$a(x)+b(x)^\rho+|u_j|^q+|\nabla g_j|^p$$

is uniformly integrable. From (3.43), we can deduce that

$$\int_\Omega F^-(x,u,\nabla u)\,dx \le C\int_\Omega a(x)+b(x)^\rho+|u|^{\frac{\gamma\rho}{\rho-1}}+|\nabla g|^p\,dx$$

whenever $u\in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$; hence it follows that $(F^-(x,u_j,\nabla u_j))_{j\in\mathbb{N}}$ is uniformly integrable under the given assumptions. □

Corollary 3.22. *Suppose that F is a Carathéodory integrand satisfying the conditions (F1), (F2) and (F3). For any fixed $g\in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$, the functional $J(\cdot;\Omega)$ restricted to the affine subspace $\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$ is swlsc. Moreover, if we assume that the condition (J1) holds as well, then the problem $\min(\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N);J)$ has at least one minimizer.*

Proof. We apply Proposition 3.21 with the constant sequence $g_j=g$ (which is clearly uniformly integrable) and any sequence $u_j\in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$ such that $u_j \overset{w}{\rightharpoonup} u$ in $\mathrm{W}^{1,p}$ (which implies boundedness in norm) to deduce that $F^-(x,u_j,\nabla u_j)$ is uniformly integrable. By Proposition 1.46, we have

$$\int_\Omega F(x,u,\nabla u)\,dx \le \liminf_{n\to\mathbb{N}}\int_\Omega F(x,u_j,\nabla u_j)\,dx,$$

which is exactly the sequential weak lower semicontinuity that we want. Note that the weak limit u is in $\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$ since the trace operator $\mathrm{Tr}\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)\to \mathrm{L}^p(\partial\Omega)$ is weakly continuous.

Suppose we assume that (J1) holds; then the sublevel sets of J are bounded in the $\mathrm{W}^{1,p}$ norm, and hence they are (sequentially) weakly precompact. This shows that J is coercive in the sense of Proposition 1.17 (1). Together which the fact that J is swlsc, Proposition 1.17 can be applied to yield a minimizer of J in $\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$. □

Quasiconvexity in the last variable of $F(x,u,z)$ has been sufficient for our results so far, but to get a stronger conclusion regarding the convergence of minimizers we shall need to strengthen the quasiconvexity assumption. There are many different types of 'strict quasiconvexity' in the literature which are not equivalent to one another. Our notion of strict convexity is defined as follows.

Assumptions: Recall our standing assumptions that $n \geq 2$, $p \geq 2$, $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain and let Λ be an open set such that $\Omega \subset\subset \Lambda$. The integrand $F\colon \Lambda \times \mathbb{R}^N \times \mathbb{R}^{N\times n} \to \mathbb{R}$ is assumed to be Carathéodory. In the following, $B = B(0;1)$ denotes the open unit ball in $\mathbb{R}^n$.

Definition 3.23 (Uniform Strict p-Quasiconvexity)**.** *We shall say that the function F is uniformly strict p-quasiconvex if there exists a $\kappa > 0$ such that for each $x_0 \in \Omega$, $u_0 \in \mathbb{R}^N$, the function $F(x_0, u_0, z) - \kappa|z|^p$ is quasiconvex in the last variable, i.e.*

$$\int_B F(x_0, u_0, z + \nabla\varphi) - F(x_0, u_0, z)\, dx \geq \kappa \int_B |z + \nabla\varphi|^p - |z|^p\, dx \tag{F3'}$$

for all $z \in \mathbb{R}^{N\times n}$ and all $\varphi \in \mathrm{W}^{1,p}_0(B;\mathbb{R}^N)$. //

Remark 3.24. The open unit ball B in the above definition can be replaced by a bounded open set $U \subset \mathbb{R}^n$ satisfying $|\partial U| = 0$. In particular, we can replace B with our domain Ω. //

Definition 3.25 (Solution Map)**.** *For each $\varepsilon \geq 0$, we define the multivalued function $\mathcal{S}^\varepsilon\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \rightrightarrows \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ to be*

$$\begin{aligned}\mathcal{S}^\varepsilon(g) &:= \varepsilon\text{-argmin}\,(\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N); J)\\ &= \{u \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N) : J(u) \leq \varepsilon + \inf_{\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)} J\},\end{aligned}$$

i.e. $\mathcal{S}^\varepsilon(g)$ is the set of all 'ε-almost minimizers' of J in $\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$. In particular, if $\varepsilon = 0$ we simply write

$$\mathcal{S}(g) := \mathrm{argmin}\,(\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N); J).$$

The multivalued function $\mathcal{S}$ is called the solution map *since it sends g to the set of all minimizers of J in $\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$.* //

In the proof of the following theorem where we utilize Young measures, it is worth noting that a more direct proof is also possible for similar situations when boundary condition is not an issue (see [42] by Evans and Gariepy). However, adapting such technique to account for the boundary conditions is not a trivial task, and the approach using Young measures is more illuminating.

Proof of Theorem 3.2. Let F be a Carathéodory integrand satisfying the conditions (F1), (F2), (F3) and (J1), $g_j \to g$ in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$, and $u_j \in \mathcal{S}^\varepsilon(g_j)$. Recall from Corollary 3.22 that $\mathcal{S}(\cdot)$ is never empty-valued.

Observe that since $(J(g_j))_{j\in\mathbb{N}}$ is bounded, (J1) implies that the sequence u_j is also bounded, hence we can extract a subsequence such that $u_{j_k} \overset{w}{\rightharpoonup} u$ for some $u \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$ (by weak continuity of the trace). We want to show that $u \in \mathcal{S}^\varepsilon(g)$.

Pick an arbitrary $v \in \mathcal{S}(g)$. It is easy to see that $(v-g)+g_j \in \mathrm{W}^{1,p}_{g_j}(\Omega;\mathbb{R}^N)$, hence

$$J(u_j) \le J(v+(g_j-g))+\varepsilon$$

by definition. The integrand F of the integral functional J satisfies the condition of Corollary 1.22; hence J is continuous in the norm topology of $\mathrm{W}^{1,p}$. As $n\to\infty$, $g_j \to g$ in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ and thus

$$\begin{aligned}\limsup_{n\to\infty} J(u_j) &\le \lim_{n\to\infty} J(v+(g_j-g))+\varepsilon \\ &= J(v)+\varepsilon \\ &= \inf_{\mathrm{W}^{1,p}_g} J+\varepsilon.\end{aligned}$$

On the other hand, Proposition 3.21 says that $(F^-(x,u_j,\nabla u_j))_{j\in\mathbb{N}}$ is uniformly integrable (because $(|\nabla g_j|^p)_{j\in\mathbb{N}}$ is uniformly integrable by Vitali's convergence theorem), so we can apply Proposition 1.46 to conclude that $J(u) \le \liminf_{n\to\infty} J(u_j)$. Thus

$$J(u) \le \liminf_{n\to\infty} J(u_j) \le \limsup_{n\to\infty} J(u_j) \le \inf_{\mathrm{W}^{1,p}_g} J+\varepsilon,$$

which implies that $u \in \mathcal{S}^\varepsilon(g)$.

Now, let us take $\varepsilon = 0$ and assume that (F3') holds, $g_j \to g$ in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$, and $u_j \in \mathcal{S}(g_j)$. Since u_j is bounded in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$, we may extract a subsequence (not relabelled) such that $u_j \overset{w}{\rightharpoonup} u$ in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$, $u_j \to u$ a.e. $x\in\Omega$, and $\nabla u_j \overset{\mathrm{Y}}{\to} \nu$ for some gradient Young measure $\nu \in \mathrm{GY}^p(\Omega;\mathbb{R}^{N\times n})$. From Theorem 3.3, the sequences $(|\nabla u_j|^p)_{j\in\mathbb{N}}$ and $(F(x,u_j,\nabla u_j))_{j\in\mathbb{N}}$ are uniformly integrable. Hence

$$\begin{aligned}\int_\Omega F(x,u(x),\nabla u(x))-\kappa|\nabla u(x)|^p\,dx &\le \int_\Omega \langle \nu_x, F(x,u(x),\cdot)-\kappa|\cdot|^p\rangle\,dx \\ &= \lim_{n\to\infty}\int_\Omega F(x,u_j(x),\nabla u_j(x))-\kappa\int_\Omega\langle\nu_x,|\cdot|^p\rangle\,dx \\ &\le \inf_{\mathrm{W}^{1,p}_g} J-\kappa\int_\Omega\langle\nu_x,|\cdot|^p\rangle\,dx,\end{aligned}$$

which implies that

$$\kappa\int_\Omega \langle\nu_x,|\cdot|^p\rangle - |\overline{\nu}_x|^p\,dx \le \inf_{\mathrm{W}^{1,p}_g} J-\int_\Omega F(x,u(x),\nabla u(x))\,dx = 0 \tag{3.44}$$

since we know from the first part of this theorem that $J(u)=\inf_{\mathrm{W}^{1,p}_g} J$.

For any $x \in \Omega$, Jensen's inequality shows that

$$|\overline{\nu}_x|^p \leq \int_{\mathbb{R}^{N\times n}} |z|^p \, d\nu_x(z) = \langle \nu_x, |\cdot|^p \rangle ,$$

so (3.44) implies that Jensen's inequality actually attains equality for almost every $x \in \Omega$, i.e. $\int_{\mathbb{R}^{N\times n}} |z|^p \, d\nu_x(z) = |\overline{\nu}_x|^p$ almost every x. However, the function $t \mapsto |t|^p$ is strictly convex, hence the probability measure ν_x degenerates to $\nu_x = \delta_{\nabla u(x)}$ for almost every $x \in \Omega$. By Proposition 1.39, we thus have $\nabla u_j \to \nabla u$ in measure.

Since the sequence $(|\nabla u_j|^p)_{j\in\mathbb{N}}$ is uniformly integrable and $\nabla u_j \to \nabla u$ in measure, Vitali's convergence theorem (Proposition A.11) implies that $\nabla u_j \to \nabla u$ in the L^p norm, thus $u_j \to u$ in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$. This concludes the proof. □

Recall that a multivalued function $T\colon (X,\tau) \rightrightarrows (Y,\tau')$ is τ-τ'-USC if for each $x \in X$ and any $U \in \mathcal{N}(T(x))$, there exists a $W_U \in \mathcal{N}(x)$ such that $T(W_U) \subset U$. Here $\mathcal{N}(x)$ denotes the family of all neighbourhoods of x (similarly for $\mathcal{N}(T(x))$).

Proof of Theorem 3.1. Let $\varepsilon \geq 0$ be fixed. The part about $\mathcal{S}^\varepsilon$ having compact value follows directly from Theorem 3.2 by taking g_j to be the constant sequence $g_j \equiv g \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$. So it is enough to show that $\mathcal{S}^\varepsilon$ (resp. $\mathcal{S}$) is $\|\cdot\|$-w-USC (resp. $\|\cdot\|$-$\|\cdot\|$-USC).

We assume for contradiction that $\mathcal{S}^\varepsilon$ is not $\|\cdot\|$-w-USC, i.e. there exists a g and a weak neighbourhood $U \in \mathcal{N}(\mathcal{S}^\varepsilon(g))$ such that for any neighbourhood $W \in \mathcal{N}(g)$, we can find $g' \in \mathcal{N}(g)$ and $u' \in \mathcal{S}^\varepsilon(g')\backslash U$. Hence there exists a sequence $g_j \to g$ and a corresponding sequence u_j satisfying $u_j \in \mathcal{S}^\varepsilon(g_j)\backslash U$.

From the growth condition (F1) (or the continuity of J in the norm topology), J is bounded near g; hence for any $W \in \mathcal{N}(g)$ there exists an $M > 0$ such that

$$\sup_{u'\in\mathcal{S}^\varepsilon(W)} J(u') \leq \sup_{g'\in W} J(g') + \varepsilon < M,$$

which, according to (J1), implies that the set $\mathcal{S}^\varepsilon(W)$ is bounded in norm, i.e. $\sup_{u'\in\mathcal{S}^\varepsilon(W)} \|u'\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)} < M'$. This shows that $\mathcal{S}^\varepsilon$ is locally bounded, hence we can find a sufficiently large open ball $B \subset \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ so that $u_j \in B$ holds for all $j \in \mathbb{N}$.

The ball B equipped with the weak topology of $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ is relatively compact and hence metrizable by a metric d, according to Proposition A.5. Define $U' = U \cap B$, then U' is a weakly open neighbourhood of $\mathcal{S}^\varepsilon(g)$ in the relative topology of (B,d) and we can easily see that $u_j \in \mathcal{S}^\varepsilon(g_j)\backslash U'$. By Theorem 3.2, we can find a subsequence g_{j_k} of g_j and a $u \in \mathcal{S}^\varepsilon(g)$ such that $u_{j_k} \overset{w}{\rightharpoonup} u$ (i.e. $d(u_{j_k}, u) \to 0$) as $k \to \infty$. However, since $u \in U'$, eventually we must have $u_{j_k} \in U'$ for sufficiently large k, which contradicts the fact that $u_j \notin U'$ for all $j \in \mathbb{N}$. This shows that our assumption that $\mathcal{S}^\varepsilon$ is not $\|\cdot\|$-w-USC is false.

Proving the last claim is simpler, we let $\varepsilon = 0$ and suppose that (F3') holds. As before, we assume for contradiction that $\mathcal{S}$ is not $\|\cdot\|$-$\|\cdot\|$-USC, so there exists g and a neighbourhood $U \in \mathcal{N}(\mathcal{S}(g))$ (in the norm topology) such that there exists a sequence $g_j \to g$ and a corresponding sequence u_j satisfying $u_j \in \mathcal{S}^{\varepsilon}(g_j) \backslash U$. By Theorem 3.2, we can find a subsequence g_{j_k} of g_j and a $u \in \mathcal{S}(g)$ such that $u_{j_k} \to u$ in norm. However, u_{j_k} is eventually in U since U is a neighbourhood of u, which is a contradiction to $u_j \notin U$. This proves that $\mathcal{S}$ is $\|\cdot\|$-$\|\cdot\|$-USC. $\square$

4 Rectifiability of the Sets of Extremals for a Singularly Perturbed Variational Integrals

The structure of the sets of minimizers (or, more generally, the sets of extremals) for a singularly perturbed strongly quasiconvex variational integral is the main object of study in this chapter, which is joint work with Prof Jan Kristensen. The unperturbed variational integral takes the form

$$J(u) = \int_\Omega F(\nabla u)\, dx,$$

where the main assumptions for our integrand $F\colon \mathbb{R}^N \to \mathbb{R}$ is the *p-growth condition* (or the *natural growth condition*)

$$|F(z)| \leq k(1 + |z|^p) \tag{F$_p$1}$$

for some $k > 0$ (so that the variational integral $J(u)$ is well-defined as a nonlinear functional $J\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N) \to \mathbb{R}$) and the *strong quasiconvexity* (or, more precisely, *strong p-quasiconvexity*), which means that

$$F - lE_p \quad \text{is quasiconvex for some } l > 0 \tag{F$_p$2}$$

where the *reference integrand* E_p is defined as

$$E_p(z) = \left(1 + |z|^2\right)^{\frac{p}{2}} - 1.$$

It is well known that strong quasiconvexity (or even strong polyconvexity) of the integrand is not enough to guarantee uniqueness of the minimizers or the extremals for the corresponding variational problems. In fact, the extremals of a strongly quasiconvex variational integral could behave quite badly, as shown in [80] by Müller and Šverák (see also [60] and [96]) that apart from the lack of uniqueness, the extremals can be very irregular. There are partial results for smooth extremals exhibiting good behaviour, however, e.g. the work by Zhang [110] where he uses the implicit function theorem to establish the existence and local uniqueness in the Hölder norm of the extremals, or [111] where he shows that smooth extremals are strong local minimizers for strongly quasiconvex variational integrals. Regarding the non-uniqueness, an analysis of the number of strong local minimizers for quasiconvex variational integrals for domains Ω with nontrivial topology is given by Taheri in [99] (see also [97]). Since it is not possible to prove either regularity nor uniqueness for a variational integral with a strongly quasiconvex integrand in general, one way to remedy the situation is to perturb the problem by adding a regularizing term to it and see if the extremals (and minimizers) of the perturbed problem behave better.

In this chapter, we study a singularly perturbed variational integral of the form

$$J_\varepsilon(u) = \int_\Omega F(\nabla u)\,dx + \varepsilon \int_\Omega |\nabla^2 u|^2\,dx, \tag{4.1}$$

where Ω is a regular domain in $\mathbb{R}^n$, $F\colon \mathbb{R}^N \to \mathbb{R}$ is a strongly quasiconvex integrand, and u is a vector-valued function in $\mathrm{W}^{2,2}(\Omega;\mathbb{R}^N)$ (with $\nabla^2 u$ denoting the Hessian of u). Recall that a **d-rectifiable set** M (in Federer's sense [44]) is a subset of a metric space that is the Lipschitz image of some bounded set in $\mathbb{R}^d$, i.e. $M = f(K)$ where f is Lipschitz and $K \subset \mathbb{R}^d$ is a bounded set. The main results of this chapter are the following:

Theorem 4.1. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with C^4 boundary, $p \in \left[2, \frac{2n}{n-2}\right)$, $g \in \mathrm{W}^{3,r}(\Omega;\mathbb{R}^N)$ where $r \geq p'$ and $r > \frac{n}{2}$, and let $F\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ be an integrand that satisfies the p-growth (F_p1) and the strong quasiconvexity (F_p2). Let $\mathcal{E} \subset \mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$ be the set of all the extremals of*

$$J_\varepsilon(u) = \int_\Omega F(\nabla u)\,dx + \varepsilon \int_\Omega |\nabla^2 u|^2\,dx.$$

Let $\delta > 0$, $R > 0$, and $\mathcal{S} := \{u \in \mathcal{E} : \|u\|_{\mathrm{W}^{1,\frac{n(p-1)}{2}+\delta}} \leq R\}$. Then $\mathcal{S}$, viewed as a metric subspace of $\mathrm{W}^{2,2}(\Omega;\mathbb{R}^N)$, is a d-rectifiable set (d depends on the parameters in (F_p1), (F_p2) and the set $\mathcal{S}$). In particular, $\mathcal{S}$ is bi-Lipschitz equivalent to a bounded subset $K \subset \mathbb{R}^d$.

Theorem 4.1 says that any bounded (in a suitable Sobolev norm) subset of the extremals is bi-Lipschitz equivalent to a bounded subset in some finite dimensional Euclidean space. In particular, this boundedness assumption is automatically satisfied for minimizers under suitable circumstances, and hence we have an immediate corollary:

Theorem 4.2. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with C^4 boundary, $n < 6$, $p \in \left[2, \frac{n+2}{n-2}\right)$, $g \in \mathrm{W}^{3,r}(\Omega;\mathbb{R}^N)$ where $r \geq p'$ and $r > \frac{n}{2}$, and let $F\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ be an integrand that satisfies the conditions (F_p1) and (F_p2). Let $\mathcal{M} \subset \mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$ be the set of all the absolute minimizers of*

$$J_\varepsilon(u) = \int_\Omega F(\nabla u)\,dx + \varepsilon \int_\Omega |\nabla^2 u|^2\,dx.$$

Then $\mathcal{M}$, viewed as a metric subspace of $\mathrm{W}^{2,2}(\Omega;\mathbb{R}^N)$, is a d-rectifiable set (d depends on the parameters in (F_p1), (F_p2)). In particular, $\mathcal{M}$ is bi-Lipschitz equivalent to a compact subset $K \subset \mathbb{R}^d$.

In this case, we obtain a stronger conclusion that the set $\mathcal{M}$ of all the minimizers of J_ε is not only bi-Lipschitz equivalent to a bounded subset $K \subset \mathbb{R}^d$, but also that

K is compact. It is not unreasonable to hope that under all these assumptions, $\mathcal{M}$ could in fact be a singleton set, i.e. that the minimizer of the perturbed variational integral must be unique, but this is false in general. The method used in the recent work [15] by Campos Cordero, Kirchheim, Kolář and Kristensen can be used to show non-uniqueness of minimizers even for the perturbed variational problem.

Remark 4.3. If $\Omega \subset \mathbb{R}^n$ is a star-shaped domain, then one can obtain a bound on the norm of the extremals in $\mathcal{E}$ using a calculation from [98] by Taheri. This would remove some restrictions on the exponent p in Theorem 4.2 and the dimension n.//

A crucial tool that we used to prove our main result is the *nonlinear Gårding inequality* which says that if F is an integrand satisfying the growth assumption (F_p1) and the strong quasiconvexity assumption (F_p2), then for a fixed $u \in C^1(\overline{\Omega}; \mathbb{R}^N)$,

$$c \int_\Omega E_p(\nabla\varphi(x))\,dx \leq \int_\Omega F_{\nabla u(x)}(\nabla\varphi(x))\,dx + C\int_\Omega E_p(\varphi(x))\,dx, \tag{4.2}$$

holds for any $\varphi \in \mathrm{W}^{1,p}_0(\Omega; R^N)$ for some constants $c, C \geq 0$. Here E_p is the *reference integrand* $E_p(z) = (1+|z|^2)^{\frac{p}{2}} - 1$, and $F_w(z) := F(w+z) - F(w) - F'(w)[z]$. This inequality is a generalization of a recent work by Campos Cordero and Kristensen [16] for the $p = 2$ case. A similar inequality was also used by Koumatos et al. (see [46] and [56]), where their proof follows closely an indirect argument used in Campos Cordero's thesis. In this chapter, we shall prove this inequality only for $p > 1$. The inequality does also hold for the case $p = 1$ but requires a slightly different method to prove some estimates related to the reference integrand E_1 (this could be done by adapting the estimates used in [49]). Most importantly, we will show that the stronger inequality

$$c \int_\Omega E_p(\nabla\varphi(x))\,dx \leq \int_\Omega F_{\nabla u(x)}(\nabla\varphi(x))\,dx \tag{4.3}$$

holds for any $\varphi \in Y$, where $Y \subset \mathrm{W}^{1,p}_0(\Omega; R^N)$ is a linear subspace of finite codimension. This implies uniqueness of minimizers up to a finite dimensional subspace, and ultimately the d-rectifiability of the set of all minimizers of J_ε.

The organization of this chapter will be as follows. First, we prove the existence of extension operators $T\colon \mathrm{W}^{1,r}(\partial Q; \mathbb{R}^N) \to \mathrm{W}^{1,r}((1+\varepsilon)Q\backslash Q; \mathbb{R}^N)$ (here Q can be a cube or a ball) that improve the integrability of $\Phi(|\nabla u|)$, where Φ is a Young-type convex function. After that, we prove the nonlinear Gårding inequality under certain assumptions, and apply it to a strongly quasiconvex variational integral whose integrand F has p-growth.

Since one requirement for applying the Gårding inequality is a smoothness condition, we use an elliptic regularity result to prove a $C^{1,\alpha}$ estimate on the set of

extremals of J_ε in a Dirichlet class with sufficiently smooth boundary datum, and after this we will prove that the Gårding inequality, which holds uniformly on a $C^{1,\alpha}$-bounded set, implies strict convexity of the functional J up to a finite dimensional subspace of $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$. Our main result regarding the d-rectifiability of the sets of extremals and minimizers is then deduced.

4.1 Extension Operators and Higher Integrability

The main goal of this section is to prove the following proposition that will be used to prove the nonlinear Gårding inequality.

Proposition 4.4. *Let $Q = \left(-\frac{1}{2},\frac{1}{2}\right)^n$ be a unit cube in $\mathbb{R}^n$ and let Φ be a convex function such that $\Phi(\alpha z) \leq \alpha^\delta\,(\Phi(z))$ for every $\alpha \geq 1$, where $\delta \geq 1$ is some fixed constant. There exists an $\varepsilon_0 > 0$ (depending only on the dimension n) such that for any given $0 < \varepsilon \leq \varepsilon_0$ and any $p \in [1,\infty)$, there exist $C = C(n,\delta)$ and an extension operator $T\colon \mathrm{W}^{1,p}(\partial Q;\mathbb{R}^N) \to \mathrm{W}^{1,p}(\mathbb{R}^n\backslash\overline{Q};\mathbb{R}^N)$ mapping u to $\widetilde{u}$ such that*

$$\int_{(1+\varepsilon)Q\backslash Q} \Phi(|\nabla\widetilde{u}(x)|)\,dx \leq \varepsilon C \int_{\partial Q} \Phi(|\nabla_\tau u(y)|)\,d\mathcal{H}^{n-1}(y) + \frac{C}{\varepsilon^{\delta-1}} \int_{\partial Q} \Phi(|u(y)|)\,d\mathcal{H}^{n-1}(y),$$

where ∇_τ denotes that tangential gradient operator. Moreover, $\widetilde{u}$ is smooth on $\mathbb{R}^n\backslash\overline{Q}$ and $\widetilde{u} = 0$ outside $(1+\varepsilon)Q$.

Additionally, we will prove the existence of another extension operator that is similar to [19] and [57]. Such an extension operator could be of independent interest, or it can be used in conjunction with a Lipschitz cut-off function to give an alternative proof to Lemma 4.13 (in the next section), which is the strategy used in a work by Campos Cordero and Kristensen to prove a similar result.

Proposition 4.5. *Let $Q = \left(-\frac{1}{2},\frac{1}{2}\right)^n$ be a unit cube in $\mathbb{R}^n$ and let Φ be a convex function such that $\Phi(\alpha z) \leq \alpha^\delta \Phi(z)$ for each $\alpha \geq 1$. Then for any given $1 < q < \frac{n}{n-1}$ and any $r \in [1,\infty)$, there exist $\varepsilon_0 > 0$, $C = C(n,q,\delta)$ and an extension operator $T\colon \mathrm{W}^{1,r}(\partial Q;\mathbb{R}^N) \to \mathrm{W}^{1,r}((1+\varepsilon_0)Q\backslash Q;\mathbb{R}^N)$ mapping u to $\widetilde{u}$ such that*

$$\int_{(1+\varepsilon_0)Q\backslash Q} \Phi(|\nabla\widetilde{u}(x)|)^q\,dx \leq C\left(\int_{\partial Q} \Phi(|\nabla_\tau u(y)|)\,d\mathcal{H}^{n-1}(y)\right)^q, \tag{4.4}$$

where ∇_τ denotes that tangential gradient operator. Moreover, $\widetilde{u}$ is smooth on $(1+\varepsilon_0)Q\backslash\overline{Q}$.

We shall see from the proof that the above result still holds if we replace the unit cube Q with the unit ball $B(0,1)$ instead.

Before proving these extension results, we shall first prove the following Poincaré inequality of Orlicz–Sobolev type, which will help us strengthen the inequalities. The proof is based on the an estimate involving the Riesz potential, similar to [47, Lemma 7.16] (see also [73] and [33] for related results).

Lemma 4.6. *Let $\Phi\colon \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing C^1 convex function which has the scaling property in the sense that for any $l \geq 1$, there exists a constant $c(l)$ such that $\Phi(lt) \leq c(l)\Phi(t)$. We also assume that Φ has q-growth, i.e. there exists an $L > 0$ such that*

$$\Phi(t) \leq L(1 + t^q).$$

Let S^d be the unit sphere in $\mathbb{R}^{d+1}$; then there exists a constant $C = C(d)$ such that for any $u \in \mathrm{W}^{1,q}(S^d; \mathbb{R}^N)$, we have

$$\int_{S^d} \Phi(|u(x) - \overline{u}|)\, d\mathcal{H}^d(x) \leq C \int_{S^d} \Phi(|\nabla_\tau u(x)|)\, d\mathcal{H}^d(x), \tag{4.5}$$

where $\overline{u} := \fint_{S^d} u(x)\, d\mathcal{H}^d(x)$.

Proof. We shall parametrize S^d by the function $T\colon B_\pi \to S^d$, where $B_\pi = B^d(0, \pi)$ is the ball in $\mathbb{R}^d$ of radius π, and T is defined by

$$T(x) = \left(\sin(|x|)\frac{x}{|x|}, -\cos(|x|)\right) \in \mathbb{R}^d \times \mathbb{R}.$$

Its derivative at x is the map

$$DT(x) = \left(\operatorname{sinc}(|x|) I_{d\times d} + \frac{\operatorname{sinc}'(|x|)}{|x|} x x^t\right) \oplus \sin(|x|) x^t.$$

We can calculate its Jacobian of transformation $JT = \sqrt{\det DT^t DT}$ to be $JT(x) = \operatorname{sinc}(|x|)^{d-1}$ (it is enough to calculate $JT(x)$ for $x = te_1$ and use rotational symmetry of T to deduce the general formula). Also, we can verify that $|DT(tx)x| = |x|$ for any $t \in (0, 1)$, $x \in B_\pi$, and that $|x| = \rho(Tx, T(0))$, where ρ is the Riemannian distance on S^d.

Let us assume first that $u \in C^1(S^d; \mathbb{R}^N)$ and denote $x_0 = T(0)$. Then

$$\begin{aligned} u(x_0) - u(Tx) &= -\int_0^1 D(u \circ T)(tx)[x]\, dt \\ &= -\int_0^1 Du(T(tx))[DT(tx)x]\, dt \end{aligned}$$

and we can integrate the above over the ball B_π to get

$$\int_{B_\pi} (u(x_0) - u(Tx)) JT(x)\, dx \leq \int_{B_\pi} \int_0^1 |Du(T(tx))||x| JT(x)\, dt\, dx.$$

Recall that $\int_{B_\pi} f(Tx)JT(x)\,dx = \int_{S^d} f(y)\,d\mathcal{H}^d(y)$ and that $\operatorname{sinc}(t)$ is decreasing on $[0,\pi]$, so $JT(x) \le JT(tx)$ for all $t \in [0,1]$. Hence

$$\begin{aligned}|S^d|(u(x_0)-\overline{u}) &\le \int_{B_\pi}\int_0^1 |Du(T(tx))||x|JT(tx)\,dt\,dx \\ &\le \int_0^\infty \int_0^\pi \int_{\{|x|=r\}} rV(xt)\,d\mathcal{H}^{d-1}(x)\,dr\,dt,\end{aligned}$$

where we set

$$V(x) := \begin{cases} |Du(T(x))|JT(x) & ; x \in B_\pi \\ 0 & ; x \notin B_\pi. \end{cases}$$

By the change of variables $z = tx$ and $s = tr$, we have

$$\begin{aligned}u(x_0)-\overline{u} &\le \frac{1}{|S^d|}\int_0^\pi\int_0^\infty\int_{\{|z|=s\}} \frac{r^{d-1}}{s^{d-1}}V(z)\,d\mathcal{H}^{d-1}(z)\,ds\,dr, \\ &= \frac{1}{|S^d|}\left(\int_0^\pi r^{d-1}\,dr\right)\int_0^\infty\int_{\{|z|=s\}}\frac{V(z)}{|z|^{d-1}}\,d\mathcal{H}^{d-1}(z)\,ds \\ &= \frac{\pi^d}{d|S^d|}\int_{B_\pi}\frac{|Du(T(z))|}{\rho(x_0,Tz)^{d-1}}JT(z)\,dz \\ &= \frac{\pi^d}{d|S^d|}\int_{S^d}\frac{|Du(y)|}{\rho(x_0,y)^{d-1}}\,d\mathcal{H}^d(y).\end{aligned}$$

The point x_0 in the above can be replaced with any $x \in S^d$ by a simple rotation and since $\int_{S^d}\frac{d\mathcal{H}^d(y)}{\rho(x_0,y)^{d-1}} < \infty$, we can apply Jensen's inequality (and use the fact that Φ is increasing and has a scaling property) to get

$$\Phi(|u(x)-\overline{u}|) \le C\int_{S^d}\frac{\Phi(|Du(y)|)}{\rho(x,y)^{d-1}}\,d\mathcal{H}^d(y)$$

for some constant $C = C(d)$. We then integrate the inequality to get

$$\begin{aligned}\int_{S^d}\Phi(|u(x)-\overline{u}|)\,d\mathcal{H}^d(x) &\le C\int_{S^d}\left(\int_{S^d}\frac{d\mathcal{H}^d(x)}{\rho(x,y)^{d-1}}\right)\Phi(|Du(y)|)\,d\mathcal{H}^d(y) \\ &\le C\int_{S^d}\Phi(|Du(y)|)\,d\mathcal{H}^d(y),\end{aligned}$$

which holds for any $u \in C^1(S^d;\mathbb{R}^N)$. The term $|Du|$ can be replaced by $|\nabla_\tau u|$ since they are equivalent.

Finally, we observe that since Φ satisfies the q-growth condition, the map $f \mapsto \int_{S^d}\Phi(|f|)\,d\mathcal{H}^d$ is continuous in the L^q norm. Since smooth functions are dense in $\mathrm{W}^{1,q}(S^d;\mathbb{R}^N)$, for any given $u \in \mathrm{W}^{1,q}(S^d;\mathbb{R}^N)$ we can take a sequence $u_n \in C^1(S^d;\mathbb{R}^N)$ such that $u_n \to u$ in $\mathrm{W}^{1,q}(S^d;\mathbb{R}^N)$ and take the limit of

$$\int_{S^d}\Phi(|u_n(x)-\overline{u}_n|)\,d\mathcal{H}^d(x) \le C\int_{S^d}\Phi(|\nabla_\tau u_n(x)|)\,d\mathcal{H}^d(x)$$

as $n \to \infty$ to obtain the inequality for u (with the same constant C). This completes the proof. □

It is not hard to see that an analogous inequality also holds for $u \in \mathrm{W}^{1,q}(rS^d;\mathbb{R}^N)$ (here rS^d is of course the sphere of radius $r>0$), but the constant C in (4.5) would depend on R as well. The only modification we need to make in the proof is replacing the parametrization T with $T(x) = \left(\sin(|x|)\frac{rx}{|x|}, -r\cos(|x|)\right)$. This fact will not be used in this work, however.

To prove Proposition 4.4 and Proposition 4.5, we will first prove corresponding results for extension operators on the half-space $\mathbb{R}^{n-1}\times\mathbb{R}_+$ of $\mathbb{R}^n$. Then the results are translated onto a unit ball $B^n \subset \mathbb{R}^n$, which is bi-Lipschitz equivalent to an n-dimensional cube, by a patching argument, and finally translated onto a cube.

The following lemma corresponds to Proposition 4.4.

Lemma 4.7. *Let $p \geq 1$ and $\Phi\colon \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing convex function such that $\Phi(\alpha z) \leq \alpha^\delta\,(\Phi(z))$ for every $\alpha \geq 1$, where $\delta \geq 1$ is some fixed constant. Let $\varepsilon > 0$ be given; then there exists an extension operator that maps any $u \in \mathrm{W}^{1,p}(\mathbb{R}^{n-1};\mathbb{R}^N)$ to $U \in \mathrm{W}^{1,p}(\mathbb{R}^n;\mathbb{R}^N)$ such that*

$$\int_{\mathbb{R}^n_+} \Phi(|\nabla U(x,t)|)\,d(x,t) \leq \varepsilon 4^\delta \int_{\mathbb{R}^{n-1}} \Phi(|\nabla u(x)|)\,dx + \frac{8^\delta}{\varepsilon^{\delta-1}} \int_{\mathbb{R}^{n-1}} \Phi(|u(x)|)\,dx. \tag{4.6}$$

U is smooth on $\mathbb{R}^{n-1}\times(0,\infty)$ and $U(x,t) = 0$ for all $t > \varepsilon$. Moreover, if we assume that $u \in \mathrm{W}^{1,p}_0(B^{n-1}(0,R);\mathbb{R}^N)$, then U is supported on $B^{n-1}(0,R+\varepsilon)\times[0,\varepsilon]$.

Proof. Choose a smooth function $\eta\colon [0,\infty) \to [0,1]$ such that $\eta(0) = 1$ and $\eta(t) = 0$ for all $t > \varepsilon$ and $|\eta'| \leq \frac{2}{\varepsilon}$. Let K be a symmetric smooth function on $\mathbb{R}^{n-1}$ that is supported on $B^{n-1}(0,1)$. Define $K_t(x) := \frac{1}{t^{n-1}}K\left(\frac{x}{t}\right)$ and let

$$\begin{aligned} U(x,t) &:= \eta(t)(K_t * u)(x) \\ &= \eta(t)\int_{\mathbb{R}^{n-1}} K_t(y)u(x-y)\,dy. \end{aligned}$$

The facts that U is smooth on $\mathbb{R}^{n-1}\times(0,\infty)$, $U(x,t) = 0$ for all $t > \varepsilon$, and that U is supported on $B^{n-1}(0,R+\varepsilon)\times[0,\varepsilon]$ if $u \in \mathrm{W}^{1,p}_0(B^{n-1}(0,R);\mathbb{R}^N)$ all follow easily from the definition of U.

For each fixed $t \leq \varepsilon$ and $i \in \{1,\dots,n-1\}$, we have

$$\begin{aligned} \nabla_x U(x,t) &= \eta(t)(K_t * \nabla u)(x) \\ &= \eta(t)\int_{\mathbb{R}^{n-1}} \frac{1}{t^{n-1}}K\left(\frac{y'}{t}\right)\nabla u(x-y')\,dy' \\ &= \eta(t)\int_{\mathbb{R}^{n-1}} K(y)\nabla u(x-ty)\,dy \end{aligned}$$

and hence

$$\begin{aligned}\Phi(|\nabla_x U(x,t)|) &\leq \Phi\left(\eta(t)\int_{\mathbb{R}^{n-1}} K(y)|\nabla u(x-ty)|\,dy\right)\\ &\leq \Phi\left(\int_{\mathbb{R}^{n-1}} K(y)|\nabla u(x-ty)|\,dy\right)\\ &\leq \int_{\mathbb{R}^{n-1}} K(y)\Phi(|\nabla u(x-ty)|)\,dy\end{aligned}$$

using the fact that $\eta \leq 1$ and Jensen's inequality. We integrate the above in the x variable over $\mathbb{R}^{n-1}$ and use Fubini's theorem to get

$$\begin{aligned}\int_{\mathbb{R}^{n-1}} \Phi(|\nabla_x U(x,t)|)\,dx &\leq \int_{\mathbb{R}^{n-1}}\int_{\mathbb{R}^{n-1}} K(y)\Phi(|\nabla u(x-ty)|)\,dy\,dx\\ &= \int_{\mathbb{R}^{n-1}} K(y)\int_{\mathbb{R}^{n-1}} \Phi(|\nabla u(x-ty)|)\,dx\,dy\\ &= \int_{\mathbb{R}^{n-1}} \Phi(|\nabla u(x)|)\,dx.\end{aligned}$$

We then integrate the above in t, taking into account $U(x,t) = 0$ for $t > \varepsilon$, to get

$$\int_{\mathbb{R}^n_+} \Phi(|\nabla_x U(x,t)|)\,d(x,t) \leq \varepsilon \int_{\mathbb{R}^{n-1}} \Phi(|\nabla u(x)|)\,dx \tag{4.7}$$

Now, to estimate the term $\partial_t U(x,t))$, we first observe that

$$\partial_t U(x,t) = \eta'(t)(K_t * u)(x) - \eta(t)\int_{B_1(0)} K(y)\nabla u(x-ty)[y]\,dy$$

and that the condition $\Phi(\alpha z) \leq \alpha^\delta\,(\Phi(z))$ implies that $\Phi(z+w) \leq 2^{\delta-1}\,(\Phi(z)+\Phi(w))$. We can apply the same technique as before to get

$$\begin{aligned}\Phi(|\partial_t U(x,t)|) &\leq 2^{\delta-1}\Phi\left(\eta'(t)|K_t * u|(x)\right) + 2^{\delta-1}\Phi\left(\eta(t)\int_{B_1(0)} K(y)|\nabla u(x-ty)||y|\,dy\right)\\ &\leq \frac{2^{2\delta-1}}{\varepsilon^\delta}\Phi\left(\int_{\mathbb{R}^{n-1}} K(y)|u(x-ty)|\,dy\right) + 2^{\delta-1}\Phi\left(\int_{B_1(0)} K(y)|\nabla u(x-ty)|\,dy\right)\\ &\leq \frac{2^{2\delta-1}}{\varepsilon^\delta}\int_{\mathbb{R}^{n-1}} K(y)\Phi(|u(x-ty)|)\,dy + 2^{\delta-1}\int_{\mathbb{R}^{n-1}} K(y)\Phi(|\nabla u(x-ty)|)\,dy\end{aligned}$$

We integrate the above in x and then in t as before to get

$$\int_{\mathbb{R}^n_+} \Phi(|\partial_t U(x,t)|)\,d(x,t) \leq \frac{2^{2\delta-1}}{\varepsilon^{\delta-1}}\int_{\mathbb{R}^{n-1}} \Phi(|u(x)|)\,dx + 2^{\delta-1}\varepsilon\int_{\mathbb{R}^{n-1}} \Phi(|\nabla u(x)|)\,dx. \tag{4.8}$$

Lastly, since $|\nabla U| \leq |\nabla_x U| + |\partial_t U|$, we can use the inequality $\Phi(|\nabla U|) \leq 2^{\delta-1}\Phi(|\nabla_x U|) + 2^{\delta-1}\Phi(|\partial_t U|)$ to get the desired estimate. The fact that $U \in \mathrm{W}^{1,p}(\mathbb{R}^n;\mathbb{R}^N)$ is not hard to verify using a standard technique (alternatively, we can let $\Phi(|z|)$ be $|z|^p$ and modify the above proof slightly to get an estimate for the $\mathrm{W}^{1,p}$ norm of U in terms of $\mathrm{W}^{1,p}$ norm of u). The smoothness of U on $\mathbb{R}^{n-1}\times(0,\infty)$ is evident from the construction. □

Proving the half-space version of Proposition 4.5 is a little bit more involved.

Lemma 4.8. *Let $\Phi\colon \mathbb{R}_+ \to \mathbb{R}_+$ be a convex function and let $f \in \mathrm{L}^\Phi(\mathbb{R}^{n-1};\mathbb{R}^d)$ be a function supported on the ball $B_R := B^{n-1}(0,R) \subset \mathbb{R}^{n-1}$. Let K be a symmetric smooth function on $\mathbb{R}^{n-1}$ that is supported on $B^{n-1}(0,1)$ and write $K_t(x) := \frac{1}{t^{n-1}}K\left(\frac{x}{t}\right)$. For any $R' \geq R$ and any $1 < q < \frac{n}{n-1}$, there exists a constant $C = C(n,q)$ such that the function $F\colon \mathbb{R}^n_+ \to \mathbb{R}^d$ defined as*

$$F(x,t) := (K_t * f)(x)$$

satisfies the estimate

$$\int_{B_{R'}\times\mathbb{R}_+} \Phi(|F(x,t)|)^q \, d(x,t) \leq C|B_{R'}|^{\frac{n}{n-1}-q}\left(\int_{B_R}\Phi(|f|)\,d\mathcal{H}^{n-1}\right)^q. \tag{4.9}$$

Proof. For a fixed $x \in \mathbb{R}^{n-1}$ and any $\delta = \delta(x) > 0$, we apply Jensen's inequality to get

$$\begin{aligned}\int_0^\infty \Phi(|F(x,t)|)^q\,dt &= \int_0^\infty \Phi(|K_t * f|(x))^q\,dt\\ &\leq \int_0^\infty ((K_t * \Phi(|f|))(x))^q\ dt\\ &= \int_0^\delta ((K_t * \Phi(|f|))(x))^q\ dt + \int_\delta^\infty ((K_t * \Phi(|f|))(x))^q\ dt.\end{aligned}$$

Recall that for a positive function g, we have the inequality $\sup_{t>0}|(g*K_t)(x)| \leq c_K Mg(x)$, where Mg is the maximal function $Mg(x) = \sup_{t>0}\fint_{B(x,t)} g(y)\,dy$; hence the first term in the above inequality is bounded by

$$\int_0^\delta ((K_t * \Phi(|f|))(x))^q\ dt \leq c\delta(M\Phi(|f|)(x))^q.$$

For the second term, we have the bound

$$\begin{aligned}\int_\delta^\infty ((K_t * \Phi(|f|))(x))^q\ dt &= \int_\delta^\infty t^{-(n-1)q}\left(\int_{B^{n-1}} K\left(\frac{y}{t}\right)\Phi(|f(x-y)|)\,dy\right)^q dt\\ &\leq \int_\delta^\infty t^{-(n-1)q}\left(\int_{B_R}\Phi(|f(y)|)\,dy\right)^q dt\\ &\leq c\delta^{1-(n-1)q}\left(\int_{B_R}\Phi(|f(y)|)\,dy\right)^q.\end{aligned}$$

Note that here we require q to satisfy $q > \frac{1}{n-1}$ so that $\int_\delta^\infty t^{-(n-1)q}\,dt < \infty$, but this is always the case for $n \geq 2$.

We may choose $\delta(x) = \left(\frac{\left(\int_{B_R}\Phi(|f(y)|)\,dy\right)}{M\Phi(|f|)(x)}\right)^{\frac{1}{n-1}}$ to get the bound

$$\int_0^\infty \Phi(|F(x,t)|)^q\,dt \leq C\left(\int_{B_R}\Phi(|f(y)|)\,dy\right)^{\frac{1}{n-1}}(M\Phi(|f|)(x))^{q-\frac{1}{n-1}},$$

where $C = C(n, q)$ does not depend on x. For $R' \geq R$, we can integrate the above inequality over the cube $B_{R'} \subset \mathbb{R}^{n-1}$ to get

$$\int_{B_{R'}\times\mathbb{R}_+} \Phi(|F(x,t)|)^q \, d(x,t) \leq C \left(\int_{B_R} \Phi(|f(y)|) \, dy \right)^{\frac{1}{n-1}} \int_{B_R} (M\Phi(|f|)(x))^{q-\frac{1}{n-1}} \, dx.$$

Since $q < \frac{n}{n-1}$, $s := q - \frac{1}{n-1} < 1$ and we have the maximal inequality

$$\int_{B_R} (M\Phi(|f|)(x))^s \, dx \leq c_s |B_{R'}|^{1-s} \left(\int_{B_R} (\Phi(|f|)(x)) \, dx \right)^s$$

(see Stein's [92, I, 8.15]), thus

$$\int_{B_{R'}\times\mathbb{R}_+} \Phi(|F(x,t)|)^q \, d(x,t) \leq C |B_{R'}|^{\frac{n}{n-1}-q} \left(\int_{B_R} \Phi(|f(y)|) \, dy \right)^q$$

for some $C = C(n, q)$. This concludes the proof. □

Lemma 4.9. *Let $\Phi\colon \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing convex function such that $\Phi(z + w) \leq \alpha(\Phi(z) + \Phi(w))$ for some $\alpha \geq 1$. Let $r \geq 1$ and $u \in \mathrm{W}^{1,r}(\mathbb{R}^{n-1}; \mathbb{R}^N)$ be a function supported on the ball $B_R = B_R^{n-1} \subset \mathbb{R}^{n-1}$. For any $T > 0$, u can be extended to a function $U \in \mathrm{W}^{1,r}(B_{R+T} \times (0,T); \mathbb{R}^N)$ (in the sense that $\mathrm{tr}\,(U) = u$) so that for any $1 < q < \frac{n}{n-1}$ we have the estimates*

$$\int_{B_{R+T}\times(0,T)} \Phi(|U(x,t)|)^q \, d(x,t) \leq C \left(\int_{B_R} \Phi(|u|) \, d\mathcal{H}^{n-1} \right)^q \tag{4.10}$$

and

$$\int_{B_{R+T}\times(0,T)} \Phi(|\nabla U(x,t)|)^q \, d(x,t) \leq C \left(\int_{B_R} \Phi(|\nabla u|) \, d\mathcal{H}^{n-1} \right)^q \tag{4.11}$$

for some $C = C(n, q, \alpha, R, T)$. Moreover, U is smooth on $B_{R+T}\times(0,T)$ and vanishes at $\partial B_{R+T} \times (0,T)$.

Proof. Let $(K_t)_{t>0}$ be standard symmetric smooth mollifiers as defined in Lemma 4.8 and define U to be

$$U(x,t) := (K_t * u)(x).$$

The fact that $U \in \mathrm{W}^{1,r}(B_{R+T} \times (0,T); \mathbb{R}^N)$, $\mathrm{tr}\,(U) = u$ on $\mathbb{R}^{n-1}$, U is smooth on $B_{R+T} \times (0,T)$, and that it vanishes at $\partial B_{R+T} \times (0,T)$, can be verified using a standard argument regarding mollifiers.

We then apply Lemma 4.8 with $f = |\nabla u|$ to get

$$\int_{B_{R+T}\times\mathbb{R}_+} \Phi(K_t * |\nabla u(x)|)^q \, d(x,t) \leq C \left(\int_{B_R} \Phi(|\nabla u(y)|) \, dy \right)^q \tag{4.12}$$

for some $C = C(n, q, R, T)$. We can estimate $|\nabla_x U|$ and $|\partial_t U|$ separately in terms of $K_t * |\nabla u(x)|$ and combine the estimates to get an estimate for $|\nabla U|$. The first half is easy since we have

$$|\nabla_x U(x,t)| = |(K_t * \nabla_x u)(x)| \le (K_t * |\nabla u|)(x)$$

from Jensen's inequality. As for the second term, direct calculation gives

$$\partial_t U(x,t) = -\int_{B_1(0)} K(y) \nabla u(x - ty)[y]\, dy,$$

and hence we also have

$$|\partial_t U(x,t)| \le \int_{B_1(0)} K(y) |\nabla u|(x - ty)\, dy = (K_t * |\nabla u|)(x)$$

in this case.

Lastly, since $|\nabla U| \le |\nabla_x U| + |\partial_t U|$, it follows that

$$\begin{aligned}
\Phi(|\nabla U(x,t)|)^q &\le \big(\alpha\Phi(|\nabla_x U(x,t)|) + \alpha\Phi(|\partial_t U(x,t)|)\big)^q \\
&\le \alpha^q 2^{q-1} \left(\Phi(|\nabla_x U(x,t)|)^q + \Phi(|\partial_t U(x,t)|)^q\right) \\
&\le (2\alpha)^q \big(\Phi(K_t * |\nabla u|(x))\big)^q
\end{aligned}$$

so we can integrate the above over $B_{R+T} \times (0,T)$ and apply (4.12) to conclude the proof for (4.11). The proof for (4.10) is simpler and follows directly from the substitution $f = u$ in Lemma 4.8. □

This concludes the proof for the half-space version of Proposition 4.5.

Next, we shall use the "patching technique" to translate the results from half space in Lemma 4.7 and Lemma 4.9 to a unit ball. The following proposition is formulated in a way that we can apply it to both cases. Note that the assumption that $\Phi(\alpha z) \le \alpha^\delta \Phi(z)$ for each $\alpha \ge 1$ below is stronger than the subadditivity assumption for Φ in Lemma 4.9 (an easy application of Jensen's inequality).

Lemma 4.10. *Let $\Phi\colon \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing convex function such that $\Phi(\alpha z) \le \alpha^\delta \Phi(z)$ for each $\alpha \ge 1$. Let $B_R^n = B^n(0,R) \subset \mathbb{R}^n$ be the n-dimensional ball of radius R, and $B := B_1^n$. Let $q, r \ge 1$ and suppose that for each $T > 0$, there is an extension operator mapping $v \in \mathrm{W}_0^{1,r}(B_R^{n-1}; \mathbb{R}^N)$ to $\widetilde{v} \in \mathrm{W}^{1,r}(B_{R+T}^{n-1} \times (0,T); \mathbb{R}^N)$ so that $\widetilde{v}$ is smooth on $B_{R+T}^{n-1} \times (0,T)$ and vanishes at $\partial B_{R+T}^{n-1} \times (0,T)$, and the estimate*

$$\int_{B_{R+T}^{n-1} \times (0,T)} \Phi(|\nabla \widetilde{v}(x,t)|)^q \, d(x,t) \le \left(\int_{B_R^{n-1}} A_0 \Phi(|v|) + A_1 \Phi(|\nabla v|)\, d\mathcal{H}^{n-1} \right)^q \tag{4.13}$$

holds for some constant $A_0, A_1 \ge 0$ (depending on T).

Then there exist $\varepsilon_0 > 0$ (depending on n) and $C = C(n,q)$ such that for any $0 < \varepsilon \le \varepsilon_0$, there is an extension operator sending $u \in \mathrm{W}^{1,r}(\partial B;\mathbb{R}^N)$ to $\widetilde{u} \in \mathrm{W}^{1,r}((1+\varepsilon)B\backslash\overline{B};\mathbb{R}^N)$ (in the sense that the trace of $\widetilde{u}$ on ∂B is equal to u) with the estimate

$$\int_{(1+\varepsilon)B\backslash B} \Phi(|\nabla\widetilde{u}|)^q\,dx \le C\left(\int_{\partial B} A_0\Phi(|u|) + A_1\Phi(|\nabla_\tau u|)\,d\mathcal{H}^{n-1}\right)^q. \tag{4.14}$$

Here $\nabla_\tau u$ denotes the tangential derivative of u on ∂B. Moreover, $\widetilde{u}$ is smooth on $(1+\varepsilon)B\backslash\overline{B}$.

Proof. It is not hard to see that for any point $x = (x_1,\dots,x_n)$ on the unit sphere $\partial B \subset \mathbb{R}^n$ has at least one coordinate x_i such that $|x_i| > \frac{1}{\sqrt{n+1}}$, hence ∂B is covered by $2n$ open sets of the form U_i, where

$$U_{2i} := \left\{x \in \partial B : x_i > \frac{1}{\sqrt{n+1}}\right\}, \quad U_{2i-1} := \left\{x \in \partial B : x_i < -\frac{1}{\sqrt{n+1}}\right\}.$$

We let (φ_i) be a smooth partition of unity subordinates to the cover $\{U_i\}$. It is easy to see that any $V \in \{U_i\}$ can be rotated to U_{2n} via some $R \in SO(n)$ so we shall focus our attention on this particular section of ∂B.

Let $U := \left\{x \in \partial B : x_n > \frac{1}{\sqrt{n+2}}\right\}$, clearly $U_{2n} \subset U$. Let

$$\varepsilon_0 := \sqrt{\frac{n+1}{n+2}} - \sqrt{\frac{n}{n+1}}, \qquad \Omega := B^{n-1}\left(0, \sqrt{\frac{n+1}{n+2}}\right)$$

and consider (for $0 < \varepsilon < \varepsilon_0$) the map $F\colon \Omega \times (-\varepsilon,\varepsilon) \to \mathbb{R}^n$ defined by

$$F(x',t) := \left(x', \sqrt{(1+t)^2 - |x'|^2}\right),$$

where we write $x' \in \mathbb{R}^{n-1}$ and $t \in \mathbb{R}$. We can see that F is a diffeomorphism onto its image $\Lambda := \mathrm{Im}(F)$ and let $G\colon \Lambda \to \Omega \times (-\varepsilon,\varepsilon)$ be its inverse, i.e. $G = F^{-1}$. Observe that the restriction of F to $\Omega \simeq \Omega \times \{0\}$, denoted by $f\colon \Omega \to \mathbb{R}^n$, is a parametrization of U (and $g = f^{-1}\colon U \to \Omega$ being its inverse), and that F maps a point $(x',t) \in \Omega \times (-\varepsilon,\varepsilon)$ to a point on $\partial B(0,1+t)$ whose first $n-1$ coordinates are x'.

Suppose that $R_i \in SO(n)$ is a rotation that maps U_i onto U_{2n} (we take R_{2n} to be the identity mapping) and write

$$G_i := G \circ R_i, \quad F_i := G_i^{-1} = R_i^{-1} \circ F, \quad f_i := F_i|_\Omega, \quad v_i := f_i^*(\varphi_i u).$$

Since $\varphi_i u$ is supported on U_i, we can see that $v_i = f_i^*(\varphi_i u)$ is supported on $\Omega' := B^{n-1}\left(0, \sqrt{\frac{n}{n+1}}\right) \subset \Omega$, hence it admits an extension to $\widetilde{v}_i \in \mathrm{W}^{1,r}(\Omega \times (0,\varepsilon);\mathbb{R}^N)$ according to our assumption. We claim that

$$\widetilde{u} := G_1^*\widetilde{v}_1 + \cdots + G_{2n}^*\widetilde{v}_{2n}$$

is the extension of u that we want. Writing $v = v_{2n}$, we will first show that $G^*\widetilde{v} = \widetilde{v} \circ G$ belong to $\mathrm{W}^{1,r}((1+\varepsilon)B\backslash\overline{B};\mathbb{R}^N)$ and satisfies the estimate (4.14).

Indeed, since $\widetilde{v}$ vanishes at $\partial B_{R+T} \times (0,T)$ (here $R = \sqrt{\frac{n}{n+1}}$ and $T = \varepsilon$), the function $G^*\widetilde{v}$ on Λ can be glued to the constant zero function on $\left((1+\varepsilon)B\backslash\overline{B}\right)\backslash\Lambda$, making it an element of $\mathrm{W}^{1,r}((1+\varepsilon)B\backslash\overline{B};\mathbb{R}^N)$. From the change of variable formula, we have

$$\begin{aligned}
\int_{(1+\varepsilon)B\backslash B} \Phi(|\nabla(G^*\widetilde{v})|)^q\,dx &= \int_\Lambda \Phi(|D(\widetilde{v}\circ G)|)^q\,dx \\
&\le \max_{\overline{\Lambda}} |DG|^{2q} \int_U \Phi(|D\widetilde{v}\circ G|)^q\,dx \\
&= \max_{\overline{\Lambda}} |DG|^{2q} \int_{\Omega\times(0,\varepsilon)} \Phi(|D\widetilde{v}|)^q |\det DF|\,d(x',t) \\
&\le \max_{\overline{\Lambda}} |DG|^{2q} \cdot \max_{\overline{\Omega\times(0,\varepsilon)}} |\det DF| \int_{\Omega\times(0,\varepsilon)} \Phi(|D\widetilde{v}|)^q\,d(x',t) \\
&\le C\left(\int_{\Omega'} A_0\Phi(|v|) + A_1\Phi(|\nabla_{x'}v|)\,dx'\right)^q,
\end{aligned}$$

where we used the estimate (4.10) from our assumption on the last line. To estimate the last term, we recall that

$$|\nabla_\tau(g^*v)(f(x'))|^2 = \sum_{1\le i,j<n} \gamma^{ij}(x')\partial_{x_i}v(x')\partial_{x_j}v(x'),$$

where $\gamma^{ij}(x')$ is an entry of the matrix $[\gamma^{ij}(x')] := [\gamma_{ij}(x')]^{-1}$, with $\gamma_{ij}(x') = \partial_{x_i}f(x')\cdot\partial_{x_j}f(x')$ (γ_{ij} is the pullback of the standard Riemannian metric on $\mathbb{R}^n$ by f). We can see that $[\gamma^{ij}(x')]$ is positive-definite for all $x' \in \Omega$, and hence there exists $m \ge 1$ such that $|\nabla_{x'}v(x')| \le m|\nabla_\tau(g^*v)(f(x'))|$ for all such x'. Thus

$$\begin{aligned}
\int_{\Omega'} \Phi(|\nabla_{x'}v|)\,d\mathcal{H}^{n-1} &\le m^\delta \int_{\Omega'} A_0\Phi(|(g^*v)\circ f|) + A_1\Phi(|\nabla_\tau(g^*v)\circ f|)\,dx' \\
&\le \max_{\overline{U}}\left(\frac{m^\delta}{Jf}\right)\int_U A_0\Phi(|g^*v|) + A_1\Phi(|\nabla_\tau(g^*v)|)\,d\mathcal{H}^{n-1},
\end{aligned}$$

where $Jf(x') = \sqrt{\det(Df(x')^tDf(x'))}$ is the Jacobian of f (direct calculation shows that $\min_{\overline{U}} Jf > 0$). Since $g^*v = \varphi_{2n}u$, we have

$$\begin{aligned}
\int_{(1+\varepsilon)B\backslash B} \Phi(|\nabla(G^*\widetilde{v})|)^q\,dx &\le C\left(\int_U A_0\Phi(|\varphi_{2n}u|) + A_1\Phi(|\nabla_\tau(\varphi_{2n}u)|)\,d\mathcal{H}^{n-1}\right)^q \\
&\le C\left(\int_U (A_0+A_1)\Phi(|u|) + A_1\Phi(|\nabla_\tau u|)\,d\mathcal{H}^{n-1}\right)^q.
\end{aligned} \tag{4.15}$$

Since $\Phi(\alpha z) \le \alpha^\delta\Phi(z)$, Jensen's inequality implies that

$$\begin{aligned}
\Phi(|\nabla\widetilde{u}|) &= \Phi(|\nabla(G_1^*\widetilde{v}_1)| + \cdots + |\nabla(G_{2n}^*\widetilde{v}_{2n})|) \\
&\le (2n)^{\delta-1}\left(\Phi(|\nabla(G_1^*\widetilde{v}_1)|) + \cdots + \Phi(|\nabla(G_{2n}^*\widetilde{v}_{2n})|)\right)
\end{aligned}$$

and since each R_i is an isometry, all the terms $\nabla(G_i^*\widetilde{v}_i)$ satisfy the same estimate (4.15) as $\nabla(G_1^*\widetilde{v}) = \nabla(G_1^*\widetilde{v}_{2n})$. Adding up all the estimates for each $\nabla(G_i^*\widetilde{v}_i)$ and replacing u with $u - \overline{u}$, where $\overline{u} = \fint_{\partial B} u \, d\mathcal{H}^{n-1}$, gives

$$\int_{(1+\varepsilon)B\setminus B} \Phi(|\nabla\widetilde{u}|)^q \, dx \leq C \left(\int_{\partial B} (A_0 + A_1)\Phi(|u - \overline{u}|) + A_1\Phi(|\nabla_\tau u|) \, d\mathcal{H}^{n-1} \right)^q$$

since shifting u by a constant does not change its gradient. Lastly, we apply the Poincaré inequality from Lemma 4.6 on the terms $\int_{\partial B} A_1\Phi(|u - \overline{u}|) \, d\mathcal{H}^{n-1}$ and use the fact that $\int_{\partial B} A_0\Phi(|u - \overline{u}|) \, d\mathcal{H}^{n-1} \leq 2^\delta \int_{\partial B} A_0\Phi(|u|) \, d\mathcal{H}^{n-1}$, which is a simple consequence of Jensen's inequality, to deduce that

$$\int_{(1+\varepsilon)B\setminus B} \Phi(|\nabla\widetilde{u}|)^q \, dx \leq C \left(\int_{\partial B} A_0\Phi(|u|) + A_1\Phi(|\nabla_\tau u|) \, d\mathcal{H}^{n-1} \right)^q. \quad \square$$

Next, we transfer the result from the unit ball $B = B^n(0,1)$ to the cube $Q = \left(-\frac{1}{2}, \frac{1}{2}\right)^n$.

Lemma 4.11. *Let $B = B^n(0,1)$, $Q = \left(-\frac{1}{2}, \frac{1}{2}\right)^n$, and $q, r \geq 1$. Let $\Phi\colon \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing convex function such that $\Phi(\alpha z) \leq \alpha^\delta \Phi(z)$ for each $\alpha \geq 1$. Suppose that the conclusions of Lemma 4.10 hold, i.e. there exist $\varepsilon > 0$, $A_0, A_1 \geq 0$, and an extension operator sending $v \in \mathrm{W}^{1,r}(\partial B; \mathbb{R}^N)$ to $\widetilde{v} \in \mathrm{W}^{1,r}((1+\varepsilon)B\setminus\overline{B}; \mathbb{R}^N)$ with the estimate*

$$\int_{(1+\varepsilon)B\setminus B} \Phi(|\nabla\widetilde{v}|)^q \, dx \leq \left(\int_{\partial B} A_0\Phi(|v|) + A_1\Phi(|\nabla_\tau v|) \, d\mathcal{H}^{n-1} \right)^q. \tag{4.16}$$

v is extended in such a way that $\widetilde{v}$ is smooth on $(1+\varepsilon)B\setminus\overline{B}$.

Then there exists a constant $C = C(n,q)$ and an extension operator sending $u \in \mathrm{W}^{1,r}(\partial Q; \mathbb{R}^N)$ to $\widetilde{u} \in \mathrm{W}^{1,r}((1+\varepsilon/2)Q\setminus\overline{Q}; \mathbb{R}^N)$ with the estimate

$$\int_{(1+\varepsilon/2)Q\setminus Q} \Phi(|\nabla\widetilde{u}|)^q \, dx \leq C \left(\int_{\partial Q} A_0\Phi(|u|) + A_1\Phi(|\nabla_\tau u|) \, d\mathcal{H}^{n-1} \right)^q. \tag{4.17}$$

Moreover, $\widetilde{u}$ is smooth on $(1+\varepsilon/2)Q\setminus\overline{Q}$ and $\widetilde{u} = 0$ outside $(1+\varepsilon/2)Q$ if $\widetilde{v} = 0$ outside $(1+\varepsilon)B$ for any $v \in \mathrm{W}^{1,r}(\partial B; \mathbb{R}^N)$.

Proof. Suppose that $\varepsilon > 0$ is the number given by Proposition 4.10. Observe that $\mathbb{R}^n$ can be written as a union of $2n$ cones $K_1, \ldots, K_{2n}$, which are defined as

$$K_{2i} := \left\{ x \in \mathbb{R}^n : \max_{j\neq i} |x_j| \leq x_i, x_i \geq 0 \right\}, K_{2i-1} := \left\{ x \in \mathbb{R}^n : \max_{j\neq i} |x_j| \leq -x_i, x_i \leq 0 \right\}.$$

We shall write $K := K_{2n}$.

Consider the mappings $T\colon K \to K$ defined as

$$\widetilde{T}(x',t) := A(x',t)(x',t), \quad \text{where} \quad A(x',t) := \frac{2t}{\sqrt{t^2 + |x'|^2}}$$

and let $R_i \in SO(n)$ be a rotation that maps K_i onto K then define $\widetilde{F}\colon \mathbb{R}^n \to \mathbb{R}^n$ to be

$$\widetilde{F}(x) = R_i^{-1} \circ T \circ R_i \quad \text{whenever } x \in K_i.$$

It can be verified that $F := \widetilde{F}\Big|_{(1+\varepsilon/2)Q\setminus Q}$ is a bijection onto $(1+\varepsilon)B\setminus B$, and that $f := \widetilde{F}\Big|_{\partial Q}$ is a bijection onto ∂B. In fact, F and f are both bi-Lipschitz continuous. We shall write $G := F^{-1}$ and $g := f^{-1}$.

Suppose that $u \in \mathrm{W}^{1,r}(\partial Q;\mathbb{R}^N)$ is given, then $v = g^*u \in \mathrm{W}^{1,r}(\partial B;\mathbb{R}^N)$ and we can extend v to $\widetilde{v} \in \mathrm{W}^{1,r}((1+\varepsilon)B\setminus\overline{B};\mathbb{R}^N)$ using our assumption. Letting $\widetilde{u} := F^*\widetilde{v}$, we can deduce that

$$\begin{aligned}
\int_{(1+\varepsilon/2)Q\setminus Q} \Phi(|\nabla\widetilde{u}|)^q\,dx &\le C\int_{(1+\varepsilon/2)Q\setminus Q} \Phi(|\nabla\widetilde{v}\circ F|)^q\,dx \\
&= C\int_{(1+\varepsilon)B\setminus B} \Phi(|\nabla\widetilde{v}|)^p|\det DG|\,dx \\
&\le C\left(\int_{\partial B} A_0\Phi(|v|) + A_1\Phi(|\nabla_\tau v|)\,d\mathcal{H}^{n-1}\right)^q \\
&\le C\left(\int_{\partial B} A_0\Phi(|v|) + A_1\Phi(|\nabla_\tau(f^*v)\circ g|)\,d\mathcal{H}^{n-1}\right)^q \\
&\le C\left(\int_{\partial Q} A_0\Phi(|f^*v|) + A_1\Phi(|\nabla_\tau(f^*v)|)\,d\mathcal{H}^{n-1}\right)^q \\
&= C\left(\int_{\partial Q} A_0\Phi(|u|) + A_1\Phi(|\nabla_\tau u|)\,d\mathcal{H}^{n-1}\right)^q,
\end{aligned}$$

following the same line of calculation as in Lemma 4.10.

Clearly, $\widetilde{u}$ is an extension of u. The fact that $\widetilde{u} \in \mathrm{W}^{1,r}((1+\varepsilon/2)Q\setminus\overline{Q};\mathbb{R}^N)$ follows from a standard technique (similar to the calculation above). The smoothness of $\widetilde{u}$ on $(1+\varepsilon/2)Q\setminus\overline{Q}$ is obvious since $\widetilde{v}$ is smooth on $(1+\varepsilon)B\setminus\overline{B}$. □

Now, we are ready for the proofs of Proposition 4.4 and Proposition 4.5.

Proof of Proposition 4.4. The estimate from Lemma 4.7 is used in Lemma 4.10 (taking $r = p$, $q = 1$, $A_0 = 4\varepsilon^\delta$ and $A_1 = \frac{8^\delta}{\varepsilon^{\delta-1}}$) and Lemma 4.11 then implies the result we want.

The fact that $\tilde{u} = 0$ outside $(1+\varepsilon)Q$ follows directly from the coordinate transformations used in Lemma 4.10 and Lemma 4.11. In particular, this means that $\tilde{u} \in \mathrm{W}^{1,p}(\mathbb{R}^n\setminus\overline{Q};\mathbb{R}^N)$. □

Proof of Proposition 4.5. This is a direct application of Lemma 4.9, Lemma 4.10 (taking $A_0 = 0$ and $A_1 = C^{1/q}$, where C is the constant appearing in the estimate (4.11) of Lemma 4.9) and Lemma 4.11 in succession. □

4.2 Nonlinear Gårding Inequality

In this section, we will establish the *nonlinear Gårding inequality* for the unperturbed variational integral

$$J(u) = \int_\Omega F(\nabla u)\,dx,$$

where F is a C^2 integrand satisfying certain additional assumptions, and Ω is a Lipschitz domain in $\mathbb{R}^n$. For each differentiable integrand G defined on some Euclidean space and a fixed w in its domain, we define the *shifted integrand* of G (also known as *Bregman divergence* or *Bregman distance*) G_w to be

$$\begin{aligned} G_w(z) &:= G(w+z) - G(w) - G'(w)[z] \\ &= \int_0^1 (1-t)G''(w+tz)[z,z]\,dt. \end{aligned}$$

For a fixed $u \in C^1(\Omega;\mathbb{R}^N)$, the nonlinear Gårding inequality states that

$$c\int_\Omega \Phi(|\nabla\varphi|)\,dx \le \int_\Omega F_{\nabla u}(\nabla\varphi)\,dx + C\int_\Omega \Phi(|\varphi|)\,dx, \tag{4.18}$$

holds for any $\varphi \in \mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^N)$. Note that the classical Gårding inequality is closely related to the case where F is a quadratic form and $\Phi(t) = t^2$. The term

$$\begin{aligned} \int_\Omega F_{\nabla u}[\nabla\varphi]\,dx &= \int_\Omega F(\nabla u + \nabla\varphi) - F(\nabla u) - F'(\nabla u)[\nabla\varphi]\,dx \\ &= J(u+\varphi) - J(u) - J'(u)[\varphi] \end{aligned}$$

reduces to a bilinear form in that case.

Ultimately, want to take $\Phi(|z|) = E_p(z)$ in (4.18) so the seemingly ad-hoc assumptions in this chapter are formulated with that goal in mind. In this thesis, we only deal with the case $p \in (1,\infty)$ but the analogous result for $p = 1$ is also true. The statement of the nonlinear Gårding inequality we shall prove first is as follows.

Proposition 4.12. *Let $\Phi\colon \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing convex function satisfying the p-growth condition such that for some given $\delta \ge 1$, $\Phi(\alpha t) \le \alpha^\delta(\Phi(t))$ for every $\alpha \ge 1$. Let $F\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ be a C^2 integrand and suppose that for any fixed $m > 0$ and $w, w_0 \in \mathbb{R}^{N\times n}$ such that $|w|, |w_0| \le m$, we can find corresponding constants l_m, C_m, C'_m and a non-decreasing, concave function $\omega_m : \mathbb{R}_+ \to [0,1]$ such that the following hold:*

- $\lim_{t\to 0}\omega_m(t) = 0$, $\omega_m(t) \le t$ *whenever* $t \le 1$, $\omega_m(t) = 1$ *for all* $t \ge 1$.

- *For any cube* $Q \subset \mathbb{R}^n$,
$$\int_Q F_w(\nabla\varphi)\,dx \geq l_m \int_Q \Phi(|\nabla\varphi|)\,dx. \tag{4.19}$$
- *For any* $z \in \mathbb{R}^{N\times n}$,
$$|F_w(z)| \leq C_m \Phi(|z|). \tag{4.20}$$
-
$$|F_w(z) - F_{w_0}(z)| \leq C'_m\,\omega_m(|w-w_0|)\Phi(|z|). \tag{4.21}$$

Let $u \in C^1(\overline{\Omega};\mathbb{R}^N)$*; then for* $m = \|\nabla u\|_\infty$ *there exists* $C = C(l_m, C_m, C'_m, n, \mu_{\nabla u})$ *such that*
$$\frac{l_m}{2}\int_\Omega \Phi(|\nabla\varphi|)\,dx \leq \int_\Omega F_{\nabla u}(\nabla\varphi)\,dx + C\int_\Omega \Phi(|\varphi|)\,dx \tag{4.22}$$
holds for all $\varphi \in \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)$*. Here* $\mu_{\nabla u}$ *is the modulus of continuity of* ∇u*.*

This proposition is inspired by the work of Campos Cordero and Kristensen that was mentioned before, where the conditions (4.19)–(4.21) presented here are extracted from some properties that $\Phi(t) = t^2$ satisfies. We will later show, in the next section, that $\Phi(|z|) = E_p(z)$ also satisfies (4.19)–(4.21) when $p > 1$, hence proving the desired Gårding inequality.

The rest of this section will be a series of lemmas that are used in proving Proposition 4.12. First, we require an extension operator that maps $u \in \mathrm{W}^{1,p}(\partial Q)$ to $\widetilde{u} \in \mathrm{W}^{1,p}((1+\varepsilon)Q\backslash Q)$ in such a way that $\int_{(1+\varepsilon)Q\backslash Q}\Phi(|\nabla\widetilde{u}|)\,dx$ can be controlled by $\int_{\partial Q}\Phi(|\nabla u|)\,d\mathcal{H}^{n-1}$. This was done in Proposition 4.4 and implies the following result.

Lemma 4.13. *Let* $m \in \mathbb{R}_+$*,* $w \in \mathbb{R}^{N\times n}$ *be such that* $|w| \leq m$*, and let* $F\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ *be a* C^2 *integrand satisfying*
$$\int_Q F_w(\nabla\varphi)\,dx \geq l\int_Q \Phi(|\nabla\varphi|)\,dx \tag{4.23}$$
for any cube $Q \subset \mathbb{R}^n$*, and*
$$|F_w(z)| \leq C_0\Phi(|z|). \tag{4.24}$$
Here $\Phi\colon \mathbb{R}_+ \to \mathbb{R}_+$ *is a convex function satisfying the conditions in Lemma 4.7. Then there exists a constant* $C = C(n, C_0, l, \delta) > 0$ *such that for any given* $\varepsilon > 0$ *and* $\varphi \in C^1(\overline{Q};\mathbb{R}^N)$*, we have the estimate*
$$\begin{aligned}\int_Q F_w(\nabla\varphi)\,dx \geq l\int_Q \Phi(|\nabla\varphi|)\,dx - \varepsilon C|Q|^{\frac{1}{n}}\int_{\partial Q}\Phi(|\nabla\varphi|)\,d\mathcal{H}^{n-1}\\ - \frac{C|Q|^{\frac{1}{n}}}{\varepsilon^{\delta-1}}\int_{\partial Q}\Phi(|Q|^{-\frac{1}{n}}|u|)\,d\mathcal{H}^{n-1}\end{aligned}$$
for all cube $Q \subset \mathbb{R}^n$*.*

Proof. It suffices to prove the claim only for $Q = \left(-\frac{1}{2}, \frac{1}{2}\right)^n$ since for any other $Q' = x_0 + rQ$ and $\varphi \in C^1(\overline{Q}'; \mathbb{R}^N)$ we may shift and rescale φ to $\phi(x) = \frac{1}{r}\varphi(x_0 + rx)$ (so that $\phi \in C^1(\overline{Q}; \mathbb{R}^N)$ and $\nabla\phi(x) = \nabla\varphi(x_0 + rx)$) to get the result for any cube Q'.

Let $\varepsilon > 0$ be given, $Q = \left(-\frac{1}{2}, \frac{1}{2}\right)^n$ and fix a function $\varphi \in C^1(\overline{Q}; \mathbb{R}^N)$. Define

$$\psi(x) = \begin{cases} \varphi(x) & ; x \in Q, \\ T(\varphi|_{\partial Q}) & ; x \in \mathbb{R}\backslash Q, \end{cases}$$

where T is the extension operator from Proposition 4.4, so $\psi \in \mathrm{W}_0^{1,p}((1+\varepsilon)Q)$. From (4.23) and (4.24), we can see that

$$\begin{aligned} \int_Q F_w(\nabla\varphi)\,dx &= \int_{(1+\varepsilon)Q} F_w(\nabla\psi)\,dx - \int_{(1+\varepsilon)Q\backslash Q} F_w(\nabla\psi)\,dx \\ &\geq l\int_{(1+\varepsilon)Q} \Phi(|\nabla\psi|)\,dx - C_0\int_{(1+\varepsilon)Q\backslash Q} \Phi(|\nabla\psi|)\,dx \\ &\geq l\int_Q \Phi(|\nabla\varphi|)\,dx - (C_0 - l)\int_{(1+\varepsilon)Q\backslash Q} \Phi(|\nabla\psi|)\,dx. \end{aligned}$$

We are done if $C_0 \leq l$, hence we shall assume further that $C_0 > l$. Since $\nabla\psi = \nabla T(\varphi|_{\partial Q})$ on $(1+\varepsilon)Q\backslash Q$, we can apply the estimate in Proposition 4.4 to get

$$\begin{aligned} \int_Q F_w(\nabla\varphi)\,dx \geq l\int_Q \Phi(|\nabla\varphi|)\,dx - \varepsilon C(C_0 - l)\int_{\partial Q} \Phi(|\nabla_\tau\varphi|)\,d\mathcal{H}^{n-1} \\ - \frac{C(n, \delta, C_0 - l)}{\varepsilon^{\delta-1}}\int_{\partial Q} \Phi(|\varphi|)\,d\mathcal{H}^{n-1}, \end{aligned}$$

and the desired estimate follows easily since $|\nabla_\tau\varphi| \leq |\nabla\varphi|$. □

The following lemma is not hard to establish.

Lemma 4.14. *Let $f \in C_c(\mathbb{R}^n)$ be a positive function and write $e = \frac{e_1 + \cdots + e_n}{\sqrt{n}}$, where $(e_i)_{i=1}^n$ are the standard basis of $\mathbb{R}^n$. Let $\mathcal{D}_k$ be the family of dyadic cubes corresponding to the lattice $2^{-k}\mathbb{Z}^n$, then for a dyadic cube $Q \in \mathcal{D}_k$ we have*

$$\int_0^{\sqrt{n}2^{-k}} \int_{re+\partial Q} f\,d\mathcal{H}^{n-1}\,dr \leq 2^n \int_{Q^*} f\,dx,$$

where Q^* is the union of Q and all of its right-adjacent cubes.

Remark 4.15. For $Q, R \in \mathcal{D}_k$, if $Q^* \cap R^* \neq \emptyset$, then Q and R are adjacent. It follows that

$$\sum_{Q \in \mathcal{D}_k} \chi_{Q^*} \leq 3^n. \tag{4.25}$$

//

Now we are ready to prove the promised Gårding inequality.

Proof of Proposition 4.12. For a cube in $\mathbb{R}^n$, we denote its center by x_Q. Let $\varphi \in C_c^1(\Omega;\mathbb{R}^N)$ and extend it to the whole $\mathbb{R}^n$ by setting $\varphi \equiv 0$ on $\mathbb{R}^n\backslash\Omega$. The function u is also extended to some $\widetilde{u}\colon \mathbb{R}^n \to \mathbb{R}^N$ so that $\widetilde{u}$ is C^1 and compactly supported. We shall denote $\widetilde{u}$ by u as well since its value outside Ω will not be relevant.

We set $m = \|\nabla u\|_\infty$ and let $\omega = \omega_{2m}$ be the modulus of continuity appeared in (4.21). Since Φ satisfies (4.19) and (4.20), we can apply the result from Lemma 4.13 with $r = |Q|^{\frac{1}{n}} \le 1$ to get

$$\begin{aligned}
\int_Q F_{\nabla u}(\nabla\varphi)\,dx &= \int_Q F_{\nabla u(x_Q)}(\nabla\varphi)\,dx + \int_Q F_{\nabla u}(\nabla\varphi) - F_{\nabla u(x_Q)}(\nabla\varphi)\,dx \\
&\ge l_m \int_Q \Phi(|\nabla\varphi|)\,dx - \varepsilon r C \int_{\partial Q} \Phi(|\nabla\varphi|)\,d\mathcal{H}^{n-1} \\
&\quad - \frac{C}{(\varepsilon r)^{\delta-1}} \int_{\partial Q} \Phi(|\varphi|)\,d\mathcal{H}^{n-1} - C\int_Q \omega(|\nabla u - \nabla u(x_Q)|)\Phi(|\nabla\varphi|)\,dx \\
&\ge \left(l_m - \omega\left(\mu\left(r\sqrt{n}\right)\right)\right) \int_Q \Phi(|\nabla\varphi|)\,dx \\
&\quad - \varepsilon r C \int_{\partial Q} \Phi(|\nabla\varphi|)\,d\mathcal{H}^{n-1} - \frac{C}{(\varepsilon r)^{\delta-1}} \int_{\partial Q} \Phi(|\varphi|)\,d\mathcal{H}^{n-1},
\end{aligned}$$

where $\mu = \mu_{\nabla u}$ is the modulus of continuity of ∇u, i.e.

$$\mu(t) := \sup\left\{|\nabla u(x) - \nabla u(y)| : x, y \in \mathbb{R}^n \text{ and } \ |x-y| \le t\right\}.$$

Since we extended u to a C_c^1 function, μ is an increasing function on $\mathbb{R}_+$ such that $\mu(0) = \lim_{t\to 0}\mu(t) = 0$ and $\mu(t) \le \operatorname{osc}(\nabla u) \le 2\max|\nabla u|$ for all $t \in \mathbb{R}_+$.

Let $\mathcal{D}_k$ be the family of dyadic cubes corresponding to the lattice $2^{-k}\mathbb{Z}^n$. Note that for any dyadic cube $Q \in \mathcal{D}_k$, $r = |Q|^{\frac{1}{n}} = 2^{-k}$, and for any $x \in \mathbb{R}\backslash\Omega$, $F_w(\nabla\varphi(x)) = F_w(0) = 0$. Hence for any $y \in \mathbb{R}^n$, we have

$$\begin{aligned}
\int_\Omega F_{\nabla u}(\nabla\varphi)\,dx &= \sum_{Q\in\mathcal{D}_k} \int_{y+Q} F_{\nabla u}(\nabla\varphi)\,dx \\
&\ge \left(l_m - \omega\left(\mu\left(\sqrt{n}2^{-k}\right)\right)\right) \int_\Omega \Phi(|\nabla\varphi|)\,dx \\
&\quad - \varepsilon 2^{-k} C \sum_{Q\in\mathcal{D}_k} \int_{y+\partial Q} \Phi(|\nabla\varphi|)\,d\mathcal{H}^{n-1} \\
&\quad - \frac{2^{k\delta}C}{\varepsilon^{\delta-1}} \sum_{Q\in\mathcal{D}_k} \int_{y+\partial Q} \Phi(|\varphi|)\,d\mathcal{H}^{n-1}.
\end{aligned}$$

We may now integrate the above over the diagonal $y \in [0, \sqrt{n}2^{-k}]e$, where e is defined as in Lemma 4.14, and apply the result from this lemma (with the following

remark) to get

$$\begin{aligned}
\sqrt{n}2^{-k}\int_\Omega F_{\nabla u}(\nabla\varphi)\,dx &= \sum_{Q\in\mathcal{D}_k}\int_{y+Q} F_{\nabla u}(\nabla\varphi)\,dx \\
&\geq \sqrt{n}2^{-k}\left(l_m-\omega\left(\mu\left(\sqrt{n}2^{-k}\right)\right)\right)\int_\Omega \Phi(|\nabla\varphi|)\,dx \\
&\quad -\varepsilon 2^{n-k}C\sum_{Q\in\mathcal{D}_k}\int_{Q^*}\Phi(|\nabla\varphi|)\,dx \\
&\quad -\frac{2^{n+k\delta}C}{\varepsilon^{\delta-1}}\sum_{Q\in\mathcal{D}_k}\int_{Q^*}\Phi(|\varphi|)\,dx. \\
&\geq \sqrt{n}2^{-k}\left(l_m-\omega\left(\mu\left(\sqrt{n}2^{-k}\right)\right)\right)\int_\Omega \Phi(|\nabla\varphi|)\,dx \\
&\quad -\varepsilon 2^{n-k}3^nC\int_\Omega\Phi(|\nabla\varphi|)\,dx \\
&\quad -\frac{2^{n+k\delta}3^nC}{\varepsilon^{\delta-1}}\int_\Omega\Phi(|\varphi|)\,dx.
\end{aligned}$$

We then divide both sides by $\sqrt{n}2^{-k}$ and deduce that

$$\begin{aligned}
\int_\Omega F_{\nabla u}(\nabla\varphi)\,dx \geq \left(l_m-\omega\left(\mu\left(\sqrt{n}2^{-k}\right)\right)-\varepsilon 6^nC/\sqrt{n}\right)\int_\Omega\Phi(|\nabla\varphi|)\,dx \\
-\frac{2^{n+k\delta}3^nC}{\sqrt{n}\varepsilon^{\delta-1}}\int_\Omega\Phi(|\varphi|)\,dx
\end{aligned}$$

holds for any $\varphi\in C_c^1(\Omega;\mathbb{R}^N)$. We then pick sufficiently large k and sufficiently small ε so that both $\omega\left(\mu\left(\sqrt{n}2^{-k}\right)\right)$ and $\varepsilon 6^nC/\sqrt{n}$ are both smaller that $l_m/4$.

Lastly, a standard approximation argument, together with the fact that Φ has p-growth and (4.20), allows us to extend the above estimate to any $\varphi\in \mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^N)$. This concludes the proof. □

4.3 The Case $\Phi(|z|)=E_p(z)$ for $p>1$

The goal of this section is to show that when $p>1$, $\Phi(|z|)=E_p(z)$ satisfies (4.19)–(4.21) from Proposition 4.12, and hence establishing the following result.

Proposition 4.16. *Let $F\colon\mathbb{R}^{N\times n}\to\mathbb{R}$ be a C^2 integrand satisfying (F_p1) and (F_p2). Let $u\in C^1(\overline{\Omega};\mathbb{R}^N)$, then there exist $\gamma=\gamma(l,\|\nabla u\|_\infty,p)$ and $C=C(k,p,l,n,\|\nabla u\|_\infty,\mu_{\nabla u})$ such that*

$$\gamma\int_\Omega E_p(\nabla\varphi)\,dx\leq\int_\Omega F_{\nabla u}(\nabla\varphi)\,dx+C\int_\Omega E_p(\varphi)\,dx \tag{4.26}$$

holds for all $\varphi\in \mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^N)$, where $\mu_{\nabla u}$ is the modulus of continuity of ∇u.

Recall that for a differentiable integrand G, the *shifted integrand* of G is

$$\begin{aligned} G_w(z) &= G(w+z) - G(w) - G'(w)[z] \\ &= \int_0^1 (1-t)G''(w+tz)[z,z]\,dt. \end{aligned}$$

In particular, we write

$$E_{p,w}(z) := E_p(w+z) - E_p(w) - E_p'(w)[z]$$

for the shifted integrand of E_p.

Remark 4.17. From the estimate (4.37) and (4.38) down below, the strong p-quasiconvexity (F_p2) implies that for any $w \in \mathbb{R}^{N\times n}$ and $\varphi \in \mathrm{W}^{1,\infty}_0(\Omega;\mathbb{R}^N)$, we have

$$\int_\Omega F(w+\nabla\varphi) - F(\nabla\varphi)\,dx \geq c\int_\Omega (1+|w|^2+|\nabla\varphi|^2)^{\frac{p-2}{2}}|\nabla\varphi|^2\,dx$$

whenever $p \in (1,2)$. Moreover, (F_p2) is in fact equivalent to

$$\int_\Omega F(w+\nabla\varphi) - F(\nabla\varphi)\,dx \geq c\int_\Omega (1+|w|^{p-2})|\nabla\varphi|^2 + |\nabla\varphi|^p\,dx$$

whenever $p \geq 2$. These are among the standard quasiconvexity assumptions in regularity theory for minimizers of the variational integral $u \mapsto \int_\Omega F(\nabla u)\,dx$. The reader is referred to [20], [17] and [36] for the case $p \in (1,2)$. Regarding the case $p \geq 2$, we refer to [40] and [48]. //

Remark 4.18. The reference integrand E_p is equivalent to $\Phi(z) = |V_p(z)|^2$, where $V_p(z) = (1+|z|^2)^{\frac{p-2}{4}}z$. In fact, calculation shows that

$$\Phi''(w)[z,z] = \langle w\rangle^{p-2}\left((2-p)\big((4-p)\theta^2-3\big)\frac{(w\cdot z)^2}{\langle w\rangle^2} + \big(2-(2-p)\theta^2\big)|z|^2\right),$$

where $\langle w\rangle = \sqrt{1+|w|^2}$ and $\theta = \frac{|w|}{\langle w\rangle}$. For $p \in (1,2)$, we can deduce the estimate

$$\begin{aligned} \Phi''(w)[z,z] &\geq (2-p)(4-p)\,\langle w\rangle^{p-2}\,|z|^2\left(\theta^4 - \frac{4\theta^2}{4-p} + \frac{2}{(4-p)(2-p)}\right) \\ &\geq \frac{2p}{4-p}\,\langle w\rangle^{p-2}\,|z|^2, \end{aligned}$$

using the fact that $\big((4-p)\theta^2-3\big) < 0$ and $\big(2-(2-p)\theta^2\big) > 0$, which also implies $\Phi''(w)[z,z] \leq p\,\langle w\rangle^{p-2}\,|z|^2$. For $p \geq 2$, similar computation (but simpler) shows that

$$2\,\langle w\rangle^{p-2}\,|z|^2 \leq \Phi''(w)[z,z] \leq \big((p-2)(p-1)+p\big)\,\langle w\rangle^{p-2}\,|z|^2,$$

hence $\Phi''(w)$ satisfies the same estimates as $E_p''(w)$ in Lemma 4.20. This shows that the conclusion of Proposition 4.22 also holds for Φ, so the integrands Φ and E_p

(including their shifted versions) are equivalent. The same computation also shows that E_p is equivalent to $|\cdot|^2 + |\cdot|^p$ for $p \geq 2$.

The function $V_p(z)$ as defined above has seen many usages in the literature, especially in the subquadratic growth case, playing similar role as our E_p. We refer to [36], [17], [2], and [18] for a more detailed discussion and usage of the V_p function. //

A lot of calculations involving E_p and its derivatives is needed in this section. To simplify some notations, we write

$$\langle z \rangle = \sqrt{1 + |z|^2}.$$

Lemma 4.19. *For each value of $r \in \mathbb{R}$ down below, there exists $c = c(r)$ and $C = C(r)$ such that for all $z, w \in \mathbb{R}^d$, we have*

1. *If $r > 0$, then*

$$c(|w|^r + |z|^r) \leq \int_0^1 (1-t)|w + tz|^r \, dt \leq C(|w|^r + |z|^r) \tag{4.27}$$

and

$$c(1 + |w|^r + |z|^r) \leq \int_0^1 (1-t) \langle w + tz \rangle^r \, dt \leq C(1 + |w|^r + |z|^r). \tag{4.28}$$

2. *If $r \in (-1, 0)$, then*

$$c(|w| + |z|)^r \leq \int_0^1 (1-t)|w + tz|^r \, dt \leq C \min\{|w|^r, |z|^r\} \tag{4.29}$$

and

$$c(1 + |w| + |z|)^r \leq \int_0^1 (1-t) \langle w + tz \rangle^r \, dt \leq C \min\{1, |w|^r, |z|^r\}. \tag{4.30}$$

Proof. First, let $r > 0$ be given. The second inequality in (4.27) follows easily from $|w + tz| \leq |w| + |z|$, so we only need to prove the first inequality. Suppose that $r \geq 2$, we write $q = 2/r$ then $t \mapsto t^{1/q}$ is convex, hence

$$\begin{aligned}
\int_0^1 (1-t)|w + tz|^r \, dt &= \int_0^1 \left((1-t)^q |w + tz|^2\right)^{\frac{1}{q}} \, dt \\
&\geq \left(\int_0^1 (1-t)^q |w + tz|^2 \, dt \right)^{\frac{1}{q}} \\
&\geq \left(\int_0^1 (1-t)^q (|w|^2 - 2t|w||z| + t^2|z|^2) \, dt \right)^{\frac{1}{q}} \\
&= (q+1)^{-\frac{1}{q}} \left(|w|^2 - \frac{2|w||z|}{(q+2)} + \frac{2|z|^2}{(q+2)(q+3)} \right)^{\frac{1}{q}}.
\end{aligned}$$

By Young's inequality, $2|w||z| \le \lambda|w|^2 + \frac{1}{\lambda}|z|^2$ and we can choose some $\lambda \in \left(\frac{q+3}{2}, q+2\right)$ so that $1 - \frac{\lambda}{q+2} = \frac{1}{q+2}\left(\frac{2}{q+3} - \frac{1}{\lambda}\right)$, hence

$$\left(|w|^2 - \frac{2|w||z|}{(q+2)} + \frac{2|z|^2}{(q+2)(q+3)}\right)^{\frac{r}{2}} \ge c(|w|^2 + |z|^2)^{\frac{r}{2}}$$

and we are done (for the case $r \ge 2$). For $r \in (0,2)$, we have

$$\begin{aligned}\int_0^1 (1-t)|w+tz|^r\,dt &= \int_0^1 (1-t)|w+tz|^{r-2}|w+tz|^2\,dt \\ &\ge (|w|+|z|)^{r-2}\int_0^1 (1-t)|w+tz|^2\,dt \\ &\ge c(|w|+|z|)^{r-2}(|w|^2+|z|^2)\end{aligned}$$

which implies the estimate we want.

To prove the inequalities (4.28), the second half is the direct consequence of $\langle w+tz\rangle \le 1+|w|+|z|$ so we only need to prove the first half. As for the first half, we use the fact that $\langle w+tz\rangle^r \ge c(1+|w+tz|^r)$ and apply the previous inequality (4.27) to get

$$\begin{aligned}\int_0^1 (1-t)\langle w+tz\rangle^r\,dt &\ge c\int_0^1 (1-t)(1+|w+tz|^r)\,dt \\ &\ge c(1+|w|^r+|z|^r).\end{aligned}$$

Now, for $r \in (-1,0)$, the first inequality in (4.29) is a trivial consequence of $|w+tz| \le |w|+|z|$. For the second inequality, we first assume, without loss of generality, that $w \ne 0$ (otherwise the inequality is trivial). Observe that $|w+tz| \ge \big||w| - t|z|\big|$ so if we write $\alpha = \frac{|z|}{|w|}$, then $|w+tz|^r \le |w|^r|1-\alpha t|^r$. Hence

$$\int_0^1 (1-t)|w+tz|^r\,dt \le |w|^r\int_0^1 (1-t)|1-\alpha t|^r\,dt = |w|^r I(\alpha),$$

where we denote $I(\alpha) := \int_0^1 (1-t)|1-\alpha t|^r\,dt$. By direct computation, we have

$$I(\alpha) = \frac{|1-\alpha|^{r+2} + \alpha(r+2) - 1}{\alpha^2(r+1)(r+2)},$$

where the right-handed side converges to $1/2$ as $\alpha \to 0$, agreeing with $I(0) = 1/2$. It is not hard to see that $I(\alpha)$ is bounded and that $I(\alpha) \le C|\alpha|^r$ for sufficiently large $|\alpha|$, thus $I(\alpha) \le C\min\{1, |\alpha|^r\}$. This implies that

$$\int_0^1 (1-t)|w+tz|^r\,dt \le C|w|^r \min\left\{1, \frac{|z|^r}{|w|^r}\right\} = C\min\left\{|w|^r, |z|^r\right\},$$

which proves (4.29).

Lastly, to prove (4.30), we observe that the first half follows directly from $\langle w+tz\rangle \le 1+|w|+|z|$. As for the second half, since $\langle w+tz\rangle \ge 1$ we thus have $\int_0^1 (1-t)\langle w+tz\rangle^r\, dt \le 1/2$ and since $\langle w+tz\rangle \ge |w+tz|$, we also have

$$\int_0^1 (1-t)\langle w+tz\rangle^r\, dt \le \int_0^1 (1-t)|w+tz|^r\, dt \le C\min\{|w|^r, |z|^r\}.$$

This implies that $\int_0^1 (1-t)\langle w+tz\rangle^r\, dt \le C\min\{1,|w|^r,|z|^r\}$ as required. □

Lemma 4.20. *For each $p>1$, E_p is a convex function whose second derivative satisfies the estimates*

$$p(p-1)\langle w\rangle^{p-2}|z|^2 \le E_p''(w)[z,z] \le p\langle w\rangle^{p-2}|z|^2 \quad \text{for } p\in(1,2) \tag{4.31}$$

and

$$p\langle w\rangle^{p-2}|z|^2 \le E_p''(w)[z,z] \le p(p-1)\langle w\rangle^{p-2}|z|^2 \quad \text{for } p\ge 2. \tag{4.32}$$

Proof. Using the fact that $\langle z\rangle'[u] = \langle z\rangle^{-1}(z\cdot u)$ and the chain rule, we can differentiate $E_p(z)=\langle z\rangle^p - 1$ to get

$$E_p'(z)[u] = p\langle z\rangle^{p-2}(z\cdot u),$$

and further differentiation yields

$$E_p''(z)[u,v] = p\langle z\rangle^{p-2}\left((u\cdot v)+\frac{p-2}{\langle z\rangle^2}(z\cdot u)(z\cdot v)\right).$$

For $p\in(1,2)$, we have

$$\begin{aligned} E_p''(w)[z,z] &= p\langle w\rangle^{p-2}\left(|z|^2-\frac{2-p}{\langle w\rangle^2}(w\cdot z)^2\right)\\ &\ge p\langle w\rangle^{p-2}|z|^2(1-(2-p)\theta)\\ &\ge p(p-1)\langle w\rangle^{p-2}|z|^2, \end{aligned}$$

where we write $\theta=\frac{|w|}{\langle w\rangle}\in[0,1)$. The upper bound for $E_p''(w)[z,z]$ can be obtained by discarding the term $\frac{2-p}{\langle w\rangle^2}(w\cdot z)^2$ in the first line of the above calculation.

For $p\ge 2$, we clearly have

$$\begin{aligned} E_p''(w)[z,z] &= p\langle w\rangle^{p-2}\left(|z|^2+\frac{p-2}{\langle w\rangle^2}(w\cdot z)^2\right)\\ &\ge p\langle w\rangle^{p-2}|z|^2. \end{aligned}$$

On the other hand, $\left(|z|^2+\frac{p-2}{\langle w\rangle^2}(w\cdot z)^2\right) \le |z|^2(1+(p-2)\theta) \le (p-1)|z|^2$, which implies the upper bound for $E_p''(w)[z,z]$.

The lower bounds for $E_p''(w)[z,z]$ in both cases imply that $E_p''(w)$ is positive definite for all $w\in\mathbb{R}^d$, hence E_p is convex. □

Remark 4.21. The above estimates are very similar to the estimates for $N_p(z) = |z|^p$, i.e.

$$p(p-1)|w|^{p-2}|z|^2 \leq N_p''(w)[z,z] \leq p|w|^{p-2}|z|^2 \quad \text{for } p \in (1,2) \tag{4.33}$$

and

$$p|w|^{p-2}|z|^2 \leq N_p''(w)[z,z] \leq p(p-1)|w|^{p-2}|z|^2 \quad \text{for } p \geq 2, \tag{4.34}$$

//

which can be deduced easily from the identity

$$N_p''(w)[z,z] = p(p-2)|w|^{p-4}(w \cdot z)^2 + p|w|^{p-2}|z|^2.$$

Proposition 4.22. *For each $M > 0$ and $p > 1$, there exist constants $c = c(M,p)$ and $C = C(M,p)$ such that whenever $|w| \leq M$, we have*

$$c\min\{|z|^2, |z|^p\} \leq E_{p,w}(z) \leq C\min\{|z|^2, |z|^p\} \qquad \text{for } p \in (1,2) \tag{4.35}$$

and

$$c\max\{|z|^2, |z|^p\} \leq E_{p,w}(z) \leq C\max\{|z|^2, |z|^p\} \qquad \text{for } p \geq 2. \tag{4.36}$$

Proof. Recall that we have

$$E_{p,w}(z) = \int_0^1 (1-t)E_p''(w+tz)[z,z]\,dt,$$

so we can combine Lemma 4.20 and Lemma 4.19 to get

$$c(1+|w|+|z|)^{p-2}|z|^2 \leq E_{p,w}(z) \leq C\min\{1, |w|^{p-2}, |z|^{p-2}\}|z|^2 \tag{4.37}$$

for $p \in (1,2)$ and

$$c(1+|w|^{p-2}+|z|^{p-2})|z|^2 \leq E_{p,w}(z) \leq C(1+|w|^{p-2}+|z|^{p-2})|z|^2 \tag{4.38}$$

for $p \geq 2$.

From the assumption that $|w| \leq M$, the above expressions can be simplified to

$$c(1+|z|)^{p-2}|z|^2 \leq E_{p,w}(z) \leq C\min\{|z|^2, |z|^p\} \tag{4.39}$$

for $p \in (1,2)$ and

$$c(|z|^2+|z|^p) \leq E_{p,w}(z) \leq C(|z|^2+|z|^p) \tag{4.40}$$

for $p \geq 2$. The desired conclusion can be drawn by inspecting the case $|z| \leq 1$ and $|z| > 1$ separately. □

Corollary 4.23. *For any $p > 1$ and $M > 0$, there exists a constant $c = c(M,p)$ such that*

$$E_{p,w}(z) \geq cE_p(z)$$

for all $w \in \mathbb{R}^d$ such that $|w| \leq M$.

Now, we are ready to prove that $\Phi(|z|) := E_p(z)$ satisfies the conditions (4.19)–(4.21) from Proposition 4.12. Regarding the condition (4.19), we observe that the strong p-quasiconvex assumption (F_p1) and Corollary 4.23 implies that for any given $m > 0$, there exists a constant $l_m = l_m(l, m, p)$ such that for $w \in \mathbb{R}^{N\times n}$ such that $|w| \leq m$, we have

$$\int_\Omega F_w(\nabla\varphi)\,dx \geq l\int_\Omega cE_{p,w}(\nabla\varphi)\,dx \geq l_m\int_\Omega E_p(\nabla\varphi)\,dx \tag{4.41}$$

for all $\varphi \in \mathrm{W}^{1,\infty}_0(\Omega;\mathbb{R}^N)$, so we are done. To show the other two conditions, we first recall that any quasiconvex function is rank-one convex and hence satisfies the assumptions of Lemma 3.13 from the previous chapter, which implies the following bounds and hence implies the condition (4.20).

Lemma 4.24. *Let $F\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ be a C^2 quasiconvex integrand satisfying (F_p1), then for each $m > 0$, there exists a constant $c = c(m)$ such that whenever $|w| \leq m$, we have*

$$|F_w(z)| \leq cE_p(z) \tag{4.42}$$

and

$$|F'_w(z)| \leq \begin{cases} c\min\{|z|, |z|^{p-1}\} & ; p \in (1, 2) \\ c\max\{|z|, |z|^{p-1}\} & ; p \geq 2. \end{cases} \tag{4.43}$$

Proof. Since F is rank-one convex, we can apply Lemma 3.13 to see that $|F'(z)| \leq K_1(1 + |z|^p)$ for all $z \in \mathbb{R}^{N\times n}$. We then let $w \in \mathbb{R}^{N\times n}$ be such that $|w| \leq m$ and let $K_2 := \max_{|z|\leq m+1} |F''(z)|$.

For $|z| \leq 1$, we have

$$|F_w(z)| = \frac{1}{2}|F''(w + tz)[z, z]| \leq \frac{1}{2}K_2|z|^2,$$

where in the above t is some number in $(0, 1)$. Similarly, we also have the estimate

$$|F'_w(z)| = |F''(w + tz)[z]| \leq K_2|z|.$$

For $|z| > 1$, we have

$$\begin{aligned} |F_w(z)| &\leq |F(z + w)| + |F(z)| + |F'(w)[z]| \\ &\leq 2k + k(|w + z|^p + |z|^p) + K_1(1 + |w|^{p-1})|z| \\ &\leq c|z|^p. \end{aligned}$$

For F'_w, we have

$$\begin{aligned} |F'_w(z)| &\leq |F'(w + z)| + |F'(w)| \\ &\leq K_1(1 + |w + z|^{p-1}) + K_1(1 + |w|^{p-1}) \\ &\leq c|z|^{p-1}. \end{aligned}$$

Combining the results from the cases $|z| \leq 1$ and $|z| > 1$ above, we may easily deduce (4.43) and

$$|F_w(z)| \leq \begin{cases} c\min\{|z|^2, |z|^p\} & ; p \in (1,2) \\ c\max\{|z|^2, |z|^p\} & ; p \geq 2, \end{cases}$$

which implies (4.42) according to Proposition 4.22. This completes the proof. □

The proof of the following estimate regarding the modulus of continuity of F'' can be found in the work by Campos Cordero and Kristensen [16] mentioned before.

Lemma 4.25. *Let $F\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ be a C^2 integrand and $m > 0$. Then there exists a non-decreasing, concave function $\omega := \omega_m : \mathbb{R}_+ \to [0,1]$ such that $\lim_{t\to 0} \omega(t) = 0$, $\omega(t) \leq t$ whenever $t \leq 1$, $\omega(t) = 1$ for all $t \geq 1$, and*

$$|F''(z) - F''(w)| \leq K\omega(|z-w|)$$

for all $z, w \in \mathbb{R}^{N\times n}$ with $|z|, |w| \leq m$. Here the constant $K = K(m)$ is defined to be $K := 2\max_{|z|\leq m} |F''(z)|$.

This estimate implies the next lemma, which shows that E_p satisfies condition (4.21).

Lemma 4.26. *Let $F\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ be a quasiconvex C^2 integrand satisfying (F_p1) and let $m \geq 1$. We take $\omega = \omega_{2m}$ to be the modulus of continuity from Lemma 4.25, then there exists $C = C(m)$ such that*

$$|F_w(z) - F_{w_0}(z)| \leq C\omega(|w-w_0|)E_p(z) \tag{4.44}$$

holds for all $z, w, w_0 \in \mathbb{R}^{N\times n}$ satisfying $|w|, |w_0| \leq m$.

Proof. For $|z| \leq m$, according to Lemma 4.25 we have

$$\begin{aligned} |F_w(z) - F_{w_0}(z)| &\leq \int_0^1 (1-t)|F''(w+tz) - F''(w_0+tz)||z|^2\,dt \\ &\leq K\omega(|w-w_0|)|z|^2. \end{aligned}$$

Now, for $|z| > m$, we let $C_0 = \sup_{B_{2m}} |F''|$ and apply Lemma 3.13 to get

$$\begin{aligned} |F_w(z) - F_{w_0}(z)| &\leq |F(w+tz) - F(w_0+tz)| + |F(w) - F(w_0)| + |F'(w) - F'(w_0)||z| \\ &\leq C(1 + |w+z|^{p-1} + |w_0+z|^{p-1})|w-w_0| \\ &\quad + K(1 + |w|^{p-1} + |w_0|^{p-1})|w-w_0| + C_0|w-w_0||z| \\ &\leq C(1+|z|^{p-1})|w-w_0| \\ &\leq C\omega(|w-w_0|)|z|^p. \end{aligned}$$

By choosing a sufficiently larger constant C, we may deduce that for $p \in (1,2)$ we have

$$|F_w(z) - F_{w_0}(z)| \leq C\omega(|w - w_0|) \min\{|z|^2, m^{2-p}|z|^p\}.$$

It is not hard to verify that $\min\{|z|^2, m^{2-p}|z|^p\} \leq m^{2-p} \min\{|z|^2, |z|\}$ whenever $p \in (1,2)$. In the case $p \geq 2$, we have

$$|F_w(z) - F_{w_0}(z)| \leq C\omega(|w - w_0|) \max\{|z|^2, m^{2-p}|z|^p\}.$$

and the inequality $\max\{|z|^2, m^{2-p}|z|^p\} \leq \max\{|z|^2, |z|\}$ instead. Therefore we can conclude, using Lemma 4.22, that

$$|F_w(z) - F_{w_0}(z)| \leq C\omega(|w - w_0|)E_p(z)$$

in both cases. □

Now that we have shown all the required conditions needed to apply Proposition 4.12, the Gårding inequality follows easily as a consequence.

Proof of Proposition 4.16. We let $\Phi(t) = (1+t^2)^{\frac{p}{2}} - 1$ and observe that $E_p(z) = \Phi(|z|)$. We can see that for all $\alpha \geq 1$, $\Phi(\alpha t) \leq \alpha^2 \Phi(t)$ if $p \in (1,2)$ and $\Phi(\alpha t) \leq \alpha^p \Phi(t)$ if $p \geq 2$.

According to (4.41), ($F_p 2$) implies (4.19). The assumptions (4.20) and (4.21) follow from (4.42) in Lemma 4.24 and Lemma 4.26, respectively. Hence we can apply Proposition 4.12 with $\Phi(t) = (1+t^2)^{\frac{p}{2}} - 1$ to deduce the estimate (4.26). □

Remark 4.27. The constant γ and C in the statement of Proposition 4.16 depends on $\|\nabla u\|_\infty$ and $\mu_{\nabla u}$, hence the conclusion also holds for all u in a fixed subset $\mathcal{S} \subset \mathrm{W}^{1,p}(\Omega; \mathbb{R}^{N\times n})$ provided that $\|\nabla u\|_\infty$'s are uniformly bounded and that ∇u's are equicontinuous for all $u \in \mathcal{S}$. In particular, this is true if $\mathcal{S}$ is bounded in some $C^{1,\alpha}$ norm (with $0 < \alpha < 1$). //

4.4 Regularized Solutions

Unfortunately, our minimizer $u \in \mathrm{W}^{1,p}_g(\Omega; \mathbb{R}^N)$ of the variational integral

$$J(u) = \int_\Omega F(\nabla u)\, dx,$$

where $F \in C^2(\mathbb{R}^{N\times n})$ is a (strongly) quasiconvex integrand with p-growth, need not be in $C^1(\overline{\Omega}; \mathbb{R}^N)$ even when g is very regular. However, we can modify our function J by adding a regularizing term $\varepsilon|\nabla^2 u|^2$ and try to find a minimizer of

$$J_\varepsilon(u) = \int_\Omega F(\nabla u) + \frac{\varepsilon}{2}|\nabla^2 u|^2\, dx, \tag{4.45}$$

in $\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$ instead. The integral makes sense as a nonlinear functional $J_\varepsilon\colon \mathrm{W}^{2,2}(\Omega;\mathbb{R}^N)\to\mathbb{R}$ provided that F satisfies the p-growth (F_p1) with exponent $p\in\left(1,\frac{2n}{n-2}\right]$ as a consequence of the Sobolev embedding theorem.

The addition of a small higher order term to fix the issue of regularity has been studied in calculus of variations, for example in nonlinear elasticity [74]. In related context, the Γ-convergence of similar functionals as $\varepsilon\to 0$ has also been studied, e.g. in the theory of phase transition [27], [25], [9], [22]. Regarding the addition of terms with even higher order derivatives to the functional $J(u)$, we refer to [18].

From the Sobolev embedding theorem, we know that such a solution $u\in\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$ is necessarily an $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ function for any $p<2^*=\frac{2n}{n-2}$, so the integral of $F(\nabla u)$ makes sense for all such p. It turns out that our minimizer u belongs to a much better space than $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$, thanks to the fact that the Euler-Lagrange equation associated with J_ε is elliptic.

In order to show that a minimizer of J_ε is regular, we will rely on a result regarding higher order elliptic equation [3, Theorem 8.1] by Agmon, which directly implies the following proposition when we take the elliptic differential operator to be the bi-Laplacian $\Delta^2=\sum_{1\le i,j\le n}\partial_i^2\partial_j^2$.

Proposition 4.28. *Let $\Omega\subset\mathbb{R}^n$ be a bounded domain whose boundary is of class C^4. Let $u\in L^q(\Omega)$, $q>1$, be a function such that there exists a constant $\lambda>0$ so that for any $\varphi\in C^4(\overline{\Omega})$ with $\varphi=0$ and $\nabla\varphi=0$ on $\partial\Omega$, there holds*

$$\left|\int u(x)\Delta^2\varphi(x)\,dx\right|\le\lambda\,\|\varphi\|_{\mathrm{W}^{4-j,p}(\Omega)} \tag{4.46}$$

for some $1\le j\le 4$. Then $u\in\mathrm{W}^{j,p'}(\Omega)$, where $\frac{1}{p}+\frac{1}{p'}=1$, and there exists a $C=C(n,p)>0$ such that

$$\|u\|_{\mathrm{W}^{j,p'}(\Omega)}\le C\left(\lambda+\|u\|_{\mathrm{L}^{p'}(\Omega)}\right). \tag{4.47}$$

Using the above result, we can prove that a minimizer u of J_ε in $\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$ is actually regular up to the boundary of Ω, provided that the boundary datum g is sufficiently smooth. For any $q\in[1,\infty)$, we shall write $q':=\frac{q}{q-1}$ and $q^*:=\frac{nq}{n-q}$.

Theorem 4.29. *Let $\Omega\subset\mathbb{R}^n$ be a bounded domain with C^4 boundary. Let $p\in(1,\frac{2n}{n-2})$, $F\colon\mathbb{R}^{N\times n}\to\mathbb{R}$ be a C^1 quasiconvex integrand with p-growth, and let g be a given function that belongs to $\mathrm{W}^{3,r}(\Omega;\mathbb{R}^N)$ for some $r\ge p'$. Then the problem of minimizing $J_\varepsilon\colon\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)\to\mathbb{R}$, where J_ε is the variational integral in (4.45), has a solution and any extremal u of J_ε in $\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$ also belongs to $\mathrm{W}^{3,r}(\Omega;\mathbb{R}^N)$ with the estimate*

$$\|u\|_{\mathrm{W}^{3,r}}\le C\left(\frac{1}{\varepsilon}\,\|F'(\nabla u)\|_{\mathrm{L}^r}+\|g\|_{\mathrm{W}^{3,r}}+\|u\|_{\mathrm{L}^r}\right) \tag{4.48}$$

for some $C=C(n,p)$. The above estimate also holds for any $0<\tilde{r}\le r$. Moreover, if $r>\frac{n}{2}$, then $u\in C^{1,\alpha}(\overline{\Omega})$ for any $\alpha\in(0,\min\{1,2-\frac{n}{r}\})$.

Proof. Since $\nabla u \mapsto F(\nabla u)$ is quasiconvex and $\nabla^2 u \mapsto |\nabla^2 u|^2$ is convex, J_ε is sequentially weakly semicontinuous in $\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$, hence a standard application of the direct method implies that J_ε has a minimizer in $\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$. Let $u \in \mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$ be a minimizer of J_ε and define

$$\mathcal{D} := \{\varphi \in C^\infty(\Omega)^N : \varphi = 0 \text{ and } \nabla\varphi = 0 \text{ on } \partial\Omega\},$$

then for any $\varphi \in \mathcal{D}$, $u + \varphi$ is in $\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$ and hence $J(u) \le J(u + \delta\varphi)$ for all $\delta \in \mathbb{R}$. The function u is thus a critical point of J', and hence satisfies the Euler-Lagrange equation $0 = -\operatorname{div} F'(\nabla u) + \varepsilon\Delta^2 u$ weakly, i.e.

$$0 = \int_\Omega F'(\nabla u) \cdot \nabla\varphi + \varepsilon\nabla^2 u : \nabla^2\varphi \, dx \tag{4.49}$$

for all $\varphi \in \mathcal{D}$.

Let us consider $v = u - g$. By definition, v belongs to $\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$, which means that $v = 0$ and $\nabla v = 0$ on $\partial\Omega$ (in the trace sense). For any given $\varphi \in \mathcal{D}$, we can perform integration by parts to get

$$\begin{aligned}\int_\Omega v\Delta^2\varphi \, dx &= \int_\Omega \nabla^2 v : \nabla^2\varphi \, dx \\ &= \int_\Omega \nabla^2 u : \nabla^2\varphi \, dx - \int_\Omega \nabla^2 g : \nabla^2\varphi \, dx \\ &= -\frac{1}{\varepsilon}\int_\Omega F'(\nabla u) \cdot \nabla\varphi \, dx + \int_\Omega \operatorname{div}(\nabla^2 g) \cdot \nabla\varphi \, dx. \end{aligned} \tag{4.50}$$

Since F is rank-one convex, $|F'(V)| \le c(1 + |V|^{p-1})$ so $F'(V) \in \mathrm{L}^{\frac{q}{p-1}}(\Omega)$ whenever $V \in \mathrm{L}^q(\Omega;\mathbb{R}^{N\times n})$. By taking $q = p$, we see that $F'(\nabla u)$ is in $\mathrm{L}^{p'}$ so

$$\left|\int_\Omega v\Delta^2\varphi \, dx\right| \le \left(\frac{1}{\varepsilon}\left\|F'(\nabla u)\right\|_{\mathrm{L}^{p'}} + \|g\|_{\mathrm{W}^{3,p'}}\right)\|\varphi\|_{\mathrm{W}^{1,p}},$$

and hence we can apply Proposition 4.28 (to each of the component v_i of v and φ_i of φ) with $j = 3$. This shows that v belongs to $\mathrm{W}^{3,p'}(\Omega;\mathbb{R}^N)$ with the estimate

$$\|v\|_{\mathrm{W}^{3,p'}} \le C\left(\frac{1}{\varepsilon}\left\|F'(\nabla u)\right\|_{\mathrm{L}^{p'}} + \|g\|_{\mathrm{W}^{3,p'}} + \|v\|_{\mathrm{L}^{p'}}\right),$$

where C depends only on n and p, which readily implies that $u \in \mathrm{W}^{3,p'}(\Omega;\mathbb{R}^N)$ and

$$\|u\|_{\mathrm{W}^{3,p'}} \le C\left(\frac{1}{\varepsilon}\left\|F'(\nabla u)\right\|_{\mathrm{L}^{p'}} + \|g\|_{\mathrm{W}^{3,p'}} + \|u\|_{\mathrm{L}^{p'}}\right). \tag{4.51}$$

By the standard Sobolev embedding theorem (see [41, 5.6.3], for example), $\mathrm{W}^{3,p'}(\Omega;\mathbb{R}^N)$ embeds continuously into $\mathrm{W}^{1,q}(\Omega;\mathbb{R}^N)$ for all $q \in [1, (p')^{**}]$, where $(p')^{**} = \frac{np}{n(p-1)-2p}$ (provided that the expression is a positive real number, otherwise $q \in [1,\infty)$), hence $\nabla u \in \mathrm{L}^{(p')^{**}}(\Omega;\mathbb{R}^{N\times n})$. More generally, if we consider (4.50)

with $\nabla u \in \mathrm{L}^q(\Omega;\mathbb{R}^{N\times n})$, we have $F'(\nabla u) \in \mathrm{L}^{\frac{q}{p-1}}(\Omega;\mathbb{R}^{N\times n})$ so the same computation would give

$$\|u\|_{\mathrm{W}^{3,\frac{q}{p-1}}} \leq C\left(\frac{1}{\varepsilon}\|F'(\nabla u)\|_{\mathrm{L}^{\frac{q}{p-1}}} + \|g\|_{\mathrm{W}^{3,\frac{q}{p-1}}} + \|u\|_{\mathrm{L}^{\frac{q}{p-1}}}\right). \tag{4.52}$$

in place of (4.51) (where we used $q = p$), as long as $\frac{q}{p-1} \leq r$. Using Sobolev embedding theorem as before, the bootstrapping argument allow us to get $\nabla u \in \mathrm{L}^{(\frac{q}{p-1})^{**}}(\Omega;\mathbb{R}^{N\times n})$ from the assumption that $\nabla u \in \mathrm{L}^q$ (and sufficiently large r). More precisely, we define an increasing function $f\colon \mathbb{R}_+ \to \overline{\mathbb{R}}_+$ by

$$f(q) := \begin{cases} (\frac{q}{p-1})^{**} = \frac{nq}{n(p-1)-2q} & ; q < \frac{1}{2}n(p-1) \\ \infty & ;\ \text{otherwise} \end{cases}$$

and see that $\nabla u \in \mathrm{L}^q(\Omega;\mathbb{R}^{N\times n})$ implies $u \in \mathrm{W}^{3,t}(\Omega;\mathbb{R}^N)$ for all $t \leq \min\{r, \frac{q}{p-1}\}$ and hence $\nabla u \in \mathrm{L}^s(\Omega;\mathbb{R}^{N\times n})$ for all $s \leq \min\{r^{**}, f(q)\}$ (or simply $s < \infty$ if $r^{**} \notin \mathbb{R}_+$).

By direct computation, f is a convex function whose only non-trivial fixed point is $\tilde{q} = \frac{1}{2}(p-2)n$, provided that this expression is greater than zero, so the iteration $q_{j+1} := f(q_j)$ with $q_0 = p$ produces a sequence that goes to infinity if $q_0 > \tilde{q}$, which is equivalent to $p < 2^* = \frac{2n}{n-2}$. This means that we may keep iterating until $\frac{q_j}{p-1} > r$ to conclude that $u \in \mathrm{W}^{3,r}(\Omega;\mathbb{R}^N)$ with the required estimate (4.49), which is obtained by substituting q in (4.52) with $(p-1)r$.

From the discussion regarding equation (4.52), (4.50) also holds if we replace r with $\tilde{r} = \frac{q}{p-1}$ for some $q < r(p-1)$. The last claim is straightforward since $\mathrm{W}^{3,r}(\Omega;\mathbb{R}^N)$ continuously embeds into $C^{1,\alpha}(\overline{\Omega};\mathbb{R}^N)$ for any $\alpha \in (0, \min\{1, 2-\frac{n}{r}\})$, provided that $r > \frac{n}{2}$. □

As we have seen, the regularizing term $\frac{\varepsilon}{2}|\nabla^2 u|^2$ makes a minimizer $u \in \mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$ of J_ε as regular as g in the class $\mathrm{W}^{3,r}(\Omega;\mathbb{R}^N)$, provided some conditions on the exponent p, r are satisfied. It is worth noting that if we assume $g \in C^3(\overline{\Omega};\mathbb{R}^N)$, then u belongs to $C^{2,\alpha}(\overline{\Omega};\mathbb{R}^N)$ for any $\alpha \in (0,1)$.

Since $|F'(\nabla u)| \lesssim 1 + |\nabla u|^{p-1}$, for any $0 < s \leq r$ we have $\|F'(\nabla u)\|_{\mathrm{L}^s} \lesssim 1 + \|\nabla u\|_{\mathrm{L}^{s(p-1)}}$. If $p \geq 2$, equation (4.48) implies that

$$\begin{aligned} \|u\|_{\mathrm{W}^{3,s}} &\leq C\left(1 + \|\nabla u\|_{\mathrm{L}^{s(p-1)}} + \|u\|_{\mathrm{L}^s}\right) \\ &\leq C\left(1 + \|u\|_{\mathrm{W}^{1,s(p-1)}}\right), \end{aligned} \tag{4.53}$$

where the constant C now also depends on ε and $\|g\|_{\mathrm{W}^{3,s}}$. Assuming that $r > \frac{n}{2}$, any $\delta > 0$ that satisfies $s = \frac{n}{2} + \frac{\delta}{p-1} \leq r$ gives the estimate

$$\|u\|_{C^{1,\alpha}} \leq C\left(1 + \|u\|_{\mathrm{W}^{1,\frac{n(p-1)}{2}+\delta}}\right). \tag{4.54}$$

By the standard embedding theorem, we have $\mathrm{W}^{2,2}(\Omega;\mathbb{R}^N) \hookrightarrow \mathrm{W}^{1,\frac{n(p-1)}{2}+\delta}(\Omega;\mathbb{R}^N)$ for some sufficiently small $\delta > 0$ precisely when $p < \frac{n+2}{n-2}$. This discussion leads to the following corollary.

Corollary 4.30. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with C^4 boundary, $p \in [2, \frac{2n}{n-2})$, let $F\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ be a C^1 quasiconvex integrand with p-growth. Let $g \in \mathrm{W}^{3,r}(\Omega;\mathbb{R}^N)$ where $r \geq p'$ and $r > \frac{n}{2}$, and let u be an extremal of J_ε in $\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$. Then for $0 < \delta \leq (p-1)\left(r - \frac{n}{2}\right)$, we have the estimate*

$$\|u\|_{C^{1,\alpha}} \leq C\left(1 + \|u\|_{\mathrm{W}^{1,\frac{n(p-1)}{2}+\delta}}\right) \tag{4.55}$$

for a suitable $\alpha = \alpha(n,p,\delta)$ and some constant $C = C(n,p,\varepsilon,\|g\|_{\mathrm{W}^{3,r}})$.

Moreover, if we assume further that $p < \frac{n+2}{n-2}$, then we have the estimate

$$\|u\|_{C^{1,\alpha}} \leq C\left(1 + \|u\|_{\mathrm{W}^{2,2}}\right) \tag{4.56}$$

for suitable $\alpha = \alpha(n,p)$ and $C = C(n,p,\varepsilon,\|g\|_{\mathrm{W}^{3,r}})$.

Note that the condition $2 \leq p < \frac{n+2}{n-2}$ can only be satisfied if $n < 6$.

4.5 Partial Convexity and the Sets of Extremals of J_ε

In this section, we will prove our main results (Theorem 4.1 and Theorem 4.2) with the help of the nonlinear E_p Gårding inequality.

First, we show that it is possible to strengthen the inequality (4.26) to

$$\int_\Omega E_p(\nabla\varphi)\,dx \leq C\int_\Omega F_{\nabla u}(\nabla\varphi)\,dx, \tag{4.57}$$

provided that we impose an additional restriction that φ belongs to a certain subspace Y of $\mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^{N\times n})$ of finite codimension. This means that we have strict convexity of the functional $\varphi \mapsto \int_\Omega F_{\nabla u}(\nabla\varphi)\,dx$ on X, which could be of independent interest. A key step to prove this result is the following proposition.

Proposition 4.31. *Let $\Omega \subset \mathbb{R}^n$ be a Lipschitz domain and let $2 \leq p < \infty$. For any given $M > 0$, there exists a closed linear subspace $Y \subset \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)$ such that Y has finite codimension and that*

$$\int_\Omega E_p(\varphi)\,dx \leq \frac{1}{M}\int_\Omega E_p(\nabla\varphi)\,dx$$

for every $\varphi \in Y$.

The proof of a similar result in the work by Campos Cordero and Kristensen is based on the decomposition of $\mathrm{W}^{1,2}_0(\Omega;\mathbb{R}^N)$ using an orthogonal basis of the eigenfunction of the Dirichlet Laplacian operator, which is not directly compatible with the reference integrand E_p. The method we will use here for proving the superquadratic ($p \geq 2$) case employs a certain wavelets-type Schauder basis of $\mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^{N\times n})$ instead of the eigenbasis used in the work by Campos Cordero and

Kristensen. The main references for the construction of such a basis is [103] and [102].

According to [103, Theorem 2.38], there exists an orthogonal basis $\{e_k\}_{k=1}^{\infty}$ of $\mathrm{L}^2(\Omega)$ that is universal in the sense that it is also simultaneously a Schauder basis for $F_{p,q}^{s,rloc}(\Omega)$, the Triebel–Lizorkin space with the refined localization property. On a Lipschitz domain $\Omega \subset \mathbb{R}^n$, we have $\mathrm{W}_0^{1,p}(\Omega) = \widetilde{F}_{p,2}^{1}(\Omega) = F_{p,2}^{1,rloc}(\Omega)$, where $\widetilde{F}_{p,q}^{k}(\Omega)$ is the homogeneous Triebel–Lizorkin space (see [103, Remark 2.17 and Theorem 4.30]). The generalization to the vector-valued space $\mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^N)$ is straightforward, which leads to the following proposition.

Proposition 4.32 (Triebel)**.** *Let $\Omega \subset \mathbb{R}^n$ be a Lipschitz domain. Then there exists a wavelet orthogonal basis $\{e_k\}_{k=1}^{\infty}$ of $\mathrm{L}^2(\Omega;\mathbb{R}^N)$ such that $|\operatorname{supp} e_k| \to 0$ as $k \to \infty$ and $\{e_k\}_{k=1}^{\infty}$ is also a Schauder basis of the space $\mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^N)$ for and $1 < p < \infty$, i.e. if $f \in \mathrm{W}_0^{1,p}(\Omega;\mathbb{R}^N)$, then f can be uniquely represented as*

$$f = \sum_{k=1}^{\infty} c_k(f) e_k, \tag{4.58}$$

where the series converges in the $\mathrm{W}^{1,p}$ norm.

Note that the series (4.58) converges unconditionally and that the linear functional $e_k^* \colon f \mapsto c_k(f)$ is continuous. The sequence $\{e_k^*\}_{k=1}^{\infty}$ is called the *biorthogonal sequence* associated with the Schauder basis $\{e_k\}_{k=1}^{\infty}$.

Now we are ready to prove Proposition 4.31.

Proof of Proposition 4.31. Let $\{e_k\}_{k=1}^{\infty}$ be the basis given by Proposition 4.32 and for an arbitrary $q > 1$, let $X^{(q)} := \mathrm{W}_0^{1,q}(\Omega;\mathbb{R}^N)$, $X_k^{(q)} := \overline{\operatorname{span}}^{\mathrm{W}^{1,q}}\{e_k, e_{k+1}, \dots\} \subset X$, i.e. $X_k^{(q)}$ is the closure in $\mathrm{W}^{1,q}$ norm of $\operatorname{span}\{e_k, e_{k+1}, \dots\}$. In this proof, we use the norm $\|\varphi\|_{\mathrm{W}_0^{1,q}} := \left(\int_\Omega |\nabla\varphi|^q \, dx\right)^{1/q}$ for simplicity.

First, we claim that for any $M > 0$, we can find $k = k(M) \in \mathbb{N}$ such that

$$\inf\left\{ \int_\Omega |\nabla\varphi|^q \, dx - M \int_\Omega |\varphi|^q \, dx \ : \ \varphi \in X_k^{(p)}, \|\varphi\|_{\mathrm{W}_0^{1,q}} = 1 \right\} > 0. \tag{4.59}$$

Assume that the above is false, then there exists $M > 0$ and a sequence $\varphi_{k_j} \in X_{k_j}^{(q)}$ (where $k_j \to \infty$) such that $\left\|\varphi_{k_j}\right\|_{\mathrm{W}_0^{1,q}} = 1$ for all $j \in \mathbb{N}$ and that

$$\lim_{j\to\infty} \int_\Omega |\nabla\varphi_{k_j}|^q \, dx - M \int_\Omega |\varphi_{k_j}|^q \, dx \le 0. \tag{4.60}$$

We know that $\int_\Omega |\nabla\varphi_{k_j}|^q \, dx = 1$ for all j by our assumption. On the other hand, since $q > 1$, the space $X^{(q)}$ is reflexive, hence by [54, Theorem 1] the basis $\{e_k\}_{k=1}^{\infty}$ is *shrinking* in the sense that $\lim_{k\to\infty} \left\| f|_{X_k^{(q)}} \right\| = 0$ for any linear functional $f \in (X^{(q)})^*$

(note that this statement corresponds to the fact that the span of the biorthogonal sequence $\{e_k^*\}_{k=1}^{\infty}$ is norm dense in $(X^{(q)})^*$). In particular, this implies that as $j \to \infty$, $\varphi_{k_j} \overset{w}{\rightharpoonup} 0$ in $\mathrm{W}_0^{1,q}(\Omega;\mathbb{R}^N)$ and hence $\varphi_{k_j} \to 0$ in L^q (by Rellich-Kondrachov theorem). However, this would imply that we have $1 - 0 \leq 0$ in (4.60), a contradiction.

Let $R^p \colon X^{(p)} \backslash \{0\} \to \mathbb{R}_+$ be the generalized p-Rayleigh quotient defined by

$$R^p[\varphi] := \frac{\int_\Omega E_p(\nabla\varphi)\,dx}{\int_\Omega E_p(\varphi)\,dx} \tag{4.61}$$

and let M_k be the infimum of R^p on $X_k^{(p)} \backslash \{0\}$, i.e.

$$M_k := \inf\left\{ R^p[\varphi] \;:\; \varphi \in X_k^{(p)} \backslash \{0\} \right\}. \tag{4.62}$$

Note that this proposition is proved if we can show that $M_k \to \infty$ as $k \to \infty$.

Now, since $p \geq 2$, we can see from (4.40) that $E_p(z) \sim |z|^2 + |z|^p$, thus

$$R^p[\varphi] \geq C\,\frac{\int_\Omega |\nabla\varphi|^2\,dx + \int_\Omega |\nabla\varphi|^p\,dx}{\int_\Omega |\varphi|^2\,dx + \int_\Omega |\varphi|^p\,dx}$$

for some generic constant $C > 0$. Using the inequality $\frac{A_1+B_1}{A_2+B_2} \geq \min\left(\frac{A_1}{A_2}, \frac{B_1}{B_2}\right)$ (which holds for $A_1, B_1, A_2, B_2 > 0$), we have for any fixed $\varphi \in X_k^{(p)} \backslash \{0\}$

$$R^p[\varphi] \geq C \min\left(\frac{\int_\Omega |\nabla\varphi|^2\,dx}{\int_\Omega |\varphi|^2\,dx}, \frac{\int_\Omega |\nabla\varphi|^p\,dx}{\int_\Omega |\varphi|^p\,dx} \right). \tag{4.63}$$

According to (4.59),

$$\lim_{k\to\infty} \inf\left\{ \frac{\int_\Omega |\nabla\varphi|^q\,dx}{\int_\Omega |\varphi|^q\,dx} \;:\; \varphi \in X_k^{(q)} \backslash \{0\} \right\} = \infty \tag{4.64}$$

for any $q > 1$. Since $p \geq 2$, we have $X_k^{(p)} \subset X_k^{(2)}$ as a linear subspace, thus

$$\inf\left\{ \frac{\int_\Omega |\nabla\varphi|^2\,dx}{\int_\Omega |\varphi|^2\,dx} \;:\; \varphi \in X_k^{(2)} \backslash \{0\} \right\} \leq \inf\left\{ \frac{\int_\Omega |\nabla\varphi|^2\,dx}{\int_\Omega |\varphi|^2\,dx} \;:\; \varphi \in X_k^{(p)} \backslash \{0\} \right\}.$$

We can then apply (4.64) with $q = 2$ and $q = p$ together with (4.63) to conclude that $M_k \to \infty$ as $k \to \infty$. This completes the proof. □

Recall that we can write $\int_\Omega F_{\nabla u}(\nabla\varphi)\,dx$ as

$$\int_\Omega F(\nabla u + \nabla\varphi) - F(\nabla u) - F'(\nabla u)[\nabla\varphi]\,dx = J(u+\varphi) - J(u) - J'(u)[\varphi].$$

As a consequence of the previous proposition, we have the following important estimate. Note that as a consequence of Proposition 4.16, the assumptions of the following proposition is satisfied when $F \colon \mathbb{R}^{N\times n} \to \mathbb{R}$ satisfies (F_p1) and (F_p2); and $\mathcal{S} \subset C^1(\overline{\Omega};\mathbb{R}^N) \subset \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ is a family of functions such that $\{\nabla u : u \in \mathcal{S}\}$ is uniformly bounded (in L^∞ norm) and equicontinuous.

Proposition 4.33. *Let $p \geq 2$, $S \subset \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$, and $F\colon \mathbb{R}^{N\times n} \to \mathbb{R}$ be an integrand with the property that there exist $l, C \geq 0$ such that for any $u \in \mathcal{S}$, the Gårding inequality*

$$\gamma \int_\Omega E_p(\nabla\psi)\,dx \leq \int_\Omega F_{\nabla u}(\nabla\psi)\,dx + C\int_\Omega E_p(\psi)\,dx \tag{4.65}$$

holds for all $\psi \in \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)$. Then there exists closed subspaces V, Y of $\mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)$ (depending on the constants γ, C) such that $\mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N) = V \oplus Y$, V is finite dimensional, and

$$\int_\Omega E_p(\nabla\varphi)\,dx \leq \frac{2^{2p-1}}{\gamma(2^{2p-1}-1)}\int_\Omega F_{\nabla u}(\nabla\varphi)\,dx \tag{4.66}$$

for any $u \in \mathcal{S}$ and any $\varphi \in Y$.

Moreover, for any $u_\alpha, u_\beta \in \mathcal{S}$, their decompositions

$$u_\alpha = g + v_\alpha + \varphi_\alpha, \quad u_\beta = g + v_\beta + \varphi_\beta, \quad \text{where} \quad v_\alpha, v_\beta \in V \quad \text{and} \quad \varphi_\alpha, \varphi_\beta \in Y,$$

satisfy the estimate

$$\begin{aligned}\int_\Omega E_p(\nabla\varphi_\alpha - \nabla\varphi_\beta)\,dx \leq \frac{2^{2p-1}C}{\gamma}\int_\Omega E_p(v_\alpha - v_\beta)\,dx + 2^p\int_\Omega E_p(\nabla v_\alpha - \nabla v_\beta)\,dx \\ - \frac{2^p}{\gamma}\big(J'(u_\beta)[u_\alpha - u_\beta] + J'(u_\alpha)[u_\beta - u_\alpha]\big).\end{aligned} \tag{4.67}$$

Proof. From Proposition 4.31, there exists a closed subspace $Y \subset \mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)$ such that

$$\int_\Omega E_p(\varphi)\,dx \leq \frac{\gamma}{2^{2p-1}C}\int_\Omega E_p(\nabla\varphi)\,dx \tag{4.68}$$

for all $\varphi \in Y$. We then let V be a complementary subspace of Y in $\mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N)$. From this, (4.66) follows immediately and any $u_\alpha, u_\beta \in \mathcal{S}$ thus admit the decompositions as in the statement of this proposition.

Using the notations from (4.65), we set $u := u_\beta$ and $\psi := u_\alpha - u_\beta$ (so that $u + \psi = u_\alpha$). The Gårding inequality states that

$$\gamma\int_\Omega E_p(\nabla u_\alpha - \nabla u_\beta)\,dx \leq J(u_\alpha) - J(u_\beta) - J'(u_\beta)[u_\alpha - u_\beta] + C\int_\Omega E_p(u_\alpha - u_\beta)\,dx.$$

By symmetry, we may reverse the role of u_α and u_β in the above inequality and conclude that

$$\begin{aligned}\gamma\int_\Omega E_p(\nabla u_\alpha - \nabla u_\beta)\,dx \leq C\int_\Omega E_p(u_\alpha - u_\beta)\,dx \\ - \big(J'(u_\beta)[u_\alpha - u_\beta] + J'(u_\alpha)[u_\beta - u_\alpha]\big).\end{aligned} \tag{4.69}$$

Recall that $E_p(z+w) \le 2^{p-1}E_p(z) + 2^{p-1}E_p(w)$, hence

$$\int_\Omega E_p(\nabla\varphi_\alpha - \nabla\varphi_\beta)\,dx \le 2^{p-1}\int_\Omega E_p(\nabla v_\alpha - \nabla v_\beta) + 2^{p-1}\int_\Omega E_p(\nabla u_\alpha - \nabla u_\beta).$$

This and (4.69) imply

$$\begin{aligned}\frac{\gamma}{2^{p-1}}\int_\Omega E_p(\nabla\varphi_\alpha - \nabla\varphi_\beta)\,dx &\le \gamma\int_\Omega E_p(\nabla v_\alpha - \nabla v_\beta) + C\int_\Omega E_p(u_\alpha - u_\beta)\,dx \\ &\quad - \big(J'(u_\beta)[u_\alpha - u_\beta] + J'(u_\alpha)[u_\beta - u_\alpha]\big) \\ &\le \gamma\int_\Omega E_p(\nabla v_\alpha - \nabla v_\beta) + 2^{p-1}C\int_\Omega E_p(v_\alpha - v_\beta)\,dx \\ &\quad + 2^{p-1}C\int_\Omega E_p(\varphi_\alpha - \varphi_\beta)\,dx \\ &\quad - \big(J'(u_\beta)[u_\alpha - u_\beta] + J'(u_\alpha)[u_\beta - u_\alpha]\big).\end{aligned}$$

Using (4.68), the third term can be absorbed into the left-hand side, which yields

$$\begin{aligned}\frac{\gamma}{2^p}\int_\Omega E_p(\nabla\varphi_\alpha - \nabla\varphi_\beta)\,dx &\le \gamma\int_\Omega E_p(\nabla v_\alpha - \nabla v_\beta) + 2^{p-1}C\int_\Omega E_p(v_\alpha - v_\beta)\,dx \\ &\quad - \big(J'(u_\beta)[u_\alpha - u_\beta] + J'(u_\alpha)[u_\beta - u_\alpha]\big).\end{aligned}$$

□

The estimate (4.67) simplifies if $J'(u_\beta)[u_\alpha - u_\beta] + J'(u_\alpha)[u_\beta - u_\alpha] \ge 0$. In particular, if both u_α and u_β are extremals of the functional J, then we have

$$\int_\Omega E_p(\nabla\varphi_\alpha - \nabla\varphi_\beta)\,dx \le \frac{2^{2p-1}C}{\gamma}\int_\Omega E_p(v_\alpha - v_\beta)\,dx + 2^p\int_\Omega E_p(\nabla v_\alpha - \nabla v_\beta)\,dx. \tag{4.70}$$

Another important case is when u_α and u_β are extremals of the perturbed functional J_ε, in which $J'(u_\alpha)[\psi] = -2\varepsilon\int_\Omega \nabla^2 u\colon \nabla^2\psi\,dx$ for any $\psi \in \mathrm{W}^{2,2}(\Omega;\mathbb{R}^N)$ (similar equation holds for u_β). In this case, we have an even stronger statement:

$$J'(u_\beta)[u_\alpha - u_\beta] + J'(u_\alpha)[u_\beta - u_\alpha] = 2\int_\Omega |\nabla^2 u_\alpha - \nabla^2 u_\beta|^2\,dx. \tag{4.71}$$

The discussion here and the results above lead to our main theorem.

Proof of Theorem 4.1: From the assumptions, Corollary 4.30 implies that there exists $\alpha > 0$ such that $\mathcal{S} \subset C^{1,\alpha}(\overline{\Omega};\mathbb{R}^N)$ and $\mathcal{S}$ is bounded in the $C^{1,\alpha}$ norm. This means the set $\{\nabla u : u \in \mathcal{S}\}$ is uniformly bounded in the $\mathrm{L}^\infty(\Omega;\mathbb{R}^{N\times n})$ norm and is equicontinuous, hence we can apply Proposition 4.33 to get the decomposition $\mathrm{W}^{1,p}_0(\Omega;\mathbb{R}^N) = V \oplus Y$, where $\dim V = d < \infty$. We choose $v_1,\dots,v_d \in V$ so that $V = \operatorname{span}\{v_1,\dots,v_d\}$.

Since $\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N) \subset \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$, any $u \in \mathcal{S}$ can be written uniquely as

$$u = g + v + \varphi \quad \text{where} \quad v \in V \ \text{ and } \ \varphi \in Y.$$

From (4.71), we know that $J'(u_\beta)[u_\alpha - u_\beta] + J'(u_\alpha)[u_\beta - u_\alpha] \geq 0$ hence (4.67) implies that for any $u, u' \in \mathcal{S}$ must be equal whenever their components in V coincide. This means that the map $\pi_V \colon V \oplus Y \to V$ defined by $\pi_V(v+\varphi) = v$ is injective on $\mathcal{S} - g$.

The fact that $\mathcal{S}$ is bounded in $\mathrm{W}^{1,\frac{n(p-1)}{2}+\delta}$ norm implies (by (4.53)) that $\mathcal{S}$ is bounded in $\mathrm{W}^{2,2}$, and hence bounded in $\mathrm{W}^{1,p}$. Therefore the set $\pi_V(\mathcal{S} - g)$ is bounded in V and its element is of the form

$$v_\alpha = \alpha_1 v_1 + \cdots + \alpha_d v_d$$

for some unique $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$. The map $I \colon V \to \mathbb{R}^d$ defined by $I(v_\alpha) = \alpha$ is a continuous linear isomorphism and $K = I \circ \pi_V(\mathcal{S} - g)$ is a bounded set in $\mathbb{R}^d$. Each $u = g + v + \varphi \in \mathcal{S}$ can thus be parametrized by $\alpha \in K$, and we may define $u_\alpha \in \mathcal{S}$ and $\varphi_\alpha \in Y$ to be the corresponding functions such that $I \circ \pi_V(u_\alpha - g) = \alpha$ and $\varphi_\alpha := \pi_Y(u_\alpha)$, i.e. $u_\alpha = g + v_\alpha + \varphi_\alpha$.

Now, for any $\alpha, \beta \in K$, we define $u_\alpha = g + v_\alpha + \varphi_\alpha$ and $u_\beta = g + v_\beta + \varphi_\beta$ as above. To get the estimate for the components $\varphi_\alpha, \varphi_\beta \in Y$, first we recall that $E_p(z) \sim |z|^2 + |z|^p$ since $p \geq 2$, thus

$$\begin{aligned}
&\int_\Omega |\nabla\varphi_\alpha - \nabla\varphi_\beta|^2 \, dx + \int_\Omega |\nabla^2\varphi_\alpha - \nabla^2\varphi_\beta|^2 \, dx \\
&\qquad \lesssim \int_\Omega E_p(\nabla\varphi_\alpha - \nabla\varphi_\beta) \, dx + \int_\Omega |\nabla^2\varphi_\alpha - \nabla^2\varphi_\beta|^2 \, dx \\
&\qquad \lesssim \int_\Omega E_p(\nabla\varphi_\alpha - \nabla\varphi_\beta) \, dx + \int_\Omega |\nabla^2 u_\alpha - \nabla^2 u_\beta|^2 + |\nabla^2 v_\alpha - \nabla^2 v_\beta|^2 \, dx \\
&\qquad \lesssim \int_\Omega E_p(v_\alpha - v_\beta) \, dx + \int_\Omega E_p(\nabla v_\alpha - \nabla v_\beta) \, dx + \int_\Omega |\nabla^2 v_\alpha - \nabla^2 v_\beta|^2 \, dx,
\end{aligned}$$

where the last inequality is obtained by applying (4.71) and (4.67). We then take the square root of both sides and, again, use the fact that $E_p(z) \sim |z|^2 + |z|^p$ to get

$$\|\varphi_\alpha - \varphi_\beta\|_{\mathrm{W}^{2,2}_0} \lesssim \|v_\alpha - v_\beta\|_{\mathrm{W}^{2,2}_0} + \|v_\alpha - v_\beta\|^{\frac{p}{2}}_{\mathrm{W}^{1,p}_0}.$$

Since V is finite dimensional, all norms on V are equivalent so $\|v_\alpha - v_\beta\|_{\mathrm{W}^{2,2}_0} \sim \|v_\alpha - v_\beta\|_{\mathrm{W}^{1,p}_0} \sim |\alpha - \beta|$. Since K is bounded and $\frac{p}{2} \geq 1$, we have $|\alpha - \beta|^{\frac{p}{2}} \leq C|\alpha - \beta|$ for some constant C depending on K, therefore

$$\|\varphi_\alpha - \varphi_\beta\|_{\mathrm{W}^{2,2}_0} \lesssim |\alpha - \beta|. \tag{4.72}$$

This means that the map $\alpha \mapsto \varphi_\alpha$ is Lipschitz, which readily implies that the map $f \colon K \to \mathcal{S}$ defined by $f(\alpha) = u_\alpha$ is a bi-Lipschitz equivalence, where $K \subset \mathbb{R}^d$ is endowed with the usual metric and $\mathcal{S}$ is viewed as a subset of $\mathrm{W}^{2,2}(\Omega; \mathbb{R}^N)$. □

As a direct consequence of this theorem, we get a statement regarding the minimizers of J_ε if we assume a tighter condition on p.

Proof of Theorem 4.2. From (4.56), we see that $\mathcal{M}$ automatically satisfies the assumptions for $\mathcal{S}$ in 4.1 if $\mathcal{M}$ is bounded in the $\mathrm{W}^{2,2}$ norm. Indeed, we know that $\mathcal{M} \subset \{u \in \mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N) : J_\varepsilon(u) = L\}$ where $L \in \mathbb{R}$ is the infimum of J_ε on $\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$. Assume, for contradiction, that there exists a sequence $u_n \in \mathcal{M}$ such that $\|u_n\|_{\mathrm{W}^{2,2}} \to \infty$ as $n \to \infty$ while keeping

$$\int_\Omega F(\nabla u_n)\,dx + \varepsilon \int_\Omega |\nabla^2 u_n|^2\,dx = L.$$

The term $\int_\Omega |\nabla^2 u_n|^2\,dx$ cannot be bounded otherwise it would imply the boundedness of u_n in the Dirichlet class $\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$, so we must have $\int_\Omega F(\nabla u_n)\,dx \to -\infty$. However, according to [21, Theorem 1.1], the assumption that F is strongly quasiconvex implies that $J(u) = \int_\Omega F(\nabla u)\,dx$ is L^p-mean coercive, so in particular J is bounded below. This is the contradiction we wanted, hence $\mathcal{M}$ is bounded in the norm of $\mathrm{W}^{2,2}(\Omega;\mathbb{R}^N)$.

All the conclusions then follow from Theorem 4.1 except the fact that K is compact. In order to see that, observe that $K = I \circ \pi_V(\mathcal{M} - g)$ so it is enough to prove that $\mathcal{M}$ is compact in $\mathrm{W}^{1,p}$ norm. However, J_ε is sequentially weakly lower semicontinuous in $\mathrm{W}^{2,2}(\Omega;\mathbb{R}^N)$ so $\mathcal{M}$ is weakly compact in $\mathrm{W}^{2,2}(\Omega;\mathbb{R}^N)$. This implies compactness in $\mathrm{W}^{1,p}$ norm and we are done. □

5 Further Developments and Related Results

In this final chapter, we shall explore the possibility of extending certain aspects of the work we did earlier, especially the well-posedness results from chapter 3 and the structure theorems from chapter 4, and discuss some related results that could be useful for this purpose. In particular, we want to employ the (generalized) Hadamard well-posedness results from chapter 3 to prove a generic uniqueness of minimizers for the problem

$$\text{minimize} \int_\Omega F(x, u, \nabla u)\, dx \quad \text{over} \quad u \in \mathrm{W}^{1,p}_g(\Omega; \mathbb{R}^N),$$

with varying Dirichlet data $g \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$. There are some difficulties associated with this goal, and we hope that the knowledge of the structure of the set of minimizers for J_ε, namely the d-rectifiability of the set of minimizers (that we proved in chapter 4), may help us prove the generic uniqueness of minimizers for the singularly perturbed problem

$$\text{minimize} \int_\Omega F(\nabla u)\, dx + \varepsilon \int_\Omega |\nabla^2 u|^2\, dx \quad \text{over} \quad u \in \mathrm{W}^{2,2}_g(\Omega; \mathbb{R}^N),$$

with varying Dirichlet data $g \in \mathrm{W}^{2,2}(\Omega; \mathbb{R}^N)$, which should be an easier problem.

There are some results regarding the generic single-valuedness of multivalued functions that suggest the possibility of expending the generalized Hadamard well-posedness result from chapter 3 to a generic uniqueness theorem for the minimization problems of the variational integral $J(u) = \int_\Omega F(x, u, \nabla u)\, dx$ in the Dirichlet class $\mathrm{W}^{1,p}_g(\Omega; \mathbb{R}^N)$. The semicontinuity property in the multivalued functions sense of the solution map $\mathcal{S}\colon \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N) \rightrightarrows \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$ is one requirement, whereas the other requirement is the so-called 'minimality' property. The work regarding this minimality property has not been fully worked out but there are some evidences that such property could hold in our setting of quasiconvex variational integrals, or at least under some suitable additional assumptions. In this last chapter, we shall outline some relevant results in this direction and suggest some possible strategies that could lead to the desired minimality result.

Beside some compelling results from convex analysis and the theory of multivalued function, there is also evidence from nearby fields that also support our hypothesis on the generic uniqueness of the solutions to the Dirichlet quasiconvex variational problems. While not directly applicable to our problem, the work on generic uniqueness of area-minimizing surfaces by Morgan indicates that our goal is not entirely unfounded. In [77], Morgan showed that almost every curve in $\mathbb{R}^3$ bounds a unique area-minimizing surface. He equipped the space of functions corresponding to the Dirichlet boundary data with a suitable measure and proved that the set of all boundary data that admit non-unique solutions is a null set

with respect to this measure (in fact, his proof also implies that this set is also a meagre set). He later showed that the result generalizes to minimization problems for elliptic variational integrands with constant coefficients in [78]. Despite these generic uniqueness results, it is worth noting that an arbitrary Dirichlet problem for minimal surfaces could have a continuum of solutions as demonstrated by Morgan [76] (see also [106] by White for a general obstruction to uniqueness of minimal surfaces).

Some of the key results regarding the generic single-valuedness of multivalued functions that could be useful for our purpose are from the paper [63] by Lassonde, which describe an abstract approach to proving generic uniqueness of a special class of multivalued functions. They can be seen as generalizations of results involving the single-valuedness of monotone operators, e.g. [83] by Phelps or [84] by Preiss and Zajíček. We hope that such results, directly as stated or suitably modified, could be applied to our quasiconvex variational problems. There are other results in this direction as well that could be fruitful for our research, in particular the work [23] by Choban, Kenderov and Revalski, and [55] by Kenderov and Moors.

An important terminology used in [63] (and some related papers such as [64]) is the *d-fragmentability*, which is defined as follows.

Definition 5.1 (d-fragmentability)**.** *Let (X, τ) be a topological space, (Y, d) a metric space. A multifunction $T\colon X \rightrightarrows Y$ is said to be* fragmented by d *if for any nonempty open set U in X and any $\varepsilon > 0$, there exists another nonempty open set $V \subset U$ such that* $\operatorname{diam} T(V) < \varepsilon$. //

The following is the first abstract generic uniqueness result based on fragmentability.

Proposition 5.2 (Lemma 1, [63])**.** *Let (X, τ) be a Baire space, (Y, d) be a metric space, and $T\colon X \rightrightarrows Y$ be a nonempty-valued mapping fragmented by d. Then, there exists a dense G_δ subset D of X such that at each point of D, T is single-valued and τ-d-USC.*

The above result is used to prove the following proposition.

Proposition 5.3 (Proposition 2, [63])**.** *Let (X, τ) be a Baire space, Y be a Banach space, and $T\colon X \rightrightarrows Y^*$ be minimal τ-w^*-CUSCO. If there is a set $C \subset Y^*$ with w^*-RNP such that $\{x \in X : T(x) \cap C \neq \emptyset\}$ is dense in X, then there is a dense G_δ set $D \subset X$ such that T is single-valued and τ-$\|\cdot\|$-USC on D. Moreover, we have $T(D) \subset \overline{C}$.*

For our purpose, we are interested in the case where $(X, \tau) = (\mathrm{W}^{1,p}(\Omega; \mathbb{R}^N), \|\cdot\|)$, $Y = \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)^*$ (so that $Y^* = \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$ by reflexivity), and $T = \mathcal{S}$, the

solution map

$$\mathcal{S} = \{u \in \mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N) : J(u) = \inf_{\mathrm{W}^{1,p}_g(\mathbb{R}^N)} J\}.$$

Note that in our case the weak* topology on Y^* is the weak topology of $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$. It is easy to see that the single-valuedness of $\mathcal{S}$ on $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ is a restatement of the generic uniqueness of minimizers of the problem $\min(\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N); J)$ with varying Dirichlet data g from the class $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$.

Certainly, if our solution map $\mathcal{S}$ has all the above properties in Proposition 5.3 then we are done. However, the assumptions that we need to impose on the integrand $F(x,u,z)$ would be too strong in order to get all the corresponding properties for $\mathcal{S}$. The p-growth condition (F1), the pointwise coercivity condition (F2) and the uniform strict quasiconvexity (F3') for the integrand F and the coercivity condition (J1) for the functional J that we have made so far only guarantee that $\mathcal{S}$ is $\|\cdot\|$-$\|\cdot\|$-USCO, but $\mathcal{S}$ is certainly not convex-valued and it is not clear if $\mathcal{S}$ is a minimal USCO map or not. For this reason, we may wish to consider the (closed) convex hulls of $\mathcal{S}$ or related functions.

Definition 5.4 (Closed Convex Hull)**.**

- *For a multivalued function $T\colon X \rightrightarrows Y$, the* closed convex hull of T *is map $\overline{\operatorname{conv}} T$ defined by*
$$(\overline{\operatorname{conv}} T)(x) := \overline{\operatorname{conv}} T(x).$$
- *For a function $J\colon X \to \overline{\mathbb{R}}$, its closed convex hull $\overline{\operatorname{conv}} J$ is the function whose epigraph is the closed convex hull of the epigraph of J.* //

Note that for a proper $J\colon X \to \overline{\mathbb{R}}$, we have $\overline{\operatorname{conv}} J = J^{**}|_X$ (see [12]).

Suppose we want to apply Proposition 5.3 directly, there are a few possible strategies that could be used to modify the multivalued function $\mathcal{S}$ so that the modified version satisfies the assumptions in Proposition 5.3 while allowing the generic single-valuedness to be transferred back to our original multivalued function $\mathcal{S}$. Three types of different modifications and associated strategies are as follow.

1. Suppose that we can prove directly that $T := \overline{\operatorname{conv}} \mathcal{S}$ is minimal τ-w^*-CUSCO then we are done.

2. If we can prove that $\mathcal{S}$ is in fact minimal $\|\cdot\|$-w-USCO, then if automatically follows that $T := \overline{\operatorname{conv}} \mathcal{S}$ is minimal $\|\cdot\|$-w-CUSCO and we are done.

3. Consider the multivalued function $\mathcal{S}'$ defined by
$$\mathcal{S}'(g) := \operatorname{argmin}(\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N); \overline{\operatorname{conv}} J),$$
if we can show that $\mathcal{S}'$ is minimal $\|\cdot\|$-w-USCO then we are done.

To understand why proving either one of the above claims is enough, for the first claim we observe that for $T := \operatorname{conv} \mathcal{S}$, we have

$$\emptyset \neq \mathcal{S}(g) \subset T(g)$$

for each $g \in \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$. If we can prove that T is minimal $\|\cdot\|$-w-CUSCO, Proposition 5.3 would imply that T is generically single-valued and hence $\mathcal{S}$ must be as well.

The problem with this approach is that there is apparently no obvious connection between $T = \overline{\operatorname{conv}}\, \mathcal{S}$ and the functional J (or, more specifically, the integrand $F(x, u, z)$). There are many ways $\overline{\operatorname{conv}}\, \mathcal{S}$ can be computed from $\mathcal{S}$ but the semicontinuity property of $\mathcal{S}$ does not seem to translate nicely to T.

As for the second claim, it is known that if $\mathcal{S}$ is $\|\cdot\|$-w-USCO, then $T = \overline{\operatorname{conv}}\, \mathcal{S}$ is $\|\cdot\|$-w-CUSCO. This follows from the following result that can be found in a survey paper about minimal USCO and minimal CUSCO maps by Holá and Holý [51].

Proposition 5.5 (Lemma 5.1, [51]). *Let X be a topological space and let Y be a Hausdorff, locally convex topological vector space. Let $G\colon X \rightrightarrows Y$ be an USCO map and $\overline{\operatorname{conv}}\, G(x)$ be compact for every $x \in X$. Then the map F defined as $F(x) = \overline{\operatorname{conv}}\, G(x)$ for every $x \in X$ is a CUSCO map.*

A stronger result which takes into account the minimality of each map can also be found in the same paper. To set the notations, we let $MU(X, Y)$ denotes the family of all minimal USCO maps from X into Y. Likewise, $MC(X, Y)$ denotes the family of all minimal CUSCO maps from X into Y.

Proposition 5.6 (Theorem 7.7, [51]). *Let X be a topological space and let Y be a Hausdorff, locally convex topological vector space such that for any compact set $K \subset Y$, $\overline{\operatorname{conv}}\, K$ is also compact. Then the map $G \mapsto \overline{\operatorname{conv}}\, G$ is a bijection from $MU(X, Y)$ onto $MC(X, Y)$.*

According to Proposition 2.30, $Y = \mathrm{W}^{1,p}(\Omega; \mathbb{R}^N)$ equipped with the weak topology satisfies the assumption in Proposition 5.6; hence if we can show that $\mathcal{S}$ is minimal $\|\cdot\|$-w-USCO, then $T = \overline{\operatorname{conv}}\, \mathcal{S}$ is also minimal $\|\cdot\|$-w-CUSCO and we can apply Proposition 5.3 on T.

The problem with this strategy is quite different from the first strategy. Following this approach, we do not need to modify anything except proving the minimality property for $\mathcal{S}$. However, this seems to be very challenging.

Regarding the third claim, it is not hard to verify that for a function $f\colon X \to \mathbb{R}$,

$$\overline{\operatorname{conv}}\, (\operatorname*{argmin}_X f) \subset \operatorname*{argmin}_X (\overline{\operatorname{conv}}\, f)$$

hence for any $g \in \mathrm{W}^{1,p}(\Omega;\mathbb{R}^N)$ we have

$$\emptyset \neq \mathcal{S}(g) \subset \overline{\mathrm{conv}}\,(\mathcal{S}(g)) \subset \operatorname*{argmin}_{\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)} (\overline{\mathrm{conv}}\, J) = \mathcal{S}'(g).$$

We can see that $\mathcal{S}'$ is automatically convex-valued. By the same reasoning as before, if we can prove that $\mathcal{S}'$ is minimal $\|\cdot\|$-w-CUSCO then Proposition 5.3 applies to $\mathcal{S}'$m which in turn implies that $\mathcal{S}$ is generically single-valued.

This method is similar to the first approach, the only different is that we convexify the functional J before computing its minimizers, whereas in the first approach the convexification process is done at the end on the set of minimizers. In some circumstances, the first and the third methods are in fact equivalent (see [50]), hence the main difficulty of this approach is very similar to the first, i.e. there is no known way of computing $\overline{\mathrm{conv}}\, J$ from $F(x,u,z)$. In fact, it is not even known whether $\overline{\mathrm{conv}}\, J$ is an integral functional or not in general.

The fact that very little is known about the structure of the set of minimizers for the quasiconvex variational integral of the form $J(u) = \int_\Omega F(x,u,\nabla u)\,dx$ is a major problem in establishing the minimality of the solution map $\mathcal{S}$ (or its convexified relatives, as described above), so it is likely that some additional assumptions may be needed to prove the generic uniqueness result we desire. In particular, the singularly perturbed problem for the variational integral J_ε

$$\text{minimize} \int_\Omega F(\nabla u)\,dx + \varepsilon \int_\Omega |\nabla^2 u|^2\,dx \quad \text{over} \quad u \in \mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$$

that we studied in chapter 4 and its associated solution map $\widetilde{\mathcal{S}}\colon \mathrm{W}^{2,2}(\Omega;\mathbb{R}^N) \rightrightarrows \mathrm{W}^{2,2}(\Omega;\mathbb{R}^N)$ defined by

$$\widetilde{\mathcal{S}}(g) := \operatorname{argmin}\left\{ J_\varepsilon(u) : u \in \mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N) \right\}$$

may be more amendable to the treatments we discuss earlier for $\mathcal{S}$. By modifying the argument in chapter 3, we can verify that $\widetilde{\mathcal{S}}$ enjoys the same kind of semicontinuity property as $\mathcal{S}$, and even thought $\widetilde{\mathcal{S}}$ is still not convex-valued, we have seen that the associated functional J_ε is partially strictly convex (in the sense that is made precise in Proposition 4.33). The d-rectifiability of $\widetilde{\mathcal{S}}(g)$ that was shown in Theorem 4.2 (or the stronger fact that $\widetilde{\mathcal{S}}(g)$ is bi-Lipschitz equivalent to a compact subset of some Euclidean space $\mathbb{R}^d$) is another supporting fact that would likely help in proving the desired generic uniqueness result in our future research.

A Appendix

This part is the collection of results in functional analysis that is used in this thesis. Some of the propositions stated here are accompanied by a proof, but most well-known results are stated without proofs. For those without proofs, references are given.

A.1 Weak Topologies and Compactness

Let X be a locally convex vector space (or a normed space, in particular). We denote the space of all continuous linear functional on X by X^*. Here is a list of some notation.

- A sequence (x_n) in X converges strongly (in the norm topology) to x is written as $x_n \to x$. If (x_n) converges weakly then we write $x_n \overset{w}{\rightharpoonup} x$.
- For a sequence (x_n^*) in X^*, we write $x_n^* \overset{w^*}{\rightharpoonup} x^*$ if (x_n^*) converges to x^* in the weak* topology of X^*.
- The *closure* of the set $S \subset X$ is denoted by $\overline{S}$. The *weak closure* of S is denoted by $\overline{S}^w$. The *weak* closure* of a subset of X^* is written in the same manner.
- The *convex hull* of a set $S \subset X$ is denoted by $\operatorname{conv} S$. The *closed convex hull* of S is denoted by $\overline{\operatorname{conv}}\, S$ while the *weakly closed convex hull* of S is denoted by $\overline{\operatorname{conv}}^w S$, i.e.

$$\overline{\operatorname{conv}}\, S = \overline{(\operatorname{conv} S)}, \quad \overline{\operatorname{conv}}^w S = \overline{(\operatorname{conv} S)}^w.$$

- The open ball of radius r centred at x in X is denoted by $B_X(x; r)$, or just $B(x; r)$ when it is clear from the context which space the ball is in. The unit sphere of X is denoted by S_X. //

The following theorems can be found across many graduate level functional analysis textbooks (or books that are closely related). A list of such books are [108] by Yosida, [37] by Edwards, [6] by Aliprantis and Burkinshaw, or the encyclopedic book [5] by Aliprantis and Border.

Theorem A.1 (Banach–Alaoglu)**.** *Let X be a normed space, then the closed unit ball of X^* is weak* compact. Consequently, a subset $S \subset X^*$ is weak* compact if and only if it is weak* closed and norm bounded.*

Proof. See [5, Theorem 6.21]. □

Theorem A.2 (Eberlein–Šmulian)**.** *A subset S of a normed space X is weakly relatively compact (resp. weakly compact) if and only if every sequence (x_n) in S has a subsequence that converges to some $x \in X$ (resp. to some $x \in S$).*

Proof. See [6, Theorem 3.40]. □

It is well-known that bounded subsets of $\mathrm{W}^{1,p}(\Omega)$ are weakly precompact if $p > 1$ but not for $p = 1$, together with another closely related fact that the weak topology on a bounded subset of $\mathrm{W}^{1,p}(\Omega)$ is metrizable if $p > 1$ but not for $p = 1$ (because $\mathrm{L}^\infty(\Omega)$ is not separable). However, it is perhaps a little less well-known that the weak topology on a weakly precompact subset of $\mathrm{W}^{1,1}(\Omega)$ is metrizable. Here we give a proof of this fact.

Lemma A.3. *Let X be a separable normed space. Then there exists a countable subset $\mathcal{F} \subset X^*$ such that $\mathcal{F}$ separates points of X.*

Proof. Let D be a countable, dense subset of S_X, the unit sphere of X. Let $\mathcal{F}$ be the family of all norming functionals of the elements in D.

For any distinct $x, y \in X$, there exists a $\lambda > 0$ such that $\|z\| = 1$ for $z = \lambda(x-y)$. Choose $z' \in D$ such that $\|z - z'\| < \frac{1}{2}$. Let $f \in \mathcal{F}$ be the norming functional of z', then

$$f(z) = f(z') + f(z - z') \geq 1 - 1/2 = 1/2.$$

This shows that $\mathcal{F}$ separates points of X. □

Lemma A.4. *Let X be a normed space and $\{f_n\}_{n=1}^\infty \subset X^*$ be a family of functionals that separates points of X. Then the map $d\colon X \times X \to [0, 1)$ given by*

$$d(x, y) := \sum_{n=1}^{\infty} \frac{2^{-n}|f_n(x - y)|}{1 + |f_n(x - y)|}$$

defines a metric on X. Moreover, the metric topology (X, d) is coarser than the weak topology (X, w).

Proof. It is not hard to verify the triangle inequality for $d(\cdot, \cdot)$ using the fact that $\rho(x) := \frac{x}{1+x}$ is concave on $[0, \infty)$ with $\rho(0) = 0$, so it is subadditive. $d(x, y) = 0$ if and only if $x = y$ follows from the fact that $\{f_n\}_{n=1}^\infty$ separates points of X.

Next, we want to show that $(X, d) \leq (X, w)$, i.e. any open subset in τ_d is open in the weak topology.

Let $B(x_0; \varepsilon)$ be a ball in (X, d), we wish to find a weak neighbourhood U of x_0 such that $U \subset B(x_0; \varepsilon)$. First, let N be a number such that $\sum_{n=N+1}^\infty 2^{-n} < \varepsilon/2$. We choose $\delta > 0$ small enough so that if $x \in X$ satisfies $|f_n(x_0 - x)| < \delta$ for $n = 1, \ldots, N$ then

$$\sum_{n=1}^{N} \frac{2^{-n}|f_n(x_0 - x)|}{1 + |f_n(x_0 - x)|} < \varepsilon/2.$$

Thus for $U := \cap_{n=1}^N \{|f_n(\cdot - x_0)| < \delta\}$, we have $U \subset B(x_0; \varepsilon)$. □

Proposition A.5. *Let X be a separable normed space and $S \subset X$ be a weakly compact subset. Then S is metrizable.*

Proof. By Lemma A.3, there is a countable subset $\mathcal{F} \subset X^*$ that separates points of X, hence Lemma A.4 implies that X has a metric topology, given by a metric d, that is coarser than (X, w).

Suppose that S is a weakly compact subset of X. Clearly (S, d) is Hausdorff and since (S, w) is compact, the inclusion $(S, w) \hookrightarrow (S, d)$ is a continuous bijection, hence a homeomorphism. This means that $(S, d) = (S, w)$, so the weak topology on S is metrizable. □

Corollary A.6. *The weak topology on a weakly compact subset of $\mathrm{W}^{1,p}(\Omega)$ is metrizable for all $p \geq 1$.*

A.2 Convergence and Compactness in $\mathbf{L}^p(\Omega)$

The following are some useful theorems regarding the compactness of the Lebesgue spaces $\mathrm{L}^p(\Omega)$ and the Sobolev spaces $\mathrm{W}^{1,p}(\Omega)$. Some good resources on this topic are [62] by Kufner, John and Fucik, [7] by Ambrosio, Fusco and Pallara, [85] by Rao, and especially [45] by Fonseca and Leoni. For more on measure theory, see [11] by Bogachev or [43] by Evans and Gariepy.

Definition A.7 (Equi-Integrability)**.** *Let $\mathcal{F} \subset \mathrm{L}^1(\Omega, \mu)$. $\mathcal{F}$ is called **equi-integrable** if for any $\varepsilon > 0$, there exists a corresponding $\delta = \delta(\varepsilon)$ such that*

$$\int_E |f| \, d\mu < \varepsilon \quad \text{for all} \quad f \in \mathcal{F}$$

whenever $\mu(E) < \delta$. //

There is a close relationship between equi-integrablility and uniform integrability. These two notions are equivalent if we add an additional hypothesis to an equi-integrable family that $\sup_{f \in \mathcal{F}} \|f\|_1 < \infty$ (see [45, Theorem 2.29]).

Definition A.8 (Uniform Integrability)**.** *Let $\mathcal{F} \subset \mathrm{L}^1(\Omega, \mu)$. $\mathcal{F}$ is called uniformly integrable if for any $\varepsilon > 0$, there exists a corresponding M such that*

$$\int_{\{|f| > M\}} |f| \, d\mu < \varepsilon \quad \text{for all} \quad f \in \mathcal{F}$$

whenever E is a measurable set satisfying $\mu(E) < \delta$. //

Definition A.9 (Convergence in measure)**.** *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Let $f, f_n \colon \Omega \to \mathbb{R}$ be measurable. the sequence $(f_n)_n$ is said to **converges in measure** to f if for any $\varepsilon > 0$*

$$\lim_{n \to \infty} \mu(\{|f - f_n| > \varepsilon\}) = 0.$$

//

Proposition A.10. *Let (Ω, μ) be such that $\mu(\Omega) < \infty$, then*

1. *If $f_n \to f$ a.e., then $f_n \to f$ in measure.*
2. *$f_n \to f$ in measure $\iff$ every subsequence of $(f_n)_n$ has a subsequence converging to f a.e.*

Proof. See [45, Theorem 2.20 and Theorem 2.22]. □

Proposition A.11 (Vitali convergence theorem)**.** *Suppose that $\mu(\Omega) < \infty$. Let $(f_n)_n \subset \mathrm{L}^p(\Omega, \mu)$ and $f \in \mathrm{L}^p(\Omega, \mu)$, where $p \in [1, \infty)$.*

Then $f_n \to f$ in L^p if and only if

1. *$f_n \to f$ in measure.*
2. *$(|f_n|^p)_n$ is equi-integrable.*

Proof. See [45, Theorem 2.24]. □

Theorem A.12 (Dunford–Pettis theorem)**.** *Let $\mathcal{F} \subset \mathrm{L}^1(\Omega, \mu)$. Then*

$\mathcal{F}$ is relatively weakly compact in L^1 $\iff$ $\mathcal{F}$ is equi-integrable and $\sup_{f \in \mathcal{F}} \|f\|_1 < \infty$.

Proof. See [45, Theorem 2.54]. □

Note that in many places here where $\mu(\Omega) < \infty$, we can replace the requirement that $\sup_{f \in \mathcal{F}} \|f\|_1 < \infty$ with the assumption that the measure μ be *nonatomic*. Note that the Lebesgue measure on $\mathbb{R}^n$ is nonatomic (see [45, Remark 1.161]).

Definition A.13 (Nonatomic Measure)**.** *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. A set $E \in \mathcal{A}$ of positive measure is said to be an* atom *if for every $F \in \mathcal{A}$ with $F \subset E$ either $\mu(F) = 0$ or $\mu(F) = \mu(E)$. The measure μ is said to be* nonatomic *if there are no atoms, i.e. if for every set $E \in \mathcal{A}$ of positive measure, there exists $F \in \mathcal{A}$ with $F \subset E$ such that $0 < \mu(F) < \mu(E)$.* //

Proposition A.14. *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space such that μ is nonatomic. Let r_n be a sequence of positive numbers such that*

$$\sum_{n=1}^{\infty} r_n \leq \mu(\Omega),$$

then there exists a sequence of mutually disjoint measurable sets A_n such that $\mu(A_n) = r_n$.

Proof. See [45, Corollary 1.21]. □

Proposition A.15. *If $\mu(\Omega) < \infty$ and μ is nonatomic, then equi-integrability of $\mathcal{F}$ implies $\sup_{f\in\mathcal{F}} \|f\|_1 < \infty$.*

Proof. Take $\varepsilon = 1$, then by equi-integrable we can find δ such that $\mu(A) < \delta$ implies that $\int_A |f|\,d\mu < 1$ uniformly in $f \in \mathcal{F}$.

Since μ is nonatomic and $\mu(\Omega) < \infty$, we can find measurable sets $A_1, \dots, A_N$, each satisfying $\mu(A_j) = \delta/2$ for all $j < N$ and $\mu(A_N) < \delta$, such that $\Omega = \bigcup_{j=1}^N A_j$. Hence

$$\int_\Omega |f|\,d\mu = \sum_{j=1}^N \int_{A_j} |f|\,d\mu \le N. \qquad \square$$

Theorem A.16 (de la Vallée–Poussin criterion, 2.31)**.** *Let μ be a finite measure on Ω, $\mathcal{F} \subset \mathrm{L}^1(\Omega,\mu)$. Then $\mathcal{F}$ is equi-integrable and $\sup_{f\in\mathcal{F}} \|f\|_1 < \infty$ if and only if there exists an increasing function $\theta\colon [0,\infty) \to [0,\infty)$ such that*

1. *$\frac{\theta(t)}{t} \to \infty$ as $t \to \infty$.*
2. *$\sup_{\mathcal{F}} \int_\Omega \theta(|f(x)|)\,d\mu < \infty$.*

Additionally, if such a function exists we may take it to be convex.

Proof. See [45, Theorem 2.29 and Remark 4.99]. $\square$

Proposition A.17 (Proposition 1.21)**.** *For a functional $J\colon \mathrm{W}^{1,p}(\Omega;\mathbb{R}^d) \to \mathbb{R}^+$ of the form 1.1, where $p \ge 1$, J is strongly lsc in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^d)$ provided its integrand L is normal and satisfies the lower bound*

$$L(x,v,A) \ge -g(x) - \alpha|v|^q - \beta|A|^p, \tag{A.1}$$

where $\alpha,\beta \ge 0$, $1 \le q \le \frac{np}{n-p}$ and $g \in \mathrm{L}^1(\Omega)$.

Proof. Suppose $(u_n)_n$ is a sequence in $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^d)$ converging to u. Without loss of generality, we may assume that $J(u_n) < \infty$ for all n and that (after passing to a subsequence) $u_n(x) \to u(x)$ and $\nabla u_n(x) \to \nabla u(x)$ pointwise a.e.-Ω.

By the Rellich-Kondrachov theorem, $u_n \to u$ strongly in $\mathrm{L}^q(\Omega)$ for $1 \le q \le \frac{np}{n-p}$ and we already knew that $\nabla u_n \to \nabla u$ strongly in $\mathrm{L}^p(\Omega)$, thus

$$\begin{aligned}
\liminf_{n\to\infty} \int_\Omega L(x,u_n,\nabla u_n) + g(x) + \alpha|u_n|^q + \beta|\nabla u_n|^p\,dx & \\
&\ge \int_\Omega \liminf_{n\to\infty} \left(L(x,u_n,\nabla u_n) + g(x) + \alpha|u_n|^q + \beta|\nabla u_n|^p\right)\,dx \\
&\ge \int_\Omega L(x,u,\nabla u) + g(x) + \alpha|u|^q + \beta|\nabla u|^p\,dx
\end{aligned}$$

by Fatou's lemma and the fact that L is normal. We may then subtract $\int_\Omega g(x) + \alpha|u|^q + \beta|\nabla u|^p\,dx$ from both sides and deduce that

$$\int_\Omega L(x,u,\nabla u)\,dx \le \liminf_{n\to\infty} \int_\Omega L(x,u_n,\nabla u_n)\,dx. \qquad \square$$

A.3 Bochner Integrals, Vector Measures and the Radon–Nikodym Property (RNP)

The integration theory of vector-valued functions has many far-reaching applications in many parts of mathematical analysis. In this section we give a brief introduction to the *Bochner integral*, which is essentially a direct generalization of Lebesgue's integration theory, and hence many theorems for Lebesgue integral continue to hold for Bochner integrals (when suitably modified). For this reason, Bochner integration is used to defined the X-valued Lebesgue space $\mathrm{L}^1(\Omega; X)$ which, together with the notion of *vector measure*, is used to define the *Radon–Nikodym Property*.

The general theory of vector-valued integration can be found in the books [37] by Edwards or [108] by Yosida. For books that are more geared toward the Radon–Nikodym Property, one should consult [14] by Bourgin and [32] by Diestel and Uhl. For a general overview of vector integration in modern calculus of variations, [45] by Fonseca is a great resource. Much of the material in this section is taken from the last three books.

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, we denote its *completion* by $(\Omega, \mathcal{F}', \mu')$. Let X be a Banach space. A function $s\colon \Omega \to X$ is called a **simple function** if there exists a finite measurable partition $\{A_i\}_{i=1}^n$ of Ω such that

$$s = \sum_{i=1}^{n} x_i \chi_{A_i}$$

for some distinct $x_i \in X$. Here χ_{A_i} is the characteristic function of the set A_i. The integral of a simple function s is defined to be

$$\int_A s \, d\mu := \sum_{i=1}^{n} x_i \mu(A \cup A_i).$$

Definition A.18 (Bochner Integrability)**.** *A function $f\colon \Omega \to X$ is said to be* Bochner Integrable *if there exists a sequence $(s_n)_n$ of simple functions such that*

1. $\lim_{n\to\infty} s_n(\omega) = f(\omega)$ *a.e.-μ.*
2. $\lim_{n\to\infty} \int_\Omega \|f(\omega) - s_n(\omega)\|_X \, d\mu = 0$.

Symbolically, we write $f \in \mathrm{L}^1(\Omega, \mu; X)$ to mean that f is Bochner integrable. //

When f is Bochner integrable and $A \in \mathcal{F}$, then $\lim_{n\to\infty} \int_A s_n \, d\mu$ exists and the value is independent of the choice of the sequence $(s_n)_n$ we choose to approximate f. In this case, we define the integral of f to be

$$\int_A f \, d\mu := \lim_{n\to\infty} \int_A s_n \, d\mu.$$

Definition A.19. *A function $f\colon \Omega \to X$ is said to be*

1. ***strongly $\mathcal{F}$-measurable*** *if there exists a sequence $(s_n)_n$ of simple functions such that*
$$f(\omega) = \lim_{n\to\infty} s_n(\omega) \quad \text{a.e.-}\mu.$$
2. ***weakly $\mathcal{F}$-measurable*** *if for any $x^* \in X^*$ the function $\langle x^*, f(\cdot)\rangle : \Omega \to \mathbb{R}$ is $\mathcal{F}$-measurable.*
3. ***weakly* $\mathcal{F}$-measurable*** *if $X = Y^*$ for some Banach space Y and for any $y \in Y$ the function $\langle f(\cdot), y\rangle : \Omega \to \mathbb{R}$ is $\mathcal{F}$-measurable.*
4. ***almost separably valued*** *if there exists $E \in \mathcal{F}$ such that $\mu(\Omega\backslash E) = 0$ and $f(E)$ is $\|\cdot\|$-separable.*
5. ***$\mathcal{F}$-Borel measurable*** *if $f^{-1}(U) \in \mathcal{F}$ for any open set U in the norm topology of X.* //

These properties are related to each other (assuming μ is finite or σ-finite). If $f\colon \Omega \to X$ is almost separably valued and $\mathcal{F}$-Borel measurable, then f is strongly $\mathcal{F}$-measurable and we can choose an approximating sequence such that

$$\|s_n(\omega)\|_X \leq \|f(\omega)\|_X \quad \text{a.e.-}\mu$$

for all $n \in \mathbb{N}$. This shows that there exists a version of a strongly $\mathcal{F}$-measurable f that is separably valued and $\mathcal{F}$-Borel measurable (see [45, Theorem 2.104] or the discussion after [14, Definition 1.1.6]).

Theorem A.20. *Let $(\Omega, \mathcal{F}, \mu)$ be a finite measure space, then the following are equivalent:*

1. *$f \in \mathrm{L}^1(\Omega, \mu; X)$.*
2. *f is weakly $\mathcal{F}'$-measurable, almost separably valued and $\int_\Omega \|f(\omega)\|_X \, d\mu'(\omega) < \infty$.*
3. *f is strongly $\mathcal{F}$-measurable and $\int_\Omega \|f(\omega)\|_X \, d\mu(\omega) < \infty$.*
4. *f is is almost separably valued, $\mathcal{F}$-Borel measurable and $\int_\Omega \|f(\omega)\|_X \, d\mu'(\omega) < \infty$.*

Proof. See [14, Theorem 1.1.7]. □

A Bochner integrable function interacts well with a bounded linear operator.

Proposition A.21. *Suppose that X and Y are Banach spaces and $T\colon X \to Y$ is a bounded linear operator, then for $f \in \mathrm{L}^1(\Omega, \mu; X)$ we have $T \circ f \in \mathrm{L}^1(\Omega, \mu; Y)$ and*

$$T\left(\int_A f \, d\mu\right) = \int_A (T \circ f) \, d\mu.$$

In particular, for a linear functional $x^* \in X^*$, we have

$$\left\langle x^*, \int_A f \, d\mu \right\rangle = \int_A \langle x^*, f(\omega) \rangle \, d\mu(\omega).$$

Proof. See [14, Theorem 1.1.8 (e)] (and perhaps [32, Theorem II.2.6] for a more general statement where T is a closed operator). □

We can define the vector-valued L^p spaces as follows:

Definition A.22. *Let $(\Omega, \mathcal{F}, X)$ be a measure space, X be a Banach space and $p \in [1, \infty]$. Then*

$$\mathrm{L}^p(\Omega, \mu; X) := \{u\colon \Omega \to X : u \text{ is strongly measurable and } \|u\|_{\mathrm{L}^p(\Omega,\mu;X)} < \infty\},$$

where $\|\cdot\|_{\mathrm{L}^p(\Omega,\mu;X)}$ is defined by

$$\|u\|_{\mathrm{L}^p(\Omega,\mu;X)} := \left(\int_\Omega \|u(\omega)\|_X^p \, d\mu(\omega)\right)^{\frac{1}{p}}$$

for $p < \infty$. For $p = \infty$, we define

$$\begin{aligned}\|u\|_{\mathrm{L}^\infty(\Omega,\mu;X)} &:= \operatorname{esssup}_{\omega\in\Omega} \|u(\omega)\|_X \\ &= \inf\{\alpha \in \mathbb{R} : \|u(\omega)\|_X < \alpha \quad \text{for almost every} \quad \omega \in \Omega\}.\end{aligned}$$

Suppose that $X = Y^*$ for some Banach space Y, we may define

$$\mathrm{L}^p_w(\Omega, \mu; X) := \{u\colon \Omega \to X : u \text{ is weakly}^* \text{ measurable and } \|u\|_{\mathrm{L}^p_w(\Omega,\mu;Y^*)} < \infty\},$$

where $\|u\|_{\mathrm{L}^p_w(\Omega,\mu;Y^*)} := \|u\|_{\mathrm{L}^p(\Omega,\mu;X)}$. //

Of course, the 'functions' in the above definition are identified if they are equal almost everywhere. Here are some important facts regarding vector-valued L^p spaces.

Proposition A.23. *Let $(\Omega, \mathcal{F}, X)$ be a measure space, X be a Banach space. Then*

1. *$\mathrm{L}^p(\Omega, \mu; X)$ is a Banach space for all $p \in [1, \infty]$.*
2. *If Ω is a separable metric space, μ is a σ-finite Radon measure and X is separable, then $\mathrm{L}^p(\Omega, \mu; X)$ is separable for all $p \in [1, \infty)$.*

See [45, Theorem 2.110] for more details.

Proposition A.24 (Riesz representation theorem for L^p)**.** *Let $(\Omega, \mathcal{F}, X)$ be a σ-finite measure space, X be a Banach space and $p \in [1, \infty)$. Then*

1. *Assume that X is separable, then $\mathrm{L}^p(\Omega, \mu; X)^* \simeq \mathrm{L}^{p'}_w(\Omega, \mu; X^*)$.*
2. *Assume that X is reflexive, then $\mathrm{L}^p(\Omega, \mu; X)^* \simeq \mathrm{L}^{p'}(\Omega, \mu; X^*)$.*

Here p' is the Hölder conjugate of p. The pairing between a functions in $f \in \mathrm{L}^p(\Omega, \mu; X)$ and its dual $g \in \mathrm{L}^p_w(\Omega, \mu; X)$ (or $\mathrm{L}^p(\Omega, \mu; X)$, respectively) is given by

$$\langle g, f \rangle := \int_\Omega \langle g(\omega), f(\omega) \rangle_{X^*, X} \, d\mu(\omega).$$

Proof. See [45, Theorem 2.112]. □

Similar to how vector-valued integration is a generalization of the real-valued Lebesgue integration theory, vector measure is defined in almost the same way as the usual signed measure0

Definition A.25 (Vector measure)**.** *Let $(\Omega, \mathcal{F})$ be a measurable space and X be a Banach space. A map $\mu \colon \mathcal{F} \to X$ is called a **vector measure** (more specifically, an X-value measure) if it is countably additive, i.e.*

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k)$$

for any sequence of pairwise disjoint $A_k \in \mathcal{F}$. The summation on the right means that the partial sums $\sum_{k=1}^{N} \mu(A_k)$ are required to form a convergence sequence in the norm topology of X. //

Just like the case for complex-valued measure, an X-valued measure is associated with a positive measure $|\mu|$, called the **total variation measure**, defined by

$$|\mu|(A) := \sup \left\{ \sum_{k=1}^{n} \mu(A \cap E_k) : (E_k)_{k=1}^n \text{ is a measurable partiation of } \Omega \right\}$$

for each $A \in \mathcal{F}$. μ is said to have *finite total variation* if $|\mu|(\Omega) < \infty$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let us consider a (rather restrictive) version of the Radon–Nikodym theorem

Theorem A.26 (Radon–Nikodym theorem)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mu \colon \mathcal{F} \to \mathbb{R}$ be a measure satisfying $\mu \ll \mathbb{P}$ (absolutely continuous with respect to $\mathbb{P}$) and $|\mu|(\Omega) < \infty$ (finite total variation). Then there exists an $f \in \mathrm{L}^1(\Omega, \mathbb{P}; \mathbb{R})$ such that*

$$\mu(A) = \int_A f \, d\mathbb{P} \quad \text{for all} \quad A \in \mathcal{F}.$$

We also write $f = \frac{d\mu}{d\mathbb{P}}$.

We may replace $\mathbb{R}$ by a general Banach space X and ask if the result still holds, i.e. if an X-valued measure $\mu\colon \mathcal{F} \to X$ can be represented by a function $f \in \mathrm{L}^1(\Omega, \mathbb{P}; X)$. In general, the answer is *no*. It turns out that we can characterize the Banach spaces in which such representation is possible, using the property called the Radon–Nikodym Property.

Definition A.27 (Average range)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let μ be a X-valued measure on $(\Omega, \mathcal{F})$ which is absolutely continuous with respect to $\mathbb{P}$. Its* ***average range*** *is defined to be*

$$AR(\mu) := \left\{ \frac{\mu(A)}{\mathbb{P}(A)} : A \in \mathcal{F}, \mathbb{P}(A) > 0 \right\}.$$

//

Definition A.28 (Radon–Nikodym Property)**.** *Let K be a closed bounded convex set in X, a Banach space. Then we say that K has the* ***Radon–Nikodym Property for*** *$(\Omega, \mathcal{F}, \mathbb{P})$ if each X-valued measure μ on $\mathcal{F}$ such that $\mu \ll \mathbb{P}$ and $AR(\mu) \subset K$, there exists $f \in \mathrm{L}^1(\Omega, \mathbb{P}; X)$ satisfying*

$$\mu(A) = \int_A f \, d\mathbb{P}$$

*for all $A \in \mathcal{F}$. K is said to have the Radon–Nikodym Property (**RNP**) if it has the Radon–Nikodym Property for all probability spaces.* //

Definition A.29. *Let $C \subset X$ be a closed convex set (possibly unbounded). We say that C has the RNP if each closed, bounded, convex subset $K \subset C$ has the RNP.* //

There are another two equivalent formulations of the RNP, using the fact that

$$\frac{d\mu}{d\mathbb{P}} = \frac{d\mu}{d|\mu|} \frac{d|\mu|}{d\mathbb{P}} \quad \text{a.s.-}\mathbb{P}$$

for a vector measure $\mu\colon \mathcal{F} \to X$ of finite total variation whenever $\frac{d\mu}{d|\mu|}$ exists.

1. X has the RNP if for any X-valued measure μ on $(\Omega, \mathcal{F})$ with finite total variation, there exists $f \in \mathrm{L}^1(\Omega, |\mu|; X)$ such that

   $$|\mu|(A) = \int_A f \, d|\mu|$$

   for all $A \in \mathcal{F}$.

2. X has the RNP if for any X-valued measure μ on $(\Omega, \mathcal{F})$ with finite total variation and $\mu \ll \mathbb{P}$, there exists $f \in \mathrm{L}^1(\Omega, \mathbb{P}; X)$ such that

   $$\mu(A) = \int_A f \, d\mathbb{P}$$

   for all $A \in \mathcal{F}$.

For more information on these two equivalent formulations of the RNP, see the discussion after [14, Example 2.1.2].

There are several other characterizations of the RNP. The following is the geometric version of the RNP.

Definition A.30.
- *For a nonempty subset A of X, $\alpha > 0$ and $x^* \in X^*$, we defined a* slice of A *to be*

$$S(x^*; A, \alpha) := \{x \in A \mid \langle x^*, x\rangle > \sigma_A(x^*) - \alpha\},$$

where σ_A is the support function *defined by $\sigma_A(x^*) := \sup\{\langle x^*, x\rangle \mid x \in A\}$.*

- *For $A \subset X^*$, its w^*-slices is defined similarly, i.e.*

$$S(x; A, \alpha) := \{x^* \in A \mid \langle x^*, x\rangle > \sigma_A(x) - \alpha\},$$

where $\sigma_A(x) := \sup\{\langle x^*, x\rangle \mid x^* \in A\}$. //

Definition A.31 (RNP, geometric form)**.** *Let X be a Banach space. A nonempty subset S of X is said to have the* Radon-Nikodym Property *(RNP) if every bounded, nonempty subset $A \subset S$ has one of the following (equivalent) property.*

1. *A is* dentable*: for each $\varepsilon > 0$, there exists $x \in A$ such that $x \notin \overline{\text{conv}}\,(A \backslash B(x, \varepsilon))$.*
2. *A contains slices of arbitrarily small diameters.*

In a similar manner, a nonempty subset S of X^* is said to have the w^*-Radon-Nikodym Property *(w^*-RNP) if every bounded, nonempty subset $A \subset S$ contains w^*-slices of arbitrarily small diameters.* //

B References

[1] E. Acerbi and N. Fusco. Semicontinuity problems in the calculus of variations. *Arch. Rational Mech. Anal.*, 86(2):125–145, 1984.

[2] E. Acerbi and N. Fusco. Regularity for minimizers of nonquadratic functionals: the case $1 < p < 2$. *J. Math. Anal. Appl.*, 140(1):115–135, 1989.

[3] S. Agmon. The L_p approach to the Dirichlet problem. I. Regularity theorems. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)*, 13:405–448, 1959.

[4] J. Alibert and B. Dacorogna. An example of a quasiconvex function that is not polyconvex in two dimensions. *Arch. Rational Mech. Anal.*, 117(2):155–166, 1992.

[5] C. D. Aliprantis and K. C. Border. *Infinite dimensional analysis*. Springer, Berlin, third edition, 2006. A hitchhiker's guide.

[6] C. D. Aliprantis and O. Burkinshaw. *Positive operators*. Springer, Dordrecht, 2006. Reprint of the 1985 original.

[7] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[8] E. Asplund. Fréchet differentiability of convex functions. *Acta Math.*, 121:31–47, 1968.

[9] M. Baía, A. C. Barroso, M. Chermisi, and J. Matias. Coupled second order singular perturbations for phase transitions. *Nonlinearity*, 26(5):1271–1311, 2013.

[10] J. M. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Rational Mech. Anal.*, 63(4):337–403, 1976/77.

[11] V. I. Bogachev. *Measure theory. Vol. I, II*. Springer-Verlag, Berlin, 2007.

[12] J. M. Borwein and J. D. Vanderwerff. *Convex functions: constructions, characterizations and counterexamples*, volume 109 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2010.

[13] J. M. Borwein and Q. J. Zhu. *Techniques of variational analysis*, volume 20 of *CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC*. Springer-Verlag, New York, 2005.

[14] R. D. Bourgin. *Geometric aspects of convex sets with the Radon-Nikodým property*, volume 993 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1983.

[15] J. Campos Cordero, B. Kirchheim, J. Kolář, and J. Kristensen. *Private Communication.*

[16] J. Campos Cordero and J. Kristensen. *Private Communication.*

[17] M. Carozza, N. Fusco, and G. Mingione. Partial regularity of minimizers of quasiconvex integrals with subquadratic growth. *Ann. Mat. Pura Appl. (4)*, 175:141–164, 1998.

[18] M. Carozza, J. Kristensen, and A. Passarelli di Napoli. Higher differentiability of minimizers of convex variational integrals. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 28(3):395–411, 2011.

[19] M. Carozza, J. Kristensen, and A. Passarelli di Napoli. A trace preserving operator and applications. *J. Math. Anal. Appl.*, 501(1):124170, 24, 2021.

[20] M. Carozza and A. Passarelli di Napoli. A regularity theorem for minimisers of quasiconvex integrals: the case $1 < p < 2$. *Proc. Roy. Soc. Edinburgh Sect. A*, 126(6):1181–1199, 1996.

[21] C. Y. Chen and J. Kristensen. On coercive variational integrals. *Nonlinear Anal.*, 153:213–229, 2017.

[22] M. Chermisi, G. Dal Maso, I. Fonseca, and G. Leoni. Singular perturbation models in phase transitions for second-order materials. *Indiana Univ. Math. J.*, 60(2):367–409, 2011.

[23] M. M. Choban, P. S. Kenderov, and J. P. Revalski. Variational principles and topological games. *Topology Appl.*, 159(17):3550–3562, 2012.

[24] G. Choquet. *Lectures on analysis. Vol. I: Integration and topological vector spaces*. Edited by J. Marsden, T. Lance and S. Gelbart. W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[25] M. Cicalese, E. N. Spadaro, and C. I. Zeppieri. Asymptotic analysis of a second-order singular perturbation model for phase transitions. *Calc. Var. Partial Differential Equations*, 41(1-2):127–150, 2011.

[26] J. B. Collier. The dual of a space with the Radon-Nikodým property. *Pacific J. Math.*, 64(1):103–106, 1976.

[27] S. Conti, I. Fonseca, and G. Leoni. A Γ-convergence result for the two-gradient theory of phase transitions. *Comm. Pure Appl. Math.*, 55(7):857–936, 2002.

[28] B. Dacorogna. *Direct methods in the calculus of variations*, volume 78 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2008.

[29] B. Dacorogna and J. Haeberly. Remarks on a numerical study of convexity, quasiconvexity, and rank one convexity. In *Variational methods for discontinuous structures (Como, 1994)*, volume 25 of *Progr. Nonlinear Differential Equations Appl.*, pages 143–154. Birkhäuser, Basel, 1996.

[30] B. Dacorogna and J. Haeberly. Some numerical methods for the study of the convexity notions arising in the calculus of variations. *RAIRO Modél. Math. Anal. Numér.*, 32(2):153–175, 1998.

[31] B. Dacorogna and P. Marcellini. A counterexample in the vectorial calculus of variations. In *Material instabilities in continuum mechanics (Edinburgh, 1985–1986)*, Oxford Sci. Publ., pages 77–83. Oxford Univ. Press, New York, 1988.

[32] J. Diestel and J. J. Uhl, Jr. *Vector measures*. American Mathematical Society, Providence, R.I., 1977. With a foreword by B. J. Pettis, Mathematical Surveys, No. 15.

[33] T. K. Donaldson and N. S. Trudinger. Orlicz-Sobolev spaces and imbedding theorems. *J. Functional Analysis*, 8:52–75, 1971.

[34] L. D'Onofrio and T. Iwaniec. The p-harmonic transform beyond its natural domain of definition. *Indiana Univ. Math. J.*, 53(3):683–718, 2004.

[35] A. L. Dontchev and T. Zolezzi. *Well-posed optimization problems*, volume 1543 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1993.

[36] F. Duzaar, J. F. Grotowski, and M. Kronz. Regularity of almost minimizers of quasi-convex variational integrals with subquadratic growth. *Ann. Mat. Pura Appl. (4)*, 184(4):421–448, 2005.

[37] R. E. Edwards. *Functional analysis. Theory and applications*. Holt, Rinehart and Winston, New York-Toronto-London, 1965.

[38] I. Ekeland. On the variational principle. *J. Math. Anal. Appl.*, 47:324–353, 1974.

[39] I. Ekeland and R. Témam. *Convex analysis and variational problems*, volume 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied

Mathematics (SIAM), Philadelphia, PA, english edition, 1999. Translated from the French.

[40] L. C. Evans. Quasiconvexity and partial regularity in the calculus of variations. *Arch. Rational Mech. Anal.*, 95(3):227–252, 1986.

[41] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[42] L. C. Evans and R. F. Gariepy. Some remarks concerning quasiconvexity and strong convergence. *Proc. Roy. Soc. Edinburgh Sect. A*, 106(1-2):53–61, 1987.

[43] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*. Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition, 2015.

[44] H. Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York, Inc., New York, 1969.

[45] I. Fonseca and G. Leoni. *Modern methods in the calculus of variations: L^p spaces*. Springer Monographs in Mathematics. Springer, New York, 2007.

[46] S. S. Ghoshal, A. Jana, and K. Koumatos. On the uniqueness of solutions to hyperbolic systems of conservation laws. *J. Differential Equations*, 291:110–153, 2021.

[47] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[48] E. Giusti. *Direct methods in the calculus of variations*. World Scientific Publishing Co., Inc., River Edge, NJ, 2003.

[49] F. Gmeineder and J. Kristensen. Partial regularity for BV minimizers. *Arch. Ration. Mech. Anal.*, 232(3):1429–1473, 2019.

[50] J.-B. Hiriart-Urruty, M. A. López, and M. Volle. The ϵ-strategy in variational analysis: illustration with the closed convexification of a function. *Rev. Mat. Iberoam.*, 27(2):449–474, 2011.

[51] Ľ. Holá and D. Holý. Minimal usco and minimal cusco maps. *Khayyam J. Math.*, 1(2):125–150, 2015.

[52] T. Iwaniec and J. Kristensen. A construction of quasiconvex functions. *Riv. Mat. Univ. Parma (7)*, 4*:75–89, 2005.

[53] T. Iwaniec and G. Martin. *Geometric function theory and non-linear analysis*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2001.

[54] R. C. James. Bases and reflexivity of Banach spaces. *Ann. of Math. (2)*, 52:518–527, 1950.

[55] P. S. Kenderov and W. B. Moors. Fragmentability and sigma-fragmentability of Banach spaces. *J. London Math. Soc. (2)*, 60(1):203–223, 1999.

[56] K. Koumatos and A. P. Vikelis. $\mathcal{A}$-quasiconvexity, Gårding inequalities, and applications in PDE constrained problems in dynamics and statics. *SIAM J. Math. Anal.*, 53(4):4178–4211, 2021.

[57] J. Kristensen. Lower semicontinuity of quasi-convex integrals in BV. *Calc. Var. Partial Differential Equations*, 7(3):249–261, 1998.

[58] J. Kristensen. Lower semicontinuity in spaces of weakly differentiable functions. *Math. Ann.*, 313(4):653–710, 1999.

[59] J. Kristensen. On the non-locality of quasiconvexity. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 16(1):1–13, 1999.

[60] J. Kristensen and A. Taheri. Partial regularity of strong local minimizers in the multi-dimensional calculus of variations. *Arch. Ration. Mech. Anal.*, 170(1):63–89, 2003.

[61] A. Y. Kruger, S. Plubtieng, and T. Seangwattana. Borwein-Preiss vector variational principle. *Positivity*, 21(4):1273–1292, 2017.

[62] A. Kufner, O. John, and S. Fučík. *Function spaces*. Noordhoff International Publishing, Leyden; Academia, Prague, 1977. Monographs and Textbooks on Mechanics of Solids and Fluids; Mechanics: Analysis.

[63] M. Lassonde. Asplund spaces, Stegall variational principle and the RNP. *Set-Valued Var. Anal.*, 17(2):183–193, 2009.

[64] M. Lassonde and J. P. Revalski. Fragmentability of sequences of set-valued mappings with applications to variational principles. *Proc. Amer. Math. Soc.*, 133(9):2637–2646, 2005.

[65] H. B. Lawson, Jr. and R. Osserman. Non-existence, non-uniqueness and irregularity of solutions to the minimal surface system. *Acta Math.*, 139(1-2):1–17, 1977.

[66] J. Lindenstrauss. On operators which attain their norm. *Israel J. Math.*, 1:139–148, 1963.

[67] J. Lindenstrauss, D. Preiss, and J. Tišer. *Fréchet differentiability of Lipschitz functions and porous sets in Banach spaces*, volume 179 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2012.

[68] P.-L. Lions. The concentration-compactness principle in the calculus of variations. The locally compact case. I. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 1(2):109–145, 1984.

[69] P.-L. Lions. The concentration-compactness principle in the calculus of variations. The locally compact case. II. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 1(4):223–283, 1984.

[70] P.-L. Lions. The concentration-compactness principle in the calculus of variations. The limit case. I. *Rev. Mat. Iberoamericana*, 1(1):145–201, 1985.

[71] P.-L. Lions. The concentration-compactness principle in the calculus of variations. The limit case. II. *Rev. Mat. Iberoamericana*, 1(2):45–121, 1985.

[72] R. Lucchetti. *Convexity and well-posed problems*, volume 22 of *CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC*. Springer, New York, 2006.

[73] J. Malý, D. Swanson, and W. P. Ziemer. Fine behavior of functions whose gradients are in an Orlicz space. *Studia Math.*, 190(1):33–71, 2009.

[74] A. Mareno and T. J. Healey. Global continuation in second-gradient nonlinear elasticity. *SIAM J. Math. Anal.*, 38(1):103–115, 2006.

[75] V. Maz'ya. *Sobolev spaces with applications to elliptic partial differential equations*, volume 342 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, augmented edition, 2011.

[76] F. Morgan. A smooth curve in R^4 bounding a continuum of area minimizing surfaces. *Duke Math. J.*, 43(4):867–870, 1976.

[77] F. Morgan. Almost every curve in $\mathbf{R}^3$ bounds a unique area minimizing surface. *Invent. Math.*, 45(3):253–297, 1978.

[78] F. Morgan. Generic uniqueness for hypersurfaces minimizing the integral of an elliptic integrand with constant coefficients. *Indiana Univ. Math. J.*, 30(1):29–45, 1981.

[79] C. B. Morrey, Jr. Quasi-convexity and the lower semicontinuity of multiple integrals. *Pacific J. Math.*, 2:25–53, 1952.

[80] S. Müller and V. Šverák. Convex integration for Lipschitz mappings and counterexamples to regularity. *Ann. of Math. (2)*, 157(3):715–742, 2003.

[81] I. Namioka and R. R. Phelps. Banach spaces which are Asplund spaces. *Duke Math. J.*, 42(4):735–750, 1975.

[82] P. Pedregal. *Parametrized measures and variational principles*, volume 30 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Verlag, Basel, 1997.

[83] R. R. Phelps. *Convex functions, monotone operators and differentiability*, volume 1364 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, second edition, 1993.

[84] D. Preiss and L. Zajíček. Stronger estimates of smallness of sets of Fréchet nondifferentiability of convex functions. In *Proceedings of the 11th winter school on abstract analysis (Železná Ruda, 1983)*, number Suppl. 3, pages 219–223, 1984.

[85] M. M. Rao. *Measure theory and integration*, volume 265 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, second edition, 2004.

[86] D. L. Renfro. Porosity, nowhere dense sets and a theorem of Denjoy. *Real Anal. Exchange*, 21(2):572–581, 1995/96.

[87] F. Rindler. *Calculus of variations*. Universitext. Springer, Cham, 2018.

[88] R. T. Rockafellar. Integral functionals, normal integrands and measurable selections. In *Nonlinear operators and the calculus of variations (Summer School, Univ. Libre Bruxelles, Brussels, 1975)*, pages 157–207. Lecture Notes in Math., Vol. 543. 1976.

[89] C. G. Simader. *On Dirichlet's boundary value problem*. Lecture Notes in Mathematics, Vol. 268. Springer-Verlag, Berlin-New York, 1972. An L^p-theory based on a generalization of Gårding's inequality.

[90] E. N. Spadaro. Non-uniqueness of minimizers for strictly polyconvex functionals. *Arch. Ration. Mech. Anal.*, 193(3):659–678, 2009.

[91] C. Stegall. Optimization of functions on certain subsets of Banach spaces. *Math. Ann.*, 236(2):171–176, 1978.

[92] E. M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[93] M. Struwe. *Variational methods*, volume 34 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer-Verlag, Berlin, fourth edition, 2008. Applications to nonlinear partial differential equations and Hamiltonian systems.

[94] M. A. Sychev. Attainment and relaxation results in special classes of deformations. *Calc. Var. Partial Differential Equations*, 19(2):183–210, 2004.

[95] M. A. Sychev. Sets of lower semicontinuity and stability of integral functionals. *J. Math. Pures Appl. (9)*, 84(8):985–1014, 2005.

[96] L. Székelyhidi, Jr. The regularity of critical points of polyconvex functionals. *Arch. Ration. Mech. Anal.*, 172(1):133–152, 2004.

[97] A. Taheri. On Artin's braid group and polyconvexity in the calculus of variations. *J. London Math. Soc. (2)*, 67(3):752–768, 2003.

[98] A. Taheri. Quasiconvexity and uniqueness of stationary points in the multidimensional calculus of variations. *Proc. Amer. Math. Soc.*, 131(10):3101–3107, 2003.

[99] A. Taheri. Local minimizers and quasiconvexity—the impact of topology. *Arch. Ration. Mech. Anal.*, 176(3):363–414, 2005.

[100] L. Tartar. Compensated compactness and applications to partial differential equations. In *Nonlinear analysis and mechanics: Heriot-Watt Symposium, Vol. IV*, volume 39 of *Res. Notes in Math.*, pages 136–212. Pitman, Boston, Mass.-London, 1979.

[101] L. Tartar. The compensated compactness method applied to systems of conservation laws. In *Systems of nonlinear partial differential equations (Oxford, 1982)*, volume 111 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 263–285. Reidel, Dordrecht, 1983.

[102] H. Triebel. *Theory of function spaces. III*, volume 100 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 2006.

[103] H. Triebel. *Function spaces and wavelets on domains*, volume 7 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2008.

[104] S. L. Troyanski. On locally uniformly convex and differentiable norms in certain non-separable Banach spaces. *Studia Math.*, 37:173–180, 1970/71.

[105] M. Volle and C. Zălinescu. Strongly adequate functions on Banach spaces. *J. Convex Anal.*, 20(3):655–668, 2013.

[106] B. White. The bridge principle for unstable and for singular minimal surfaces. *Comm. Anal. Geom.*, 2(4):513–532, 1994.

[107] K.-O. Widman. Hölder continuity of solutions of elliptic systems. *Manuscripta Math.*, 5:299–308, 1971.

[108] K. Yosida. *Functional analysis*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the sixth (1980) edition.

[109] L. Zajíček. On σ-porous sets in abstract spaces. *Abstr. Appl. Anal.*, (5):509–534, 2005.

[110] K. Zhang. Energy minimizers in nonlinear elastostatics and the implicit function theorem. *Arch. Rational Mech. Anal.*, 114(2):95–117, 1991.

[111] K. Zhang. Remarks on quasiconvexity and stability of equilibria for variational integrals. *Proc. Amer. Math. Soc.*, 114(4):927–930, 1992.

Index

E_p, 94
$E_{p,w}$, 114
F_w, 96, 109
$J(u)$, 10, 94
$J(u;E)$, 65
J_ε, 13, 95
$T\colon X \rightrightarrows T$, 14
T^{-1} (for multivalued functions), 14
V_p, 114
$\operatorname{Dom} T$, 14
$\mathrm{L}^p(\Omega,\mu;X)$, 144
$\mathrm{L}^p_w(\Omega,\mu;X)$, 144
$\mathcal{N}(x)$, 14
$\mathcal{S}$, 90
$\mathcal{S}^\varepsilon$, 90
$\Gamma(X)$, 44
$\mathrm{W}^{1,p}_g(\Omega;\mathbb{R}^N)$, 11, 13
$\mathrm{W}^{2,2}_g(\Omega;\mathbb{R}^N)$, 13
$\mathrm{W}^{-1,p'}(\Omega;\mathbb{R}^N)$, 39
α-angle small, 43
α-cone meagre, 43
$\operatorname{argmin}(X;I)$, 15

$\langle\cdot\rangle$, 115
$\overline{\text{conv}}$
  for $\overline{\mathbb{R}}$-valued functions, 134
  for multivalued functions, 134
  for subsets of a vector space, 137
cont , 44
∂f (subdifferential), 15, 45
$\partial_\varepsilon f$, 45
dom , 44
$\overline{\mathbb{R}}$, 10
ε-$\operatorname{argmin}(X;I)$, 15

$\min(X;I)$, 12
∇f (Fréchet derivative), 47
$\nabla^2 u$, 8, 13, 95
σ-porosity, 42
d-fragmentability, 133
d-rectifiability, 95
f^* (Fenchel conjugate), 46
p-growth, 27, 65, 94
p', 12, 39
p^*, 25

Asplund spaces, 49

biorthogonal sequences, 126

Carathéodory integrand, 24
coercivity
  coercivity for a function I, 23
  coercivity for a functional J, 65
  pointwise strong coercivity for an integrand F, 55
  strong coercivity for a functional J, 38, 51
CUSCO, 17
  minimal CUSCO, 17

dyadic cubes, 61

equi-integrablility, 139

Fenchel conjugate, 46

internal cusp, 62

meagre, 41
multivalued functions, 14

nonatomic, 140
normal integrand, 24
nowhere dense, 41

polyconvexity, 29
proper, 23, 44

quasiconvexity, 27
  strong quasiconvexity, 94
  uniform strict quasiconvexity, 58, 65, 90

rank-one convexity, 29
residual, 41
RNP (Radon–Nikodym Property), 146
    for closed and bounded convex sets, 146
    for closed but unbounded convex sets, 146
    geometric form, 147

subdifferential, 45

uniform integrability, 139
USC, 16
USCO, 17
    minimal USCO, 17

well-posedness
    generalized Hadamard well-posedness, 20
    generic well-posedness, 22
    Hadamard well-posedness, 20
    Tykhonov well-posedness, 18

Young measures, 32
    gradient Young measures, 34
    homogenous gradient Young measures, 34