CONCENTRATION AND COMPENSATED COMPACTNESS TECHNIQUES IN PDE

SAMANTHAK THIAGARAJAN

ABSTRACT. We discuss the applications of the calculus of variations to solving various nonlinear problems. We emphasize the role of compactness, which provides us with strong convergence results. In problems where there is a lack of compactness, we discuss applications of concentration compactness techniques, which analyze exactly why we do not have strong convergence and what limiting behavior we can expect from our sequences. Finally, for problems that present us with additional structure, such as convexity, we discuss compensated compactness principles, which use these additional tools to obtain convergence results. We assume knowledge of real analysis and introductory functional analysis.

CONTENTS

1. INTRODUCTION AND BACKGROUND INFORMATION

Compactness is a key tool used everywhere in analysis, especially optimization problems. Given a bounded set of numbers, we may extract a convergent subsequence, which may then be an appropriate optimizer for such a problem.

However, with sequences of functions, there are complex behaviors we have to take into account that cause convergence to fail. In general, there are four ways a sequence of functions can fail to converge: it may concentrate onto a smaller and smaller set, disperse to infinity, translate and move towards infinity while retaining its shape, or oscillate with rapidly increasing oscillations.

Date: August 31, 2023.

The principle of concentration compactness characterizes the types of limiting behavior a weakly converging sequence of functions can have. Broadly speaking, this principles says are three "fundamental" limiting behaviors for such a sequence $\{u_n\}$:

(1) $\{u_n\}$ concentrates onto a compact set (Concentration).
(2) $\{u_n\}$ vanishes, so its mass escapes to infinity (Vanishing).
(3) $\{u_n\}$ splits into two masses diverging away from each other (Dichotomy).

When analyzing certain variational problems, when we do not have access to compactness, we can use this principle. In practice, the best case we can hope for is the concentration case, as this resembles compactness. In addition, this restricts us to a bounded domain, which rules out specific types of limiting behavior, and on which we have tools we can use to obtain convergence results. However, a sequence concentrating onto too small of set may converge weakly to a measure, such as a Dirac mass, whereas we want convergence to a function. So, we have to ensure this case does not occur as well.

In variational problems, we take a minimizing sequence and try to prove convergence, in some sense, to a minimizer. In doing so, we try to rule out vanishing and dichotomy, as we cannot obtain convergence results in those cases, and study the concentration case.

Another topic we discuss is compensated compactness. This is a technique that aims to find and use hidden structure present in nonlinear PDE to obtain convergence results. An example of this is the Div-Curl Lemma. In general, there is nothing we can say about the product of weakly converging sequences. However, if we have a modest assumption on certain linear combinations of derivatives, then we can say that the product of these sequences does indeed converge weakly to the product of weak limits.

We now present several important definitions and results that will be used throughout the paper following [3].

Definition 1.1 (Sobolev Space)**.** The Sobolev Space $W^{1,p}(\mathbb{R}^d)$ is the set of functions for which the function and its weak derivative are in L^p. We define

$$||u||_{W^{1,p}(\mathbb{R}^d)} = (||u||^p_{L^p(\mathbb{R}^d)} + ||Du||^p_{L^p(\mathbb{R}^d)})^{1/p},$$

for $1 \leq p < \infty$. This can be extended analogously for $p = \infty$, where

$$||u||_{W^{1,\infty}(\mathbb{R}^d)} = |||u| + |Du|||_{L^\infty(\mathbb{R}^d)},$$

and can be extended to $W^{k,p}(\mathbb{R}^d)$, where we have more derivatives. We also define $W^{k,p}_0(U)$, which is the set of $W^{k,p}(U)$ functions that are also 0 on ∂U.

Remark 1.2. To appropriately define functions on a measure zero boundary and functions defined almost everywhere, we use the trace operator, a bounded linear operator that maps functions in $W^{1,p}(U)$ to $L^p(\partial U)$.

Theorem 1.3 (Sobolev Inequality)**.** *Assume $1 \leq p < d$. Then, there exists $C = C(p, d)$ such that*

$$||u||_{L^{p^*}(\mathbb{R}^d)} \leq C||Du||_{L^p(\mathbb{R}^d)},$$

for all $u \in C^1_c(\mathbb{R}^d)$, where $p^* = pd/(p - d)$.

Remark 1.4. We provide a brief discussion of how this exponent arises. First, assume we have an inequality such as

$$||u||_{L^q(\mathbb{R}^d)} \leq C||Du||_{L^p(\mathbb{R}^d)}.$$

We show there is only one such q for which such an inequality holds, namely $q = p^*$. To do so, let $u \in C_c^\infty(\mathbb{R}^d)$ that is not identically zero and let $\lambda \in \mathbb{R}$ be positive. Now, define $u_\lambda(x) = u(\lambda x)$. We note that

$$\int_{\mathbb{R}^d} |u_\lambda(x)|^q \, dx = \int_{\mathbb{R}^d} |u(\lambda x)|^q \, dx = \frac{1}{\lambda^d} \int_{\mathbb{R}^d} |u(t)|^q \, dt.$$

A similar calculation shows $||Du_\lambda||_p^p = \lambda^{p-d}||Du||_p^p$. Applying the inequality to $u_\lambda(x)$, we see that

$$\frac{1}{\lambda^{d/q}}||u||_{L^q(\mathbb{R}^d)} \leq C\frac{\lambda}{\lambda^{d/p}}||Du||_{L^p(\mathbb{R}^d)} \implies ||u||_{L^q(\mathbb{R}^d)} \leq C\lambda^{1+d/q-d/p}||Du||_{L^p(\mathbb{R}^d)}.$$

Now, if $1 + d/q - d/p > 0$, we may let $\lambda \to 0$ to obtain that $||u||_{L^q(\mathbb{R}^d)} = 0$, a contradiction, and similarly let $\lambda \to \infty$ if $1 + d/q - d/p < 0$. Therefore, $1 + d/q - d/p = 0$, so $q = dp/(d - p)$. This exponent is known as p^*, the Sobolev conjugate.

Remark 1.5. If $U \subset \mathbb{R}^d$ is bounded, open, and with a C^1 boundary, then the analog of the Sobolev Inequality for functions $u \in W^{1,p}(U)$ is

$$||u||_{L^{p^*}(U)} \leq C||u||_{W^{1,p}(U)}.$$

This follows by extension.

Theorem 1.6 (Rellich-Kondrachov Compactness)**.** *If $U \subset \mathbb{R}^d$ is bounded, open, and with a C^1 boundary, with $1 \leq p < d$, then $W^{1,p}(U)$ is compactly embedded into $L^q(U)$ for all $1 \leq q < p^*$.*

Remark 1.7. We give context to this embedding as follows. Letting $\tau_h u$ be the translation of u by the vector h, the space $W^{1,p}(U)$ can alternatively be described as the functions u for which we have an estimate of the form

$$||u - \tau_h u||_p \leq C|h|,$$

for all sufficiently small h. This is a type of "L^p-Lipschitz" condition. As a technicality, because u is only defined on U, the term $u - \tau_h u$ is only defined on a slightly smaller domain than U. However, if we now have a bounded sequence of functions in $W^{1,p}$, this uniform Lipschitz property enables us to apply Arzela-Ascoli and obtain a subsequence that converges in L^p through a diagonalization argument. As we are on a bounded domain, L^p space inclusions give us a convergent subsequence in L^q for $1 \leq q < p$. Finally, to obtain this property for $p < q < p^*$, we use the L^p interpolation inequality

$$||u||_{L^q(U)} \leq ||u||_{L^p(U)}||u||_{L^{p^*}(U)}.$$

Theorem 1.8 (Poincare Inequality)**.** *Let $U \subset \mathbb{R}^d$ be bounded and open. Then, there exists $C = C(p, d, U)$ such that, for all $u \in W_0^{1,p}(U)$,*

$$||u||_{L^p(U)} \leq C||Du||_{L^p(U)}.$$

Proof. We provide a brief proof of this statement using Theorem 1.6. For the sake of contradiction, assume there exists a sequence $\{u_k\}$ such that

$$||u_k||_{L^p(U)} > k||Du_k||_{L^p(U)}.$$

This contradicts the existence of such a C because the k's are increasing and we cannot obtain a uniform bound on them. Without loss of generality, assume $||u_k||_{L^p(U)} = 1$. Then, we have that $||Du_k||_{L^p(U)} < 1/k$. We thus have boundedness in $W^{1,p}(U)$. Applying Rellich-Kondrachov, after passing to a subsequence (still denoted $\{u_k\}$), $u_k \to u$ $L^p(U)$. This strong convergence ensures $||u||_{L^p(U)} = 1$. These results ensure u has zero derivative, and so is constant. Because $||u||_{L^p(U)} = 1$, $u \neq 0$. However, this contradicts that $u \in W_0^{1,p}(U)$, as u is non-zero on the boundary. □

2. Weak Convergence

Weak convergence is an important concept in PDEs. To weaken the notion of strong convergence on a space X, we instead test convergence against members of the dual space X^*. To be specific, in $L^p(U)$ for instance, we say a sequence of functions $\{f_k\} \subset L^p(U)$ converges weakly to f if

$$\int_U f_k g \to \int_U fg,$$

for all $g \in (L^p(U))^* = L^q(U)$, where q is the Holder conjugate of p. One natural question that arises is: if ϕ is smooth, then does $\phi(f_k) \rightharpoonup \phi(f)$ whenever $f_k \rightharpoonup f$? Letting $\phi(x) = x^2$, we will show this is not the case. We define $f_n(x) = \sin(nx)$. Applying Lemma 2.2, $f_n \rightharpoonup 0$ but $f_n^2 \rightharpoonup 1/2$. We conclude that not all smooth functions are continuous with respect to weak convergence. In fact, the situtation is much worse, as the only functions that are continuous with respect to weak convergence are affine linear, which we prove below.

Lemma 2.1. *If $F(f_k) \rightharpoonup F(f)$ in $L^2(0,1)$ for all $f_k \rightharpoonup f$ in $L^2(0,1)$, then F is affine linear.*

Proof. For the sake of contradiction, assume F is not affine linear. So, there exist a, b and some $\lambda \in (0,1)$ such that

$$F(\lambda a + (1-\lambda)b) \neq \lambda F(a) + (1-\lambda)F(b).$$

We now define a sequence of functions $\{f_k\}$ as follows: partition $(0,1)$ into intervals $(i, i+1/k]$ of length $1/k$. Then, we define f_k to be a on the first λ portion of $(i, i+1/k]$ and b on the rest, and similarly for each other interval. Now, by boundedness of our sequence of functions and as step functions are dense in $L^2(0,1)$, we only need to test weak convergence against step functions. We want to show $f_k \rightharpoonup \lambda a + (1-\lambda)b$, as this is exactly the linearity we want to obtain a contradiction. Testing against the characteristic function of an interval I χ_I, we see that

$$\left| \int_0^1 f_k \chi_I - [\lambda a + (1-\lambda)b]|I| \right| \leq \frac{4\max(a,b)}{k} \to 0,$$

as the only error term comes from the two intervals $[i, i+1/k]$ that may not overlap perfectly with I. As testing against χ_I for intervals I is sufficient, we see that $f_k \rightharpoonup \lambda a + (1-\lambda)b$. Applying the same argument, we have $F(f_k) \rightharpoonup \lambda F(a) + (1-\lambda)F(b)$. However, by assumption at the beginning of the proof, $F(f_k) \rightharpoonup F(f) = F(\lambda a + (1-\lambda)b)$. As weak limits are unique, we obtain the desired contradiction. □

We also prove a lemma about rescalings and weak convergence. This gives context to the idea that weak convergence is related to "averaging".

Lemma 2.2. *Let $Q = [0,1]^d$ and $f\colon Q \to \mathbb{R}$ be Q-periodic with $f \in L^2(Q)$. Let $f^n(x) = f(nx)$. Prove $f_n \rightharpoonup (f)_Q = \int_Q f$.*

Proof. We want to show that $\int_Q f^n g \to \int_Q f \int_Q g$ for all $g \in L^2(Q)$. We first observe that f^n are uniformly bounded in $L^2(Q)$. Indeed,

$$\left(\int_Q (f^n)^2\right)^{1/2} = \left(\sum_{i=1}^{n^d} \int_{Q_i^n} f^2(nx)\, dx\right)^{1/2} = \frac{1}{n^d}\left(\sum_{i=1}^{n^d} \int_Q f^2(x)\, dx\right)^{1/2},$$

where Q_i^n is a partition of Q into n^d cubes of side length $1/n$. Note that we have used the periodicity of f. The last term on the right above is just $||f||_{L^2(Q)}$. By this uniform bound on f in $L^2(Q)$, it suffices to test weak convergence against $\phi \in C_c^\infty(Q)$, a dense subset. We now see that

$$\int_Q f^n \phi = \sum_{i=1}^{n^d} \int_{Q_i^n} f(nx)\phi(x)\, dx = \sum_{i=1}^{n^d} \frac{1}{n^d} \int_Q f(t)\phi(c_i^n + t/n)\, dt,$$

where we have made a change of variables and used the periodicity of f. Here, c_i^n denotes the bottom left corner of the cube Q_i^n. This sum almost resembles a Riemann sum. We now want to isolate the part that resembles a Riemann sum and show the other components are negligible. From here, we Taylor expand to write $\phi(c_i^n + t/n) = \phi(c_i^n) + D\phi(c_i^n) \cdot t/n + O(|t/n|^2)$. Thus,

$$\int_Q f^n \phi = \int_Q f \sum_{i=1}^{n} \frac{1}{n^d}\phi(c_i^n) + \sum_{i=1}^{n^d} \frac{1}{n^d} \int_Q f(t) D\phi(c_i^n) \cdot t/n + O(|t/n|^2)\, dt.$$

The first term approaches $\int_Q f \int_Q \phi$ as $n \to \infty$ by definition of the Riemann integral. For the second term, we bound $|t| \leq 1$ to obtain the second term is less, in absolute value, than C/n ($C \in \mathbb{R}$), which tends to 0 as $n \to \infty$. The third term also tends to zero in a similar manner. □

Remark 2.3. While we proved Lemma 2.2 for the rescalings $f(nx)$ for $n \in \mathbb{N}$, it is also true for $f(x/\varepsilon)$ for $\varepsilon > 0$. The only difference in the proof is the handling of a negligible error term.

3. Concentration Compactness

3.1. **Compactness on Bounded Domains.**

In light of 1.3, the Sobolev inequality, we may ask whether there is an optimal constant for that inequality and whether this is attained for some specific function u. To obtain this constant, we consider the following:

$$\inf\left\{\frac{||Du||_p}{||u||_q} \mid u \in W_0^{1,p}(U), u \neq 0\right\}.$$

By the Sobolev inequality, we always have $||Du||_p/||u||_q \geq C$, and so we take the infimum to obtain the optimal constant. We now present the following theorem.

Theorem 3.1. *For $1 \leq q < p^*$, there is an optimal constant for the Sobolev inequality and it is attained for some $u \in W_0^{1,p}(U)$.*

Proof. We begin by noting that we may instead consider the simplified minimization problem

$$\min\left\{\int_U |Du|^p \mid u \in W_0^{1,p}(U), \int_U |u|^q = 1\right\}.$$

Here, we have also raised the norms to their respective powers for simplicity. Let m denote the infimum. To show it is attained, and in fact a minimum, we take a minimizing sequence $\{u_n\}_{n\in\mathbb{N}}$ so $\int_U |Du_n|^p \to m$. The minimum is finite, which follows by considering any $u \in C_c^\infty(U)$. After passing to a subsequence (which we continue to denote by $\{u_n\}$), this ensures $\{Du_n\}$ is bounded in L^p. The Poincare inequality ensures this controls the L^p norms of $\{u_n\}$, and so $\{u_n\}$ is bounded in $W^{1,p}$. This boundedness provides us with a subsequence (still denoted by $\{u_n\}$) that converges weakly in $W_0^{1,p}(U)$ to some $u \in W_0^{1,p}(U)$. As we are on a bounded domain, we apply the Rellich-Kondrachov compactness theorem, so $u_n \to u$ strongly in $L^q(U)$ and $L^p(U)$. As $\int_U |u_n|^q = 1$ for all n, strong convergence implies that $\int_U |u|^q = 1$, and so the limit u is admissible. In addition, as $I[\cdot] = ||D(\cdot)||_p^p$ and $I[u_n] \to m$, Rellich-Kondrachov gives that $I[u] = m$. Thus, u is a minimizer.

Before we conclude, we present the PDE that the minimizer u solves. To find this PDE, we use a parameterization technique that is discussed in thorough detail in Theorem 3.7. We find that

$$\Delta_p u = \lambda u^{q-1}$$

for some $\lambda > 0$, where Δ_p is the p-Laplacian, defined as

$$\Delta_p u = \operatorname{div}(|Du|^{p-2} Du).$$

□

3.2. **Abstract Principle.** We have seen that the Rellich-Kondrachov compactness theorem played a critical role in showing the minimizer was attained in the constraint set. After taking a minimizing sequence, we used this compactness theorem to show the weak limit of the sequence had an L^q norm of 1. We do not have access to this theorem on unbounded domains. The general principle of concentration compactness is concerned with what can happen with a sequence of functions with fixed mass.

Indeed, consider a sequence $\{u_n\} \subset L^2(\mathbb{R}^d)$ such that $\int_{\mathbb{R}^d} |u_n|^2 = \lambda$ for all n, and assume $u_n \rightharpoonup 0$ in $L^2(\mathbb{R}^d)$. Following [4], we can imagine a few scenarios for how this occurs.

(1) u_n disperses. So, u_n approaches 0 a.e. and the L^2 mass spreads out as $n \to \infty$. An example is $u_n(x) = e^{-(x/n)^2}/n$.
(2) u_n "moves" to infinity. So, u_n still approaches 0 a.e. but the L^2 mass shifts arbitrarily far away as $n \to \infty$. An example is $u_n = \chi_{[n,n+1]}$.
(3) u_n concentrates. So, as $n \to \infty$ the L^2 mass is contained in a smaller and smaller set, with larger and larger spikes. An example is $u_n(x) = ne^{-(nx)^2}$.
(4) u_n oscillates. So, as $n \to \infty$, u_n has more and more high oscillations. An example is $u_n(x) = \sin(2n\pi x)\chi_{[0,1]}$.

On a bounded domain, the first two possibilities cannot occur. So, on an unbounded domain, not only do we not have access to the powerful Rellich-Kondrachov compactness theorem, but there are additional, complicating behaviors that u_n can exhibit that make analyzing such a sequence difficult. We now prove the following theorem which characterizes all the possible limiting behaviors that a sequence with fixed $L^1(\mathbb{R}^d)$ mass can have.

Theorem 3.2. *Let ρ_n be a non-negative sequence of real-valued functions on $\mathbb{R}^d$ with $||\rho_n||_{L^1(\mathbb{R}^d)} = \lambda$ for all n. Then, there exists a subsequence $\{\rho_{n_k}\}$ for which we have exactly one of following possibilities:*

1. *(Compactness) There exist translations $\{y_n\}$ such that for all $\varepsilon > 0$, there exists some R such that $\sup_k ||\rho_{n_k}||_{L^1(B_R(y_k))} > \lambda - \varepsilon$.*
2. *(Vanishing) For all $R > 0$,*
$$\lim_{k\to\infty} \sup_{y\in\mathbb{R}^d} \int_{B_R(y)} \rho_{n_k} = 0.$$
3. *(Dichotomy) Two diverging humps arise. So, there exists $0 < \alpha < \lambda$ and $\rho_k^1, \rho_k^2 > 0$, in L^1, such that for all $\varepsilon > 0$, there exists k_0 such that, if $k \geq k_0$, then*
$$||(\rho_k^1 + \rho_k^2) - \rho_{n_k}||_{L^1} < \varepsilon, \ \left|\int_{\mathbb{R}^d} \rho_k^1 - \alpha\right| < \varepsilon, \ \left|\int_{\mathbb{R}^d} \rho_k^2 - (\lambda - \alpha)\right| < \varepsilon$$

.

3.3. **Failure of Compactness.** We now discuss the embedding $W^{1,q}(U) \subset L^{q^*}(U)$, for $1 \leq q < n$, which follows from the Sobolev inequality, and illustrate how it is, in fact, not a compact embedding.

Theorem 3.3. *The embedding $W^{1,p}(B_2) \subset L^{p^*}(B_2)$ is not compact (B_2 denotes the ball of radius 2 centered at the origin).*

Proof. We begin by defining $u_\varepsilon(x) = u(x/\varepsilon)$ for some $u \in W^{1,p}(B_2)$ such that $\text{supp}(u) \subset B_1$ and $u \neq 0$ everywhere. We now define a family of functions $v_\varepsilon = \varepsilon^{-d/p^*} u_\varepsilon$. As u is supported in B_1, as ε gets smaller, as soon as $x/\varepsilon \notin B_1$, $u_\varepsilon(x) = 0$, and so $v_\varepsilon \to 0$ a.e. By change of variables, $||v_\varepsilon||_{L^{p^*}(B_1)} = ||u||_{L^{p^*}(B_1)}$ and $||Dv_\varepsilon||_{L^p(B_1)} = ||Du||_{L^p(B_1)}$. As discussed with the Sobolev Inequality, p^* is the only exponent that ensures these rescalings are the same. We then observe that
$$||v_\varepsilon||^p_{L^p(B_1)} = \int_{B_1} \varepsilon^{-dp/p*} u^p(x/\varepsilon)\, dx = \varepsilon^{d(1-p/p^*)} \int_{B_1} u^p(y)\, dy,$$
and therefore $\{v_\varepsilon\}$ is bounded in $W^{1,p}(B_1)$. We have thus obtained the desired contradiction, as $v_\varepsilon \to 0$ a.e. but $\{v_\varepsilon\}$ has constant $L^{p^*}(B_1)$ norm and is bounded in $W^{1,p}(B_1)$. □

We now present a theorem following [2] characterizing exactly how this compactness fails.

Theorem 3.4. *Assume $1 < p < d$,*
$$f_k \to f \text{ in } L^p_{loc}\ , \quad Df_k \rightharpoonup Df \text{ in } L^p(\mathbb{R}^d; \mathbb{R}^d).$$
Also, assume
$$|Df_k|^p \rightharpoonup \mu \text{ in } \mathcal{M}(\mathbb{R}^d)\ , \quad |f_k|^{p^*} \rightharpoonup \nu \text{ in } \mathcal{M}(\mathbb{R}^d).$$
Then, there exists a countable set J, points $\{x_j\}_{j\in J} \subset \mathbb{R}^d$ and non-negative weights $\{\mu_j, \nu_j\}_{j\in J}$ such that
$$\nu = |f|^{p^*} + \sum_{j\in J} \nu_j \delta_{x_j} \quad \mu \geq |Df|^p + \sum_{j\in J} \mu_j \delta_{x_j}.$$

In addition,

$$\nu_j \leq C_p^{p^*} \mu_j^{p^*/p},$$

where C_p is the optimal constant for the Sobolev inequality. Finally, if $f \equiv 0$ and $\nu(\mathbb{R}^d)^{1/p^*} \geq C_p \mu(\mathbb{R}^d)^{1/p}$, then ν is concentrated at a single point.

Proof. To begin, consider the case where $f \equiv 0$, with $\phi \in C_c^\infty(\mathbb{R}^d)$. Applying the Sobolev inequality,

$$\left(\int_{\mathbb{R}^d} |\phi f_k|^{p^*} \right)^{1/p^*} \leq C_p \left(\int_{\mathbb{R}^d} |D(\phi f_k)|^p \right)^{1/p}.$$

Applying the product rule, $D(\phi f_k) = f_k D\phi + \phi D f_k$. Noting that $f \equiv 0$ and the definitions of ν and μ,

$$\left(\int_{\mathbb{R}^d} |\phi|^{p^*} \, d\nu \right)^{1/p^*} \leq C_p \left(\int_{\mathbb{R}^d} |\phi|^p \, d\mu \right)^{1/p}.$$

As we arbitrarily chose $\phi \in C_c^\infty(\mathbb{R}^d)$, for any Borel set E, we obtain $\nu(E)^{1/p^*} \leq C_p \mu(E)^{1/p}$ by approximating χ_E. Consider the set $D = \{x \in \mathbb{R}^d \mid \mu(x) > 0\}$. By definition of μ, as weakly convergent sequences are bounded, $\mu(\mathbb{R}^d) < \infty$. Thus, D is not uncountable and there exist points $\{x_j\}$ and weights $\{\mu_j\}$ such that $\mu \geq \sum_j \mu_j \delta_{x_j}$. We don't have any information on the behavior of μ off of these Dirac masses, so we cannot obtain an equality, but rather an inequality. We now want to obtain a characterization of the measure ν.

Recalling the definition of absolute continuity in the context of measures, we see that ν is absolutely continuous with respect to μ. Invoking the Radon-Nikodym Theorem, there exists an L^1 function $D_\mu \nu$ such that $\nu(E) = \int_E D_\mu(\nu) \, d\mu$, where, by the Lebesgue Differentiation Theorem, μ-a.e.,

$$D_\mu \nu(x) = \lim_{r \to 0} \frac{\nu(B_r(x))}{\mu(B_r(x))}.$$

As long as $\mu(B_r(x)) \neq 0$,

$$\frac{\nu(B_r(x))}{\mu(B_r(x))} \leq C_p^{p^*} \mu(B_r(x))^{p^*/(p-1)}.$$

Now, for any $x \notin D$, by definition of D, we see that $\mu(B_r(x)) \to 0$ as $r \to 0$, giving that $D_\mu \nu(x) = 0$ if $x \notin D$, μ-a.e. The same reasoning shows that ν is a finite measure, and thus the existence of weights $\{\nu_j\}$ with $\nu_j = D_\mu \nu(x_j) \mu_j$ such that $\nu = \sum_j \nu_j \delta_{x_j}$. This completes the proof if $f \equiv 0$.

Next, assume additionally that $\nu(\mathbb{R}^d)^{1/p^*} \geq C_p \mu(\mathbb{R}^d)^{1/p}$. We thus have that $\nu(\mathbb{R}^d)^{1/p^*} = C_p \mu(\mathbb{R}^d)^{1/p}$. We thus see that

$$\left(\int_{\mathbb{R}^d} |\phi|^{p^*} \, d\nu \right)^{1/p^*} \leq C_p \left(\int_{\mathbb{R}^d} |\phi|^p \, d\mu \right)^{1/p} \leq C_p \left(\int_{\mathbb{R}^d} |\phi|^{p^*} \, d\mu \right)^{1/p^*} \mu(\mathbb{R}^d)^{1/n},$$

where we have proceeded by applying Holder's Inequality with p^*/p and n/p (which are Holder conjugates). By approximation, this gives

$$\nu(E) = C_p^{p^*} \mu(\mathbb{R}^d)^{n/(n-p)} \mu(E) = C_p^p \nu(\mathbb{R}^d)^{p/n} \mu(E).$$

We see that

$$\left(\int_{\mathbb{R}^d} |\phi|^{p^*} \, d\nu \right)^{1/p^*} \leq C_p \left(\int_{\mathbb{R}^d} |\phi|^p \, d\mu \right)^{1/p} = \nu(\mathbb{R}^d)^{-1/n} \left(\int_{\mathbb{R}^d} |\phi|^p \, d\nu \right)^{1/p}.$$

Approximating again, we see that for any E such that E is Borel, $\nu(E)^{1/p^*}\nu(\mathbb{R}^d)^{1/n} \leq \nu(E)^{1/p}$. If $\nu(E) \neq 0$, then we see that $\nu(\mathbb{R}^d) \leq \nu(E)$, and as $\nu(E) \leq \nu(\mathbb{R}^d)$ by properties of measures, $\nu(E) = \nu(\mathbb{R}^d)$. We now claim this implies ν is concentrated at a single point. For the sake of contradiction, assume we have a rectangle E such that $\nu(E) > 0$. Now, we split E into two disjoint Borel sets, E_1 and E_2. Consider the case where $\nu(E_1), \nu(E_2) > 0$. We know $\nu(E_1) = \nu(E_2) = \nu(\mathbb{R}^d)$. Additionally, $\nu(E_1) + \nu(E_2) = \nu(E) = \nu(\mathbb{R}^d)$. Thus, $\nu(\mathbb{R}^d) = 0$, which is impossible. Thus, only one of $\nu(E_1)$ and $\nu(E_2)$ is positive. Repeatedly applying this process shows ν is concentrated at a single point.

For the general case, we define an auxiliary function $g_k = f_k - f$. As we now have that $g_k \to 0$ strongly in L^p_{loc}, the results for the zero case hold for g_k. Now, $|Dg_k|^p \rightharpoonup \mu - |Df|^p$ and $|g_k|^{p^*} \rightharpoonup \nu - |f|^{p^*}$ in $\mathcal{M}(\mathbb{R}^d)$, which follow by Lemma 3.5, proved below. This completes the proof.

□

3.4. **Minimizer on Global Domain.** As we discussed earlier, we cannot simply apply the Rellich compactness theorem to obtain a strong convergence result on unbounded domains because minimizing sequences can exhibit complicated limiting behaviors. Instead, we have to resort to different techniques. Before doing so, we will prove the Brezis-Lieb Lemma, which is an improvement of Fatou's Lemma. This lemma allows us to express the limit of the L^q norms of a sequence $\{f_k\}$ in terms of the L^q norms of $f_k - f$ and f. We take inspiration from [1].

Lemma 3.5. *If $f_k \rightharpoonup f$ in $L^q(U)$, $f_k \to f$ a.e. in U, and $1 \leq q < \infty$, then*

$$\lim_{k\to\infty}(||f_k||_q^q - ||f_k - f||_q^q) = ||f||_q^q.$$

Proof. We first note that we can prove the estimate

$$|a+b|^q \leq (\varepsilon+1)|a|^q + C(\varepsilon)|b|^q,$$

where C is a constant depending on ε and q.

To prove this, first assume $a < b$. Letting $f(x) = x^q$, the Mean Value Theorem implies there exists a $c \in (a, b)$ such that

$$(a+b)^q - b^q = qac^{q-1} < q(a\delta)\left(\frac{b}{\delta^{1/(q-1)}}\right)^{q-1} \leq \delta^q a^q + \frac{b^q(q-1)}{\delta^{q/(q-1)^2}}.$$

We can obtain the desired inequality from this inequality by appropriating choosing δ and applying Young's inequality in a different order.

From here, we define functions $\{g_k^\varepsilon\}$ to be such that

$$g_k^\varepsilon = ||f_k|^q - |f_k - f|^q - |f|^q| - \varepsilon|f_k - f|^q.$$

As $f_k \to f$ a.e., $g_k^\varepsilon \to 0$ a.e. as $k \to \infty$. The inequality at the beginning of the proof allows to say $g_k^\varepsilon \leq C_\varepsilon|f|^q$. Thus, g_k^ε is bounded from above by a function in L^q, which is in L^1 as we are on a bounded domain. In addition, $g_k^\varepsilon \to 0$ a.e. Thus, by the Lebesgue Dominated Convergence Theorem, $g_k^\varepsilon \to 0$ in L^1. Thus,

$$\limsup_{k\to\infty}\int_U ||f_k|^q - |f_k - f|^q - |f|^q| \leq \varepsilon \sup_k \int_U |f_k - f|^q.$$

The boundedness of $\{f_k\}$ and f in L^q ensures that the term on the left is, in fact, arbitrarily small. This completes the proof. □

Before we proceed to solving two minimization problems, we present one more theorem. The importance of this theorem is to addressing the vanishing condition that can happen with functions with fixed L^1 mass. Importantly, it says that if we have functions disappearing in L^p of a ball of radius 1 over all of $\mathbb{R}^d$, then we can obtain strong convergence to 0 in L^q for certain q.

Theorem 3.6. *Let $p < d$. Assume u_k is a bounded sequence of functions in $W^{1,p}(\mathbb{R}^d)$ and it satisfies*

$$\lim_{k\to\infty} \sup_{x\in\mathbb{R}^d} \int_{B_1(x)} |u_k(y)|^p \, dy = 0,$$

Then $u_k \to 0$ in $L^q(\mathbb{R}^d)$ for any $q \in (p, p^*)$.

Proof. Using the interpolation inequality for L^p norms and the Sobolev inequality, for any $q \in (2, 2^*)$ we see that

$$||u||_{L^q} \leq ||u||_p^\theta ||u||_{p^*}^{1-\theta} \leq C||u||_p^\theta ||Du||_p^{1-\theta},$$

for some $\theta \in (0,1)$. Thus, it suffices to prove the result for any $q \in (p, p^*)$, as the inequalities above then give us the convergence for all of (p, p^*). We now consider the sequence u_k. Let Q denote an integer cube of length 1. Choosing $q \in (p, p^*)$ so $q(1-\theta)/p^* = p/p^*$,

$$\int_Q |u_k|^q \leq C \left(\int_Q |u_k|^p \right)^{q\theta/p} ||Du_k||_{L^{p^*}(Q)}^p.$$

Noting that we can cover $\mathbb{R}^d$ by countably many cubes Q_n,

$$\begin{aligned} ||u_k||_{L^q(\mathbb{R}^d)}^q &\leq C \sum_n \left(\int_{Q_n} |u_k|^p \right)^{q\theta/p} ||Du_k||_{L^{p^*}(Q_n)}^p \\ &\leq C \sup_n \left(\int_{Q_n} |u_k|^p \right)^{q\theta/p} \sup_k ||u_k||_{W^{1,p}(\mathbb{R}^d)}^p. \end{aligned}$$

Here, we note that u_k is a bounded sequence in $W^{1,p}(\mathbb{R}^d)$. Then, by the assumption of the proof, we see that the supremum over n goes to zero as $k \to \infty$. Thus, $u_k \to 0$ in L^p for any $p \in (p, p^*)$. □

We can now use these theorems and concentration compactness techniques to solve the following minimization problem.

Theorem 3.7. *If $p < d$, then for $p < q < p^*$,*

$$\min \left\{ \int_{\mathbb{R}^d} |Du|^p + |u|^p \mid \int_{\mathbb{R}^d} |u|^q = \lambda \right\},$$

is attained.

Proof. Let $E_\lambda = \{u \in W^{1,p} \mid \int_{\mathbb{R}^d} |u|^q = \lambda\}$, $I[u] = \int_{\mathbb{R}^d} |Du|^p + |u|^p$, and $J(\lambda) = \min_{E_\lambda} I[u]$.

Step 1: Our first step is to reduce the problem of finding the minimizer for $J(\lambda)$ to $J(1)$. To do so, let $v \in E_1$. Now, we let $u(x) = \lambda^\alpha v(x)$, where $\lambda > 0$. We want to choose α so $u \in E_\lambda$. If $||u||_q^q = \lambda$, then

$$\int_{\mathbb{R}^d} |\lambda^\alpha v(x)|^q \, dx = \lambda^{\alpha q} \int_{\mathbb{R}^d} |v(x)|^q \, dx = \lambda,$$

so $\lambda^{\alpha q-1} = 1$, and so $\alpha = 1/q$. Substituting into $I[u]$, we obtain that the common coefficient on each term is $\lambda^{p/q}$. We thus see that we can apply this invertible transformation to move from any sequence in E_λ to E_1. In other words, there is a sequence in E_1 corresponding to E_λ which induces a coefficient in the action of the functional I on the sequence. Thus, we see that $J(\lambda) = \lambda^{p/q} J(1)$.

Step 2: We now intend to prove some bounds on $J(1)$. The absolute values ensure $J(1) \geq 0$. We want to show this inequality is strict. To see this, we first apply the L^p interpolation inequality to some $u \in E_1$, so for some $\theta \in (0,1)$,

$$1 = ||u||_q \leq ||u||_p^\theta ||u||_{p^*}^{1-\theta} \leq C||u||_p^\theta ||Du||_p^{1-\theta} \leq CI[u],$$

where the last step follows from Young's inequality. Thus, we have $1/C \leq I[u]$ for all $u \in E_1$, and so $J(1) > 0$. Then, choosing $u \in C_c^\infty(\mathbb{R}^d)$, $J(1) < \infty$.

Step 3: Now, let $\{u_n\}$ be a minimizing sequence. We will pass to several subsequences, but still denote our subsequences by $\{u_n\}$ for brevity. As $\{u_n\}$ is bounded in $W^{1,p}(\mathbb{R}^d)$, by reflexivity, we may pass to a weakly convergent subsequence, and so $u_n \rightharpoonup u$ in $W^{1,p}(\mathbb{R}^d)$. By definition of the $W^{1,p}$ norm, this implies that $u_n \rightharpoonup u$ in L^p and $Du_n \rightharpoonup Du$ in L^p. In addition, by assumption, u_n is bounded in L^q, so $u_n \rightharpoonup u$ in L^q. The Rellich compactness theorem gives us strong convergence in L^p_{loc}, which in turn gives us pointwise convergence a.e.

Step 4: We now apply the Lemma 3.5. Because $u_n \rightharpoonup u$ in L^p and $u_n \to u$ pointwise a.e.,

$$||u_n||_p^p = ||u||_p^p + ||u - u_n||_p^p + o(1).$$

We may also apply this lemma to $\{Du\}$ in L^p and $\{u_n\}$ in L^q. In particular, we see that

$$1 = ||u_n||_q^q = ||u||_q^q + ||u - u_n||_q^q + o(1).$$

Now, let $\lambda = ||u||_q^q$ and $1 - \lambda = \lim_{n\to\infty} ||u - u_n||_q^q$. This makes u a candidate for $J(\lambda)$ and $u - u_n$ a candidate for $J(1-\lambda)$. We now combine the equalities from the Brezis-Lieb lemma applied to $\{u_n\}$ and $\{Du_n\}$ in L^p to obtain

$$I[u_n] = ||u_n||_p^p + ||Du_n||^p = ||u||_p^p + ||Du||^p + ||u - u_n||_p^p + ||Du - Du_n||_p^p + o(1),$$

and we note that the right-hand side is $I[u] + I[u - u_n] + o(1)$. Now, u and $u - u_n$ are candidates for two different minimization problems, $J(\lambda)$ and $J(1-\lambda)$, respectively. So, $I[u] \geq J(\lambda)$. $J(\lambda)$ is continuous in λ as shown earlier, as $J(\lambda) = \lambda^{p/q} J(1)$. Thus, as $||u - u_n||_q^q \to 1 - \lambda$, we may pass this limit through J. Finally, note that, by definition of $\{u_n\}$, $I[u_n] \to J(1)$. Thus, we see that $J(1) \geq J(\lambda) + J(1-\lambda)$.

Step 5: First, we observe, by the lower semicontinuity of $||\cdot||_q$, that $||u||_q \leq \liminf_{n\to\infty} ||u_n||_q$. Thus, if $\lambda = 1$, then $||u||_q^q = 1$, which implies our minimizer is in the constraint set and indeed attained. So, we want to show $\lambda = 1$. First, we show $\lambda = 0$ or $\lambda = 1$. By Step 1, we have

$$1 \geq \lambda^{p/q} + (1-\lambda)^{p/q}.$$

To show this occurs for $\lambda = 0$ or $\lambda = 1$, we let $f(\lambda) = \lambda^{p/q} + (1-\lambda)^{p/q}$. Taking the derivative, we see that

$$f'(\lambda) = \frac{p}{q}\lambda^{p/q-1} - \frac{p}{q}(1-\lambda)^{p/q-1}.$$

If $f'(\lambda) = 0$ (for $\lambda \in [0,1]$), we must have that $\lambda = 1/2$. Taking the second derivative, we find that $f''(1/2) < 0$, so $\lambda = 1/2$ is in fact a maximum of $f(\lambda)$ on

$[0, 1]$. The other extrema must therefore occur at $\lambda = 0$ and $\lambda = 1$, for which we have $f(0) = f(1) = 1$.

Now, for the sake of contradiction, assume $\lambda = 0$. This implies $||u||_q = 0$, and so $u = 0$. We know $u_k \rightharpoonup u$ in $W^{1,p}(\mathbb{R}^d)$, which implies $u_k \to u$ in $L^p_{\text{loc}}(\mathbb{R}^d)$ by the Rellich-Kondrachov compactness theorem. As $u = 0$, the condition in Lemma 3.6 is met, and thus the lemma implies that $u_k \to 0$ in $L^q(\mathbb{R}^d)$, as $p < q < p^*$. However, this is a contradiction as $||u_k||_q = 1$ for all k, by construction of the problem. Therefore, $\lambda > 0$, and so $\lambda = 1$, which shows the minimizer is attained. Before we conclude, let us describe the PDE this minimizer solves.

Step 6: We now present the PDE the minimizer solves. Let u be the minimizer for $J(\lambda)$ and choose some $\phi \in C_c^\infty(\mathbb{R}^d)$. Now, define a function

$$v(t) = \frac{u + t\phi}{||u + t\phi||_q},$$

for $t \in \mathbb{R}$. The motivation for this is to parameterize our functional I around the minimizer. We now define $i(t) = I[v(t)]$. As $v(0) = u$, the minimizer for $J(\lambda)$, $i'(0) = 0$. Writing out this equation will give us the PDE that u solves. We observe

$$i(t) = ||u + t\phi||_q^{-p} \left(\int_{\mathbb{R}^d} |Du + tD\phi|^p + |u + t\phi|^p \right).$$

Let $f(t) = ||u + t\phi||_q^{-p}$ and $g(t)$ be the quantity in parentheses. Note that

$$f'(t) = \frac{-p}{q} \left(\int_{\mathbb{R}^d} |u + t\phi|^q \right)^{-p/q-1} \left(\int_{\mathbb{R}^d} q|u + t\phi|^{q-2}(u + t\phi)(\phi) \right),$$

$$g'(t) = \int_{\mathbb{R}^d} \left[p|Du + tD\phi|^{p-2}(Du + tD\phi)(D\phi) + p|u + t\phi|^{p-2}(u + t\phi)(\phi) \right].$$

Applying the product rule and expanding the equation $i'(0) = 0$, we find

$$||u||_q^{-q} \left(\int_{\mathbb{R}^d} |u|^{q-2} u\phi \right) \left(\int_{\mathbb{R}^d} |Du|^p + |u|^p \right) = \int_{\mathbb{R}^d} |Du|^{p-2} Du \cdot D\phi + |u|^{p-2} u\phi.$$

This is the weak formulation of the PDE u solves. The PDE itself is

$$\frac{I[u]}{\lambda} u^{q-1} = -\Delta_p u + u^{p-1}.$$

□

We now apply these concentration compactness techniques to another minimization problem. This problem is concerned with attaining the optimal constant for the Sobolev inequality.

Theorem 3.8. *The following minimum is attained:*

$$\min \left\{ \frac{||Du||_{L^p(\mathbb{R}^d)}}{||u||_{L^{p^*}(\mathbb{R}^d)}} \mid u \in W^{1,p}(\mathbb{R}^d) \right\}.$$

Proof. We note that this minimization problem is the same as showing

$$\min \left\{ \int_{\mathbb{R}^d} |Du|^p \mid \int_{\mathbb{R}^d} |u|^{p^*} = \lambda \right\}$$

is attained for each $\lambda \geq 0$. Denote our functional by $I[\cdot]$ and our constraint set by E_λ. The proof of this statement follows the same program as Theorem 3.7. Letting

C be the optimal constant for the Sobolev Inequality, the PDE the minimizer solves is

$$-\Delta_p u = C^p u^{p^*-1}.$$

□

4. Compensated Compactness

4.1. **Convexity and Lower Semicontinuity.**

In dealing with minimization problems, we often encounter a functional of the form

$$I[w] = \int_U F(Dw)\,dx,$$

and want to minimize it over, say, the set $S = \{u \in W^{1,q}(U) \mid w = g \text{ on } \partial U\}$, where $F\colon \mathbb{R}^d \to \mathbb{R}$ is smooth. Typically, we have a quadratic growth assumption on F, so there exists a C such that, for all p,

$$|F(p)| \leq C(1 + |p|^2).$$

This is a reasonable assumption for minimization over $H^1(U) = W^{1,2}(U)$ because then, for all $u \in H^1(U)$, $\int_U F(Du) < \infty$. We observe the theorems and ideas present in [2]. We will discuss a theorem relating convexity and lower semicontinuity, but first let's imagine we have a minimizer u to the problem above and that u is smooth. We will see if we can obtain information about F. Let $v \in C_c^0(U)$ be Lipschitz. Then, let's define $i(t) = I[u + tv]$. Because u is the minimizer of I, i must have a minimum at $t = 0$. In addition, it's second derivative is therefore non-negative at $t = 0$. Thus, we see that

$$i''(0) = \int_U \sum_{i,j} F_{y_i,y_j}(Du) v_{x_i} v_{x_j} \geq 0,$$

where $1 \leq i, j \leq d$. We will abbreviate this sum as $F_{y_i,y_j}(Du) v_{x_i} v_{x_j}$. We now want to choose a specific v to learn something about F. To this end, let $\eta \in C_c^\infty$, $\xi \in \mathbb{R}^d$, and $\varepsilon \in \mathbb{R}$. Next, let $z\colon \mathbb{R} \to \mathbb{R}$ be defined so

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

and extended to be periodic with period 2. This is known as a "sawtooth" function. We now define

$$v(x) = \varepsilon \eta(x) z\left(\frac{x \cdot \xi}{\varepsilon}\right).$$

We will now use v as our test function in the equality for $i''(0)$. First, note that $v_{x_i} v_{x_j}$ equals

$$\left(\varepsilon \eta'(x) z\left(\frac{x \cdot \xi}{\varepsilon}\right) + \eta(x) z'\left(\frac{x \cdot \xi}{\varepsilon}\right)\xi_i\right)\left(\varepsilon \eta'(x) z\left(\frac{x \cdot \xi}{\varepsilon}\right) + \eta(x) z'\left(\frac{x \cdot \xi}{\varepsilon}\right)\xi_j\right).$$

As $\varepsilon \to 0$, because η and z are bounded, $\varepsilon \eta'(x) z(x \cdot \xi/\varepsilon) \to 0$. Substituting into $i''(0) \geq 0$ and letting $\varepsilon \to 0$, we are left with

$$\lim_{\varepsilon \to 0} \int_U F_{y_i,y_j}(Du) \eta^2(x) \left(z'\left(\frac{x \cdot \xi}{\varepsilon}\right)\right)^2 \xi_i \xi_j \geq 0.$$

As $\eta^2(x) \geq 0$, this implies $\langle F_{y_i,y_j}(Du(x))\xi, \xi\rangle \geq 0$. This is the definition of convexity if we can take $Du(x) = p$ for any p.

Finally, before presenting the theorem, we introduce the notion of weak lower semicontinuity.

Definition 4.1. A functional ϕ is said to be weakly lower semicontinuous if

$$\liminf_{n\to\infty} \phi(u_n) \geq \phi(u)$$

for all sequences $\{u_n\}$ such that $u_n \rightharpoonup u$.

We now present a connection between lower semicontinuity and convexity.

Theorem 4.2. *Weak lower semicontinuity of $I[\cdot]$, with respect to weak convergence in $W^{1,q}(U)$, is equivalent to the convexity of F.*

Proof. First, assume $I[\cdot]$ is weakly lower semicontinuous. Assume our domain is $Q = (0,1)^n$, and let $p \in \mathbb{R}^d$. Let $v \in C_c^\infty(Q)$ be a test function. We now split Q, for $k \in \mathbb{N}$, into 2^{kn} disjoint, identical cubes. The side length of each of these cubes is thus $1/2^k$. For each k, denote this collection by $\{Q_i\}_{i=1}^{kn}$ with centers $\{x_i\}$. Then, for $x \in Q_i$, define

$$u_k(x) = \frac{1}{2^k} v(2^k(x - x_i)) + p \cdot x.$$

The first term vanishes as $k \to \infty$, and so let $u(x) = p \cdot x$. By the boundedness of v on Q, we obtain that $u_k \to u$ in $W^{1,q}(Q)$, which gives weak convergence in $W^{1,q}(Q)$. By the definition of lower semicontinuity, $I[u] \leq \liminf_{k\to\infty} I[u_k]$. Thus,

$$L^n(Q)F(p) \leq \liminf_{k\to\infty} I[u_k] = \int_Q F(Dv + p),$$

where we have used the fact $u_k \rightharpoonup u$ in $W^{1,q}(U)$, and so $Du_k \rightharpoonup Du$ in $W^{1,q}(U)$ as well, along with 2.2. Here, $u = p \cdot x$ is a smooth minimizer and, as $\operatorname{supp}(v) \subset Q$, u and $\{u_k\}$ all have the same boundary conditions, so u is a smooth minimizer lying in the same constraint class. So, u is a smooth minimizer of the minimization problem

$$\min \left\{ \int_Q F(Dw) \mid w = u \text{ on } \partial Q \right\},$$

and so $D^2F(p)\xi \cdot \xi \geq 0$ for all p, so F is convex.

Now, assume F is convex. We want to show if $u_k \rightharpoonup u$, then $\liminf_k I[u_k] \geq I[u]$. To simplify this, we first assume we have coefficients $\{b_j\}_1^m$ and $\{c_j\}_1^m$ so

$$F(p) = \max_{1\leq j\leq m} (b_j \cdot p + c_j).$$

So, we have a finite set of affine functions of which F is the maximum of. We assume this because, if F is convex, then we can find a set of affine functions for which F is the supremum of. So, proving this theorem for a finite set is sufficient because we can then apply the Monotone Convergence Theorem to obtain the result when we have a supremum instead of a maximum.

We now define a set

$$E_j = \{x \in U \mid F(Du(x)) = b_j \cdot Du(x) + c_j\}.$$

By the definition of F, $U = \cup_j E_j$. Assuming the affine functions are distinct, $E_i \cap E_j = \emptyset$ if $i \neq j$. We now see that

$$I[u] = \int_U F(Du)\,dx = \sum_{j=1}^m \int_{E_j} b_j \cdot Du(x) + c_j\,dx$$
$$\leq \liminf_{k\to\infty} \sum_{j=1}^m \int_{E_j} F(Du_k)\,dx,$$

which follows from linearity of the integrand and the assumption on the structure of F. Then, we note that the last line is $\liminf_{k\to\infty} I[u_k]$. Thus, the functional $I[\cdot]$ is weakly lower-semicontinuous. □

In Lemma 2.1, we see that a functional that is continuous with respect to weak convergence is affine linear. Here, we see a similar idea. Functionals of the form $I[u] = \int F(Du)$ that are weakly lower semicontinuous with respect to weak convergence are convex, and convex functions are in turn the supremum of affine linear functions.

For a typical minimization problem on bounded domains, we utilize compactness from Rellich-Kondrachov to conclude that, for a minimizing sequence $\{u_n\}$, $I[u_n] \to I[u]$, so the weak limit of $\{u_n\}$ is indeed a minimizer. On unbounded domains, even though we don't have access to compactness, and thus strong convergence in L^2, we may still obtain convergence of the energies using the convexity argument in Theorem 4.2, and so we see convexity compensating for the lack of compactness.

We next present the following theorem about the convergence of the energies and its relationship to weak convergence. Before doing so, we define an important term.

Definition 4.3. F is said to be uniformly strictly convex if there exists a $\gamma > 0$ such that

$$\xi^T D^2 F(p) \xi \geq \gamma |p|^2$$

for all $p, \xi \in \mathbb{R}^d$.

Theorem 4.4. *With assumptions of quadratic growth and uniform strict convexity of F, if $u_k \rightharpoonup u$ in $W^{1,2}$ and $I[u_k] \to I[u]$, then $u_k \to u$ in $W^{1,2}(U)$.*

Proof. Applying the Taylor Series expansion on F (as we know F is smooth) and the uniform convexity assumption, given $x, y \in \mathbb{R}^d$,

$$F(y) \geq F(x) + DF(x) \cdot (y - x) + \frac{\gamma}{2}|y - x|^2.$$

Applying this to Du_k and Du, we see that

$$I[u_k] \geq I[u] + \int_U DF(Du)(Du - Du_k) + \frac{\gamma}{2}\int_U |Du - Du_k|^2.$$

We observe that $DF(Du) \in L^2(U)$ by the assumptions on F. Indeed, $|DF(Du)| \leq C(1 + |Du|)$. We have $u_k \rightharpoonup u$ in $W^{1,2}(U)$ and $I[u_k] \to I[u]$, and thus

$$\frac{\gamma}{2}\int_U |Du - Du_k|^2 \leq I[u_k] - I[u] + \int_U DF(Du)(Du - Du_k) \to 0.$$

Thus, $Du_k \to Du$ in $L^2(U)$. The Sobolev inequality then gives that $u_k \to u$ in $L^{2^*}(U)$, and as we are on a bounded domain, we have $u_k \to u$ in $L^2(U)$. Combining these, we see that $u_k \to u$ in $W^{1,2}(U)$. □

Remark 4.5. We note that this is an example of convexity providing strong compactness. Indeed, we proved that quadratic growth and uniform strict convexity alone imply $\{u_k\}$ is bounded in $H^1(U) = W^{1,2}(U)$. We then obtain a subsequence (still denoted $\{u_n\}$) that converges weakly in $H^1(U)$. Finally, convexity of F ensures $I[\cdot]$ is lower semicontinuous and so this ensures $I[u_n] \to I[u]$.

4.2. **Div-Curl Lemma.** We now present a theorem dealing with the product of weakly convergent sequences. First, we discuss the need for such a theorem. In section 2, we showed that $\sin(nx) \rightharpoonup 0$ but $\sin^2(nx) \rightharpoonup 1/2$. Thus, the problem of the product of such sequences is nontrivial. In general, we cannot obtain information about such a product, but when we have additional information on the structure of certain derivatives of each sequence, we can. We present the following theorem, taking inspiration from [5].

Theorem 4.6. *Let $\Omega \subset \mathbb{R}^d$ be an open, bounded set. $\{v_n\}, \{\operatorname{div}(v_n)\}, \{w_n\}, \{\operatorname{curl}(w_n)\}$ are bounded in $L^2(\Omega)$, and $v_n \rightharpoonup v$ and $w_n \rightharpoonup w$ in $L^2(\Omega)$. Then, $v_n \cdot w_n \rightharpoonup v \cdot w$ in $L^2(\Omega)$.*

Remark 4.7. We define $\operatorname{div}(v_n) = \sum_i \partial_i v^i$, where v^i denotes the i-th component of v. The divergence of a function is a vector. We also define $\operatorname{curl}(w_n)_{i,j} = \partial_j w^i - \partial_i w^j$. The curl of a function is a matrix.

Remark 4.8. For the case where the functions w_n are in fact potentials on a bounded domain, we can prove the Div-Curl Lemma directly. So, assume $w_n = Dz_n$ for some sequence of functions $\{z_n\}$ with the same assumptions on the sequences $\{v_n\}$ and $\{w_n\}$. In addition, assume $z_n \to z$ in $L^2(U)$ and $Dz_n \rightharpoonup Dz$ in $L^2(U)$. Then, we note that

$$\int_U v_n \cdot w_n \phi = \int_U v_n \cdot Dz_n \phi = -\int_U Dv_n \cdot z_n \phi - \int_U v_n \cdot D\phi z_n,$$

where we have applied integration by parts. We now have a z_n coefficient in each integral. We know $z_n \to z$ in $L^2(U)$ and, additionally, $v_n \rightharpoonup v$ and $Dv_n \rightharpoonup Dv$ in $L^2(U)$. As the product of weak and strong convergence is strong, we see that

$$\int_U v_n \cdot w_n \phi \to -\int_U Dv \cdot z\phi - \int_U v \cdot D\phi z = \int_U v \cdot w\phi.$$

This is a simpler proof of the Div-Curl Lemma for this special case. We now present the proof of the general Div-Curl Lemma.

Proof. We want to show that

$$\int_\Omega v_n \cdot w_n \phi \to \int_\Omega v \cdot w\phi$$

for all $\phi \in C_c^\infty(\Omega)$. To show this, fix such a ϕ and choose $\psi \in C_c^\infty$ such that $\psi \equiv 1$ on $\operatorname{supp}(\phi)$. Define

$$\tilde{v_n} = \begin{cases} \phi v_n & \text{in } \Omega \\ 0 & \text{in } \mathbb{R}^d \setminus \Omega \end{cases}, \quad \text{and} \quad \tilde{w_n} = \begin{cases} \psi w_n & \text{in } \Omega \\ 0 & \text{in } \mathbb{R}^d \setminus \Omega \end{cases}.$$

Applying Plancherel, $\int_{\mathbb{R}^d} \widehat{\tilde{v}}_n \cdot \overline{\widehat{\tilde{w}}_n} = \int_\Omega v_n \cdot w_n \phi$. We now want to make sure the Fourier Transform of $\tilde{v_n}$ and $\tilde{w_n}$ preserves the properties of v_n and w_n. As

$v_n \rightharpoonup v$, we may directly compute that $\widehat{\tilde{v}}_n \rightharpoonup \widehat{\tilde{v}}$ in $L^2(\mathbb{R}^d)$, as does $\{w_n\}$. Applying Plancherel, $\{\widehat{\tilde{v_n}}\}$ and $\{\widehat{\tilde{w_n}}\}$ are bounded in $L^2(\mathbb{R}^d)$. Similarly, $\sum_i \widehat{\tilde{v}}_n^i(\xi)$ and $\mathrm{curl}(\widehat{\tilde{w}}_n)$ are bounded in $L^2(\mathbb{R}^d)$.

Now, let's look at the convergence of $\int_{\mathbb{R}^d} \widehat{\tilde{v}}_n \cdot \overline{\widehat{\tilde{w}}_n}$. We consider the convergence over B_R and then over $\mathbb{R} \setminus B_R$. In this case, B_R is considered as the set of low frequencies and $\mathbb{R} \setminus B_R$ as the set of high frequencies.

For the low frequencies, first note that

$$|\widehat{\tilde{v}}_n(\xi)| = \left| \int_{\mathbb{R}^d} \tilde{v}_n(t) e^{-2\pi i t \cdot \xi}\, dt \right| \leq \int_\Omega |\phi||v_n| \leq \sup_\Omega |\phi||\Omega|^{1/2}||v_n||_{L^2(\Omega)},$$

where $|\Omega|$ denotes the Lebesgue measure of Ω. So, $||\widehat{\tilde{v}}_n||_\infty \leq C_\Omega ||v_n||_{L^2(\Omega)}$. In addition, we see that

$$\hat{\tilde{v}}_n(\xi) = \int_{\mathbb{R}^d} \tilde{v}_n e^{-2\pi i t \cdot \xi}\, dt = \int_{\mathbb{R}^d} \eta v_n e^{-2\pi i t \cdot \xi}\, dt \to \int_{\mathbb{R}^d} \eta v e^{-2\pi i t \cdot \xi}\, dt = \widehat{\phi v},$$

as $\phi e^{-2\pi i t \cdot \xi} \in L^2(\mathbb{R}^d)$ because ϕ is supported in Ω. So, $\widehat{\tilde{v}}_n \to \widehat{\tilde{v}}$ pointwise. We may apply the same argument to $\widehat{\tilde{w}}_n$. Then, using the Lebesgue Dominated Convergence Theorem, we see that

$$\int_{B_R} \widehat{\tilde{v}}_n \widehat{\tilde{w}}_n - \widehat{\tilde{v}}\widehat{\tilde{w}} = \int_{B_R} \widehat{\tilde{w}}_n(\widehat{\tilde{v}}_n - \widehat{\tilde{v}}) + \int_{B_R} \widehat{\tilde{v}}(\widehat{\tilde{w}}_n - \widehat{\tilde{w}}) \to 0.$$

This gives convergence of the low frequencies.

To tackle the high frequencies, we make use of the following lemma:

Lemma 4.9. *If $V, W \in \mathbb{C}^d$, then for all $1 \leq j \leq d$,*

$$\xi_j \sum_i V_i \overline{W_i} = \overline{W_j} \sum_i \xi_i V_i + \sum_i [\xi_j \overline{W_i} - \xi_i \overline{W_j}] V_i.$$

Let $V = \widehat{\tilde{v}}_n$ and $W = \widehat{\tilde{w}}_n$. We have

$$|\xi_j| \widehat{\tilde{v}}_n \cdot \widehat{\tilde{w}}_n \leq |\widehat{\tilde{w}}_n||\mathrm{div}(\widehat{\tilde{v}}_n)| + |\mathrm{curl}(\widehat{\tilde{w}}_n)||\widehat{\tilde{v}}_n|.$$

Integrating both sides and applying Holder's inequality, we find that the LHS is bounded in $L^1(\mathbb{R}^d)$. Here, we have used the uniform boundedness of $\widehat{\tilde{v}}_n$, $\mathrm{div}(\widehat{\tilde{v}}_n)$, $\widehat{\tilde{w}}_n$, and $\mathrm{curl}(\widehat{\tilde{w}}_n)$. Summing over j, we see that $|\xi|\widehat{\tilde{v}}_n \cdot \widehat{\tilde{w}}_n \in L^1(\mathbb{R}^d)$, where we have noted that $|\xi| \leq \sum_i |\xi_i|$. From here, we see that

$$\left| \int_{\mathbb{R}^d \setminus B_R} \widehat{\tilde{v}}_n \cdot \widehat{\tilde{w}}_n \right| \leq \frac{1}{R} \int_{\mathbb{R}^d \setminus B_R} |\xi| \widehat{\tilde{v}}_n \cdot \widehat{\tilde{w}}_n.$$

Thus, the contribution off of a ball can be made arbitrarily small. This is why we split into low and high frequencies: We can control the contribution off of a ball of large radius, and so if we have convergence on B_R for all R, then we desired convergence on $\mathbb{R}^d$ is achieved. This completes the proof. □

4.3. **Homogenization.** We now introduce homogenization following [2]. To begin, we consider the PDE

$$\begin{cases} -\mathrm{div}(A(x/\varepsilon) D u^\varepsilon) = f & \text{in } U \\ u^\varepsilon = 0 & \text{on } \partial U \end{cases}.$$

If A is periodic, then $A(x/\varepsilon)$ has rapid oscillations that repeat by periodicity. Investigating the system at the ε-level allows us to analyze the system at a specific, microscopic level. However, directly computing characteristics of this system, such as its solutions, is much too computationally expensive because of the rapid oscillations. Instead, we approximate the system for small ε by sending $\varepsilon \to 0$. We can envision this as dividing the domain into cells, each of which is on the ε scale, where the system is the same across each cell. Instead of directly describing this system, we instead investigate its average behavior. This is what homogenization is: looking for the "average" on an incredibly small scale. This naturally leads to two questions:

(1) What is the limit u of $\{u^\varepsilon\}_{\varepsilon>0}$ as $\varepsilon \to 0$? What equation does it solve? Surprisingly, we find that the homogenized matrix for the PDE u solves is not A.
(2) In what sense does $u_\varepsilon \to u$ as $\varepsilon \to 0$?

As a note, we may sometimes also use short-hand notation to indicate the first condition of the PDE as $-(a_{ij}(x/\varepsilon)u^\varepsilon_{x_i})_{x_j}$. The convenience of this will become clear soon. We assume a few things about this PDE. We assume an L^∞ condition on A, so $|A(y)| \leq C$ and that A is periodic on on the unit cube $Q \subset \mathbb{R}^d$. In addition, we assume A is uniformly elliptic, so $\xi^T A(y)\xi \geq \nu|\xi|^2$ for $\nu > 0$. Now, assume we have weak solutions $\{u^\varepsilon\}$ to this PDE. Using u^ε as our test function, we see that u^ε is uniformly bounded in $W^{1,2}_0(U)$.

This uniform boundedness furnishes us with a subsequence, still denoted $\{u^\varepsilon\}$, that converges weakly to some $u \in W^{1,2}_0(U)$. What PDE does u solve? Before we discuss this, we need a lemma.

Lemma 4.10 (Fredholm Alternative)**.** *Let $H^1_\#(Q)$ be the set of functions periodic on the unit cube $Q \subset \mathbb{R}^d$, and let $V = \{u \in H^1_\#(Q) \mid \int_Q u = 0\}$. Assume $f \in L^2(Q)$. Then, there exists a unique $w \in V$ such that, for all $\phi \in V$,*

$$\int_Q ADw \cdot D\phi = \int_Q f\phi.$$

Remark 4.11. The proof follows using the Riesz-Representation Theorem.

First, we consider the following system of PDEs:

$$\begin{cases} -(a_{ij}(y)w^l_{y_i})_{y_j} = (a_{il}(y))_{y_i} & \text{in } \mathbb{R}^d \\ w^l & \text{is } Y\text{-periodic} \end{cases},$$

where $1 \leq l \leq n$. The first condition is shorthand notation for the equation $-\text{div}(ADw^l) = \text{div}(A_l)$, where A_l is the l-th column of A. These PDEs are known as **corrector problems**. The Fredholm alternative provides us with a solution w^l, for each l, to this PDE, as we note that $\int_Y (a_{il}(y))_{y_i} = 0$ by the periodicity of A. Using these functions w^l, we consider the coefficients

$$\tilde{a}_{il} = \int_Y a_{ij}(y)(\delta_{jl} + w^l_{y_j}(y)),$$

and the corresponding PDE:

$$\begin{cases} -(\tilde{a}_{il}u_{x_i})_{x_l} = f & \text{in } U \\ u = 0 & \text{on } \partial U \end{cases}.$$

We now prove the following theorem:

Theorem 4.12. *u solves the PDE above.*

Proof. By our boundedness conditions, for each j, $a_{ij}(x/\varepsilon)u^\varepsilon_{x_i}$ converges weakly to, say, $\xi^j : \mathbb{R}^d \to \mathbb{R}$ (where we have denoted our subsequence as ε). We now see that

$$\int_U fv = \int_U a_{ij}(x/\varepsilon)u^\varepsilon_{x_i} v_{x_j} \implies \int_U fv = \int_U \xi \cdot Dv,$$

by taking limits. We now define a set of functions, known as **correctors**, for a fixed l, by $v^\varepsilon(x) = x_l + \varepsilon w^l(x/\varepsilon)$. We now note that

$$(a_{ij}(x/\varepsilon)v^\varepsilon_{x_j})_{x_i} = (a_{ij}(x/\varepsilon)\delta_{jl})_{x_i} + (a_{ij}(x/\varepsilon)(w^l(x/\varepsilon))_{x_j})_{x_i} = 0.$$

So, v^ε solves $(a_{ij}(x/\varepsilon)v^\varepsilon_{x_j})_{x_i} = 0$, weakly. Let $\eta \in C^\infty_c(U)$ be a test function and define $v = \eta v^\varepsilon$. We see that

$$\begin{aligned}\int_U f\eta v^\varepsilon &= \int_U a_{ij}(x/\varepsilon)u^\varepsilon_{x_i}(\eta v^\varepsilon)x_j \\ &= \int_U a_{ij}(x/\varepsilon)u^\varepsilon_{x_i}\eta_{x_j}v^\varepsilon + a_{ij}(x/\varepsilon)u^\varepsilon_{x_i}\eta v^\varepsilon_{x_j} \\ &= \int_U a_{ij}(x/\varepsilon)u^\varepsilon_{x_i}\eta_{x_j}v^\varepsilon - u^\varepsilon(a_{ij}(x/\varepsilon)\eta v^\varepsilon_{x_j})_{x_i} \\ &= \int_U a_{ij}(x/\varepsilon)u^\varepsilon_{x_i}\eta_{x_j}v^\varepsilon - u^\varepsilon a_{ij}(x/\varepsilon)\eta_{x_i}v^\varepsilon_{x_j}.\end{aligned}$$

We have boundedness of the w^l in L^2, which gives that $v^\varepsilon \to x_l$ in $L^2(U)$. Noting that rescalings converge to the average, we see that

$$a_{ij}(x/\varepsilon)v^\varepsilon_{x_j} = a_{ij}(x/\varepsilon)(\delta_{jl} + w^l(x/\varepsilon)_{x_j}) \rightharpoonup \tilde{a}_{il},$$

in $L^2(U)$. Taking the limit as $\varepsilon \to 0$ on each side of

$$\int_U f\eta v^\varepsilon = \int_U a_{ij}(x/\varepsilon)u^\varepsilon_{x_i}\eta_{x_j}v^\varepsilon - u^\varepsilon a_{ij}(x/\varepsilon)\eta_{x_i}v^\varepsilon_{x_j}$$

and noting the convergence results we have already proved, we obtain

$$\int_U f\eta x_l = \int_U \xi^j x_l \eta_{x_j} - u\tilde{a}_{il}\eta_{x_i},$$

where we have noted that the product of strong and weak convergence is strong. Indeed, we know that $u^\varepsilon \to u$ strongly in L^2 by the Rellich compactness theorem. We have an integral equality involving ξ. Indeed,

$$\int_U f\eta x_l = \int_U \xi^l\eta + \xi^j\eta_{x_j}x_l,$$

where we have again used summation notation. Importantly, the derivative of x_l picks out the l-th component. Noting that $\tilde{a}_{il}$ is constant, we apply integration by parts to see that

$$\int_U u_{x_i}\tilde{a}_{il}\eta = \int_U \xi^l\eta.$$

Now, we're almost done! Noting the integral equality that ξ satisfies, from the beginning, we have

$$\int_U fv = \int_U \xi \cdot Dv = \int_U \tilde{A}Du \cdot Dv.$$

This completes the proof. □

We now present a non-trivial application of the Div-Curl Lemma to the sequence $\{u^\varepsilon\}$.

Theorem 4.13. *Given $\{u^\varepsilon\}$ as defined in Theorem 4.12, the energies $A(x/\varepsilon)Du^\varepsilon \cdot Du^\varepsilon \rightharpoonup A^* Du \cdot Du$.*

Proof. We know $Du^\varepsilon \rightharpoonup Du$. Note that $\text{curl}(Du^\varepsilon) = 0$. Next, we know $A(x/\varepsilon)Du^\varepsilon \rightharpoonup A^* Du$, and $\text{div}(A(x/\varepsilon)Du^\varepsilon) = f \in L^2(U)$. We now have two weakly converging sequences, one with vanishing curl and one with constant, bounded divergence. Thus, we can apply the Div-Curl Lemma to say that the product of these sequences converges weakly to the product of weak limits. So, $A(x/\varepsilon)Du^\varepsilon \cdot Du^\varepsilon \rightharpoonup A^* Du \cdot Du$. □

5. Acknowledgements

I would like to thank David Bowman for being a phenomenal mentor and guiding me through this REU. He gave me many interesting, difficult problems but also spent a significant amount of time discussing these problems and the topics behind them with me, for which I cannot thank him enough. I would also like to thank Peter May for making this REU possible.

References

[1] Wikipedia contributors, *Compensated compactness*. https://en.wikipedia.org/wiki/Brezis-Lieb_lemma.

[2] Lawrence C. Evans, *Weak convergence methods for nonlinear partial differential equations*, American Mathematical Society, Providence, Rhode Island, 1990.

[3] ———, *Partial differential equations*, American Mathematical Society, Providence, Rhode Island, 2010.

[4] Mathieu Lewin, *Describing lack of compactness in Sobolev spaces*, Université de Cergy-Pontoise, France, 2010. Lecture - Taken from unpublished lecture notes "Variational Methods in Quantum Mechanics" written for a course delivered at the University of Cergy-Pontoise in 2010.

[5] Christophe Prange, *Compensated compactness*. https://prange.perso.math.cnrs.fr/documents/coursEDMI2016_lecture3.pdf.