Semilinear Elliptic Partial Differential Equations with the Critical Sobolev Exponent

Melusi Manqoba Mavuso
Supervisor: Dr Francois Ebobisse-Bille

A dissertation presented for the degree of
Master of Science

Department of Mathematics and Applied Mathematics
Faculty of Science, University of Cape Town
South Africa

February 6, 2017

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science at the University of the Cape Town. It has not been submitted before for any degree or examination in any other University.

Signed by candidate Signature Removed

February 6, 2017

Abstract

We present how variational methods and results from linear and non-linear functional analysis are applied to solving certain types of semilinear elliptic partial differential equations (PDEs). The ultimate goal is to prove results on the existence and non-existence of solutions to the Semilinear Elliptic PDEs with the Critical Sobolev Exponent. To this end, we first recall some useful results from functional analysis, including the Sobolev spaces, which provide a natural setting for the idea of weak or generalised solutions. We then present linear PDE theory, including eigenvalues of the Dirichlet Laplacian operator. We discuss the Direct Methods of Calculus of Variations and Critical Point Theory, together with examples of how these techniques are applied to solving PDEs. We show how the existence of solutions to semilinear elliptic equations depends on the exponent of the growth of the non-linear term. This then naturally leads to the discussion of the critical Sobolev exponent, where we present both positive and negative results.

Acknowledgements

First, I would like to acknowledge the help, guidance and patience from my supervisor Dr. Francois Ebobisse-Bille.

I would also like to thank Mr Armin van Buuren for keeping me in 'a state of trance' during the long nights of typing this dissertation.

Contents

Chapter 1

Introduction and Background

The study of Partial Differential Equations (PDEs) has a long and rich history that dates back to the works of Euler, d'Alembert, Langrange and Laplace in the eighteenth century. Nowadays PDEs are ubiquitous in many branches of science, due to their success in describing the dynamics of many natural processes. As a result, finding and analysing properties of solutions to PDEs have become central to research in the field of mathematical analysis. A detailed history and survey of some techniques for solving PDEs is given by [Brezis and Browder, 1998].

Ideally, when presented with a PDE, one would like to find an analytical solution to the equation, usually by employing well-established standard techniques for solving PDEs. Some of these techniques include the method of characteristics and separation of variables, and they provide methods to explicitly construct a sufficiently smooth solution to the PDE — called a *classical solution*. Unfortunately, though, these techniques only work for a few elementary PDEs. For the vast majority of PDEs, usually the best that can be done is to prove, often in a non-constructive way, that a solution exists, rather than explicitly construct one. The usefulness of some of these existence techniques is that they are often accompanied by stable numerical algorithms that approximate the solution to arbitrary precision. One such numerical approximation scheme is the celebrated *Finite Element Method* (see [Ciarlet, 2002]).

This shift towards existence results has opened the door to methods from linear and non-linear functional analysis to be applied to the field. The motivation for the application of such methods stems from the observation that, from a high-level point of view, solving a PDE can be viewed as solving an equation of the form

$$Tu = v,$$

in u, for some operator $T : E \to F$ between two function spaces E and F, where $v \in F$ is fixed. The identification of the right spaces E and F, and right operator T, is one of the greatest success of modern PDE theory, and has lead researchers to consider what is known as *generalised or weak solutions* to PDEs.

The aim of this dissertation is twofold. First, we want to present this alternative view of solving PDEs by specifying the right spaces E and F, the right operator T, and establish the results from functional analysis that are applicable to solving PDEs. Secondly, we want to show how these ideas can be applied to solving specific types of linear and semilinear elliptic PDEs. The main idea is to build up this theory in the first few chapters and then apply it in solving the notorious semilinear elliptic PDEs with the so-called *critical Sobolev exponent* in later chapters.

Enormous research has been done in this direction and we aim to present the main results from a different point of view (some with different proofs and approaches), and in a manner that is accessible to beginning graduate students. Our work was heavily influenced by, and will at times mimic (in principle), the works of [Brezis, 2010], [Brézis and Nirenberg, 1983],[Evans, 1998], [Badiale and Serra, 2010] and [Struwe, 1990]. We begin with a brief historical motivation.

1.1 Dirichlet's Principle

Let $\Omega \subseteq \mathbb{R}^N$ be an open and connected subset (a domain) that is also bounded (finite diameter) with a smooth boundary $\partial\Omega$ and $g : \partial\Omega \to \mathbb{R}$ be continuous. We consider *Laplace's equation*:

$$\begin{cases} -\Delta u = 0 & \text{in } \Omega \\ u = g & \text{on } \partial\Omega. \end{cases} \tag{1.1}$$

Here $\Delta u = \sum_{i=1}^{N} \frac{\partial^2 u}{\partial x_i^2}$ is the *Laplacian* of u. A *classical solution* of (1.1) is a function $u \in C^2(\Omega) \cap C(\bar{\Omega})$ that satisfies (1.1) for every $x \in \overline{\Omega}$.

Solutions to this equation are called *harmonic functions*, and methods of finding classical solutions to (1.1) are well known and discussed in many standard textbooks (see [Evans, 1998], for instance). We take a different point of view and consider an indirect way of characterising solutions to (1.1), by viewing them as *critical points* of a certain functional. First let $u \in C^2(\Omega) \cap C(\bar{\Omega})$ be a classical solution of (1.1) and $v \in C^2(\Omega) \cap C(\bar{\Omega})$ with $v = 0$ on $\partial\Omega$. Multiplying (1.1) by v and integrating by parts gives

$$\int_\Omega \nabla u \cdot \nabla v \, dx = 0, \tag{1.2}$$

where $\nabla w = \left(\frac{\partial w}{\partial x_1}, \dots, \frac{\partial w}{\partial x_n} \right)$ is the gradient of w. Let $\mathcal{A} := \left\{ w \in C^2(\Omega) \cap C(\bar{\Omega}) : w = g \text{ on } \partial\Omega \right\}$ and define the functional $I : \mathcal{A} \to \mathbb{R}$ by

$$I(w) := \frac{1}{2} \int_\Omega |\nabla w|^2 \, dx, \;\; w \in \mathcal{A}.$$

Note that I is actually well-defined even for $w \in C^1(\Omega)$ — more on this later in the chapter. The *first variation* of I at u in the direction v is the directional derivative

$$\delta I(u; v) := \lim_{t \to 0} \frac{I(u + tv) - I(u)}{t} = \int_\Omega \nabla u \cdot \nabla v \, dx.$$

Thus, (1.2) says that $\delta I(u; v) = 0$ for any $v \in C^2(\Omega) \cap C(\bar{\Omega})$ with $v = 0$ on $\partial\Omega$; that is, $\delta I(u; \cdot) = 0$, which prompts us to think of u as a *critical point* of I (we cannot as yet make this link since $\mathcal{A}$ is not a vector space when $g \neq 0$). For any $w \in \mathcal{A}$ we have,

$$I(w) = \frac{1}{2} \int_\Omega |\nabla w|^2 \, dx = \frac{1}{2} \int_\Omega |\nabla(u + w - u)|^2 \, dx$$

$$= \frac{1}{2} \int_\Omega |\nabla u|^2 \, dx + \int_\Omega \nabla u \cdot \nabla(w - u) \, dx + \frac{1}{2} \int_\Omega |\nabla(w - u)|^2 \, dx \geq \frac{1}{2} \int_\Omega |\nabla u|^2 \, dx = I(u)$$

since $w - u = 0$ on $\partial\Omega$. Thus u minimizes I over all functions w in $\mathcal{A}$. So we have just shown that every classical solution u of (1.1) minimizes I, and the first variation of I vanishes at u.

Conversely, let $u \in \mathcal{A}$ be a minimizer of I (or even just a point where the first variation vanishes), and $v \in C^2(\Omega) \cap C(\bar{\Omega})$ with $v = 0$ on $\partial\Omega$. Define $h : \mathbb{R} \to \mathbb{R}$ by

$$h(t) = I(u + tv),\ t \in \mathbb{R}.\ (h \text{ is well-defined since } u + tv \in \mathcal{A})$$

The function h has a global minimum at $t = 0$, thus

$$0 = h'(0) = \delta I(u; v) = \int_\Omega \nabla u \cdot \nabla v\, dx = \int_\Omega (-\Delta u) v\, dx.$$

Since this is true for any such v, we can conclude, from the *Fundamental Lemma of Calculus of Variations*, that $-\Delta u = 0$ in Ω. Thus, we have (heuristically) shown:

Theorem 1.1.1 (Dirichlet's Principle)**.** *A function $u \in \mathcal{A}$ is a solution to (1.1) if and only if u minimizes I on $\mathcal{A}$.*

This result establishes the equivalence between two problems in mathematics: solving PDEs and minimizing functionals (or, as we shall see later, finding critical points). This is analogous to single-variable calculus, where we often solve minimization problems by solving algebraic equations (setting the derivative equal to 0) and vice versa. The difference here is that we have replaced the notion of a function of a real variable with a functional defined on an appropriate function space, and the notion of an algebraic equation is replaced by a differential equation. However, not all PDEs have an equivalent minimization problem. The PDEs whose solutions correspond to minimizers (or more generally, critical points) of certain functionals are said to be of *variational form*, and we limit our study exclusively to such equations.

Dirichlet's principle was established by, among others, Riemann in 1851 (see [Riemann, 1851] and [Riemann, 1857]) and he gave it its name. As the story goes, Riemann further claimed — without rigorous justification — that a unique minimizer of the functional I, and consequently a solution to (1.1), exists (note that we only proved the equivalence of a minimizer and a solution to Laplace's equation without claiming the existence of either). However, even though a minimizer does indeed exist in this example, Weierstrass later showed that in general, such a fact cannot be taken for granted, by giving an example of a functional that is bounded below but does not have a minimizer. Here is a slight modification of his counterexample.

Theorem 1.1.2. *Let $\mathcal{A} := \{u \in C^1([-1, 1]) : u(-1) = -1, u(1) = 1\}$ and define $J : \mathcal{A} \to \mathbb{R}$ by*

$$J(u) := \int_{-1}^{1} \left(x \frac{du}{dx} \right)^2 dx,\ u \in \mathcal{A}.$$

Then $\inf_{u \in \mathcal{A}} J(u) = 0$ is not attained.

One can prove this result by observing that the sequence $(u_n) \subseteq \mathcal{A}$, where $u_n(x) = \frac{\arctan nx}{\arctan n}$ for $x \in [-1, 1]$, is such that $\lim_{n\to\infty} J(u_n) = 0$, while there clearly is no $u \in \mathcal{A}$ such that $J(u) = 0$ (otherwise such u would be constant on $[-1, 1]$, contradicting the boundary conditions). This theorem begs the question: which minimization problems have solutions?

To put this discussion into a broader perspective, again let $\Omega \subseteq \mathbb{R}^N$ be a bounded domain with smooth boundary $\partial\Omega$, $f : \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ be a C^1 function and $g : \partial\Omega \to \mathbb{R}$ be continuous. For non-trivial results, f is usually assumed to satisfy some growth and regularity conditions, but we will delay the discussion of these conditions. Let $\mathcal{A} := \{w \in C^1(\Omega) \cap C(\bar{\Omega}) : w = g \text{ on } \partial\Omega\}$ and define the *energy functional* $I : \mathcal{A} \to \mathbb{R}$ by

$$I(w) := \int_\Omega f(x, w(x), \nabla w(x))\, dx, \;\; w \in \mathcal{A}.$$

Assume that $u \in \mathcal{A}$ is a minimizer of I on $\mathcal{A}$, $v \in C^1(\Omega) \cap C(\overline{\Omega})$ with $v = 0$ on $\partial\Omega$ and define $h : \mathbb{R} \to \mathbb{R}$ by $h(t) := I(u + tv)$. Since h has a global minimum at $t = 0$, we have

$$0 = h'(0) = \delta I(u; v) = \int_\Omega \left(\frac{\partial f}{\partial u} v + \sum_{i=1}^{N} \frac{\partial v}{\partial x_i} \frac{\partial f}{\partial p_i} \right) dx \tag{1.3}$$

where $f = f(x, u, p)$ for $p = (p_1, \ldots, p_N)$. If we further assume that $u \in C^2(\Omega)$, then we can conclude (after integrating by parts and using the fundamental lemma of calculus of variations) that u satisfies the *Euler-Lagrange Equation* ([Gelfand and Fomin, 1963])

$$\begin{cases} \dfrac{\partial f}{\partial u} - \displaystyle\sum_{i=1}^{N} \frac{\partial}{\partial x_i} \frac{\partial f}{\partial p_i} = 0 & \text{in } \Omega \\ u = g & \text{in } \partial\Omega. \end{cases} \tag{1.4}$$

Conversely, classical solutions of (1.4) correspond to points where the first variation of u vanishes (i.e. they satisfy (1.3) for all appropriate v). However, it is generally not possible to proceed from (1.3) to (1.4) when $u \notin C^2(\Omega)$. That is, not all minimizers of I are classical solutions to the Euler-Lagrange equation (1.4). As an example (see [Buttazzo et al., 1998]) we take $\Omega = (-1, 1)$, $g(-1) = 0, g(1) = 1$ and

$$I(w) = \int_{-1}^{1} w^2(x)(2x - w'(x))^2\, dx, \;\; w \in \mathcal{A}.$$

Then u defined by

$$u(x) = \begin{cases} x^2 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

is a minimizer of I on $\mathcal{A}$ that clearly does not belong to $C^2(\Omega)$. A good reference for many such one-dimensional examples is [Ball and Mizel, 1987].

However, even though we cannot interpret $u \in \mathcal{A}$ satisfying $\delta I(u; \cdot) = 0$ as a solution to (1.4) in the classical sense, we can still think of it as a *generalised* or *weak solution* to (1.4). This makes sense because every classical solution is a weak solution, while the converse is clearly not true. Furthermore, the only reason for a weak solution u not to be a classical solution is if u does not have enough regularity for (1.4) to make sense. Thus, to recover a classical solution from a weak solution, we need to justify the *regularity* of the weak solution.

In summary, we have observed that classical solutions of certain PDEs can be characterised as minimizers — or more generally, critical points — of certain functionals. This allows us to prove the existence of solutions to certain PDEs by proving that the corresponding functional has a minimizer or critical point. However, in trying to take advantage of this insight, we encountered two practical problems:

1. Not every functional that is bounded below has a minimizer or critical point.

2. Critical points of certain functionals need not have enough regularity to be classical solutions to their corresponding Euler-Lagrange equation.

In the next sections we outline how we will go about solving these problems in the dissertation, with specific reference to special semilinear elliptic equations. Unsurprisingly, addressing the problems involves a good choice of $\mathcal{A}$, and placing appropriate conditions on I.

1.2 Direct Methods and Critical Points

Direct Methods

We will first attack the problem of existence of minimizers. To this end, we recall and prove one of the fundamental theorems of optimization.

Theorem 1.2.1 (Weiestrass' Extreme Value Theorem)**.** *Let X be a sequentially compact topological space and $I : X \to \mathbb{R}$ be a continuous function (or even sequentially continuous). Then I is bounded below and attains its infimum.*

Proof. Let $\mu := \inf_{u \in X} I(u)$.
Step 1: By definition of μ, there exists a sequence $(u_n) \subseteq X$ such that $I(u_n) \to \mu$; i.e., $I(u_n) = \mu + o(1)$, where $o(1)$ is a sequence converging to 0.
Step 2: Since X is sequentially compact, we can extract a subsequence — still denoted by (u_n) — such that $u_n \to u$ for some $u \in X$.
Step 3: Since I is (sequentially) continuous, we also have that $I(u_n) \to I(u)$; hence $-\infty < I(u) = \mu$ as required. ■

We deliberately highlighted the three steps in the above proof to carefully examine the importance of each hypothesis on I and X for this result to hold. We summarise them below:

1. The existence of the sequence (u_n) such that $I(u_n) = \mu + o(1)$ requires no hypotheses placed on I and X; it is just a consequence of the definition of the infimum. This sequence is called a *minimizing sequence.*

2. The existence of a convergent subsequence is a compactness requirement, and is often too strong in many applications. The same conclusion holds if we assume that X is compact.

3. The last condition is continuity. We actually do not need the full strength of continuity, but it is enough to be able to conclude that

$$I(u) \leq \liminf_{n \to \infty} I(u_n) = \mu.$$

   That is, for I to be *sequentially lower semicontinuous.*

In our applications, $X = \mathcal{A}$ will be a function space, and will generally not be compact with respect to its natural topology (usually a norm topology). However, we do gain compactness — at least on bounded sets — by using a weaker topology. Loosely speaking, a weaker topology has fewer open sets, and hence more compact sets. However, this reduction in open sets could also tamper with the continuity of I, since fewer open sets results in fewer continuous functions. Hence, a good choice for a topology is one that optimally balances these two competing requirements.

The procedure for establishing the existence of a minimizers via the above three-step procedure is an example of the so-called *Direct Methods of Calculus of Variations* that was pioneered by Gauss, Lord Kelvin, Dirichlet, Riemann, Hilbert and Tonelli. We will explore this in more detail in Section 3.2.

Critical Points

Taking a step back, we recall that the reason that the search for solutions to PDEs leads to searching for minimizers of certain functionals is that minimizers are critical points. However, not every critical point arises as a minimizer. In general, a functional may have critical points of saddle type. Let X be a Banach space and $I \in C^1(X, \mathbb{R})$ be a functional on X. A point $u \in X$ is a *critical point* of I if $I'(u) = 0$, where I' is the *Frechet derivative* of I, and we call $c = I(u)$ the corresponding *critical value*.

In establishing the existence of critical points, the *sublevel sets* $I^a := \{u \in X : I(u) \leq a\}, a \in \mathbb{R}$ play a decisive role. Indeed, if $c \in \mathbb{R}$ is a critical value, then there is a sufficiently small $\epsilon > 0$ such that $I^{c+\epsilon}$ and $I^{c-\epsilon}$ are not homeomorphic. That is, there is a change in the topology of the sub-level sets when passing through c. This change in topology actually occurs more generally when c is a *Palais-Smale (PS) Level* for I (see [Palais and Smale, 1964] and [Ambrosetti and Rabinowitz, 1973]); that is, when there exists a sequence $(u_n) \subseteq X$ such that $I(u_n) \to c$ and $I'(u_n) \to 0$ as $n \to \infty$ (such a sequence is called a $(PS)_c$ sequence). A critical value c is also a (PS) level, since if $I(u) = c$ and $I'(u) = 0$, then the constant sequence $u_n = u$ is clearly a $(PS)_c$ sequence. This idea is precisely stated in the *deformation theorem*, which says that if c is not a PS level for I, then there is a sufficiently small $\epsilon > 0$ such that the set $I^{c+\epsilon}$ can be continuously deformed to $I^{c-\epsilon}$. In short, the only obstruction to deformations is the presence of PS sequences.

It is not enough to know that c is a PS level though — we need c to be a critical value to establish the existence of a critical point. Obtaining a critical value from a PS level c is achieved by introducing the *Palas-Smale compactness condition* at level c ($(PS)_c$ condition) on I: every $(PS)_c$ sequence has a (strongly) convergent subsequence. Indeed, if c is a PS level for I and I satisfies $(PS)_c$ then for a $(PS)_c$ sequence $(u_n) \subseteq X$, there exists $u \in X$ such that, up to a subsequence,

$$I(u) = \lim_{n\to\infty} I(u_n) = c \text{ and } I'(u) = \lim_{n\to\infty} I'(u_n) = 0.$$

Thus u is a critical point of I.

One set of conditions that ensures the existence of $(PS)_c$ sequences for I — and consequently, critical points if I satisfies $(PS)_c$ — is provided by the celebrated *Mountain Pass Theorem* (MPT) of [Ambrosetti and Rabinowitz, 1973]. This theorem gives geometric conditions that guarantee the existence of such sequences. We discuss critical point theory, and particularly the MPT, in Section 3.3.

1.3 Weak Solutions and Sobolev Spaces

After discussing the methods of finding solutions (either directly or via critical points), we turn to the definition of a solution to a PDE, or equivalently, the domain $\mathcal{A}$ of the energy functional that contains the competing functions. The optimization procedure described above establishes solutions as limits of sequences, so we need to ensure that $\mathcal{A}$ is (at least) sequentially closed in an appropriate topology. For second order PDEs, it may seem natural at first to take $\mathcal{A}$ to be a subset of $C^2(\Omega) \cap C(\overline{\Omega})$.

After all, any classical solution must at least be twice differentiable for us to justify taking second derivatives. However, it turns out that these function spaces are not well-behaved under taking limits and establishing compactness results, thus making it difficult to establish the existence of solutions using the methods described above.

The solution to this problem is to relax the definition of a solution to a PDE, and choose $\mathcal{A}$ to be strictly bigger than $C^2(\Omega) \cap C(\overline{\Omega})$. In fact, looking back at Dirichlet's Principle we can see that the functional I is well-defined even on $C^1(\Omega)$. That is, we do not need the second derivative of u to even exist to make sense of $I(u)$. However, once we have found a minimizer u of I, we cannot immediately interpret it as a classical solution to Laplace's equation, since it need not belong to $C^2(\Omega) \cap C(\overline{\Omega})$. We can, however, think of u as a kind of *weak solution* to Laplace's equation.

The weakening of the definition of a solution is achieved by choosing $\mathcal{A}$ to be a subset of a *Sobolev space.*

The Sobolev Spaces (see [Adams and Fournier, 2003]), as studied in Chapter 2, are constructed by weakening the notion of the derivative. Indeed, if $u \in C^1(\Omega)$ and $v \in C_c^\infty(\Omega)$, then integration by parts gives

$$\int_\Omega u \frac{\partial v}{\partial x_i}\, dx = -\int_\Omega \frac{\partial u}{\partial x_i} v\, dx \tag{1.5}$$

for $i = 1, 2, \ldots, N$. The boundary terms are zero since $v = 0$ on the boundary. Note that (1.5) still makes sense even when $u \in L^1_{\text{loc}}(\Omega)$, with an appropriate function $w \in L^1_{\text{loc}}(\Omega)$ replacing $\dfrac{\partial u}{\partial x_i}$ on the right hand side (assuming such a function exists). If (1.5) holds for every $v \in C_c^\infty(\Omega)$, the function w is then called the *weak partial derivative* (or simply, the *weak derivative*) of u and is also denoted by $\dfrac{\partial u}{\partial x_i}$. This abuse of notation is justified since if the classical (strong) derivative exists, then the weak derivative also exists and the two coincide (up to a set of measure zero).

In general, if α is a multi-index and $u \in L^1_{\text{loc}}(\Omega)$, then u has an αth weak partial derivative if there exists a unique function $D^\alpha u \in L^1_{\text{loc}}(\Omega)$ such that

$$\int_\Omega u D^\alpha v\, dx = (-1)^{|\alpha|} \int_\Omega D^\alpha u\, v\, dx \text{ for every } v \in C_c^\infty(\Omega).$$

One then defines the Sobolev spaces $W^{k,p}(\Omega)$ for $k = 1, 2, \ldots,$ and $p \geq 1$ as

$$W^{k,p}(\Omega) := \{u \in L^p(\Omega) : \text{ for every } |\alpha| \leq k, D^\alpha u \text{ exists and belongs to } L^p(\Omega)\}.$$

For $1 \leq p < \infty$ and $k = 1, 2, \ldots$, the norm

$$||u||_{W^{k,p}(\Omega)} := \left(\sum_{|\alpha| \leq k} ||D^\alpha u||_p^p \right)^{1/p}, \quad u \in W^{k,p}(\Omega) \tag{1.6}$$

turns $W^{k,p}(\Omega)$ into a Banach space and an analogous inner product turns $H^k(\Omega) := W^{k,2}(\Omega)$ into a Hilbert space. We then have $C_c^\infty(\Omega)$ as a subspace of $W^{k,p}(\Omega)$ and denote its closure with respect to the norm $||\cdot||_{W^{k,p}(\Omega)}$ by

$$W_0^{k,p}(\Omega) := \overline{C_c^\infty(\Omega)}.$$

We also set $H_0^k(\Omega) := W_0^{k,p}(\Omega)$.

A property of Sobolev spaces that we use frequently in applications to second order PDEs is the fact that $H_0^1(\Omega)$ is continuously embedded into $L^q(\Omega)$ for any $1 \leq q \leq 2^* := \frac{2N}{N-2}$, and that the embedding is compact for $1 \leq q < 2^*$. The latter implies that for any $1 \leq q < 2^*$, every bounded sequence in $H_0^1(\Omega)$ has a strongly convergent subsequence in $L^q(\Omega)$. The embedding is not compact when $q = 2^*$, and this has serious consequences for the existence or non-existence of solutions to certain semilinear PDEs — more about this later in the chapter.

A useful special case is when the dimension $N = 1$. Here the Sobolev spaces are compactly embedded into $L^q(\Omega)$ for any $q < \infty$ when Ω is a bounded interval. Also, every function $u \in W^{1,p}(\Omega)$ has a continuous representative, in the sense that there exists a unique absolutely continuous function $\tilde{u}$ such that $u(x) = \tilde{u}(x)$ a.e. $x \in \Omega$. In fact, this has a generalization to any dimension if $p > N$, and is a consequence of *Morrey's Inequality* (see [Evans, 1998]).

So we will choose $\mathcal{A}$ to be a subset of an appropriate Sobolev space. The steps in finding a solution can then be summarised into the following modified version of the program proposed by [Brezis, 2010].

1. We define the notion of a *weak solution* to a PDE. Such a solution will belong to a Sobolev space. For example, a weak solution to the Euler-Lagrange equation (1.4) must satisfy (1.3) for all $v \in C_c^\infty(\Omega)$.

2. We identify a functional whose critical points are weak solutions to the PDE. As mentioned before, not all PDEs are of this variational form, but we will limit our study to these.

3. We then prove that the functional identified above does indeed have a critical point. This can be achieved either through minimization or via methods of critical point theory.

4. Finally, we show that the weak solution has enough regularity to be a classical solution. This step is usually very delicate and we will only prove it in simple special cases in this dissertation.

Next, we discuss how these steps will be applied to specific problems that are of interest to us.

1.4 Linear and Semilinear Elliptic PDEs

In this dissertation we limit our focus to equations of the form

$$\begin{cases} -\Delta u = f(x,u) & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{1.7}$$

where Ω is a bounded domain in $\mathbb{R}^N$ with smooth boundary $\partial\Omega$, $N \geq 3$ and $f : \Omega \times \mathbb{R} \to \mathbb{R}$ is a continuous function. We will solve (1.7) for different functions f throughout Chapter 3 and Chapter 4.

Linear Equations

Chapter 3 begins with a study of the linear equation

$$\begin{cases} -\Delta u + cu = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{1.8}$$

where $c \geq 0, c \in L^\infty(\Omega)$ and $f \in L^2(\Omega)$. Weak solutions of (1.8) correspond to critical points of $I : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$I(u) := \frac{1}{2}\int_\Omega |\nabla u|^2 + cu^2\, dx - \int_\Omega fu\, dx,\ \ u \in H_0^1(\Omega).$$

We use the Riesz representation theorem in Hilbert spaces to prove the existence of a unique solution to (1.8). We then proceed to study the eigenvalue problem:

$$\begin{cases} -\Delta u + cu = \lambda u & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{1.9}$$

Non-trivial solutions to (1.9) are called *eigenvectors* or *eigenfunctions*, and the corresponding values of λ for which such non-trivial solutions exist are called *eigenvalues* of the linear operator $L := -\Delta + c$. We prove that L has a countably infinite sequence of eigenvalues $0 < \lambda_1 \leq \lambda_2 \leq \ldots$. The corresponding eigenvectors $(\varphi_n)_{n=1}^\infty$ can be chosen to form an orthonormal basis of $L^2(\Omega)$, or even $H_0^1(\Omega)$ with an appropriate norm.

The proof of this result uses the celebrated *Spectral Theorem for Compact Self-Adjoint Operators* (which we recall in the next chapter) in the following way. The inverse operator $K := L^{-1}$ is first shown to be a linear, compact, self-adjoint operator on $L^2(\Omega)$. Then we invoke the spectral theorem to obtain a countably infinite sequence of positive eigenvalues $(\mu_n)_{n=1}^\infty$ such that $\lim\limits_{n\to\infty} \mu_n = 0$. It is then easy to see that $(\lambda_n) = (1/\mu_n)$ is a sequence of eigenvalues for L with the same eigenvectors. The smallest eigenvalue λ_1 is of great importance throughout this dissertation. Its corresponding eigenvector φ_1 is *simple*, and can be chosen to be positive in Ω. Finally, we show that the eigenvalues satisfy

$$\lambda_n = \inf_{u \in E_{n-1}^\perp \setminus \{0\}} \frac{\int_\Omega |\nabla u|^2 + cu^2\, dx}{\int_\Omega u^2\, dx} = \inf_{\substack{u \in E_{n-1}^\perp \\ ||u||_2^2 = 1}} \int_\Omega |\nabla u|^2 + cu^2\, dx = \int_\Omega |\nabla \varphi_n|^2 + c\varphi_n^2\, dx, \tag{1.10}$$

where $E_0 = \{0\}, E_n := \text{Span}\{\varphi_1, \ldots, \varphi_n\}$ for $n \geq 1$ and $E_n^\perp$ is the orthogonal complement in $L^2(\Omega)$.

Semilinear Equations: Sub-critical Exponents

The story of non-linear equations is vastly different and more challenging compared to its linear counterpart. Let us return to the original problem (1.7), whose weak solutions correspond to critical points of the functional $I : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$I(u) = \frac{1}{2}\int_\Omega |\nabla u|^2\, dx - \int_\Omega F(x,u)\, dx,\ \ u \in H_0^1(\Omega)$$

where $F(x,t) := \int_0^t f(x,s)\, ds$ is a primitive of f. An important example that we study in detail is when $f(x,u) = |u|^{p-2}\, u$ for some $p > 2$:

$$\begin{cases} -\Delta u = |u|^{p-2}\, u & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{1.11}$$

This problem has been studied extensively in the literature. As discussed below, the exponent p plays a crucial role in determining whether or not (1.11) has non-trivial solutions. Another particularly important variation is that of positive solutions to (1.11) when $p = 2^* := \frac{2N}{N-2}$ — the so-called *critical Sobolev exponent*. These are solutions to

$$\begin{cases} -\Delta u = u^{2^*-1} & \text{in } \Omega \\ u > 0 & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{1.12}$$

Equation (1.12) is related to the *Yamabe problem* ([Lee and Parker, 1987]).

When f grows sub-linearly, (1.7) is easily shown to have weak solutions using direct methods (i.e. minimizing I on $H_0^1(\Omega)$). The super-linear case presents problems since the functional I is no-longer bounded from below. We consider two alternatives to finding critical points of I in the special case when $f(x,u) = |u|^{p-2}\, u$ for $2 < p < 2^*$.

For the first method, we use the Mountain Pass Theorem to prove the existence of critical points. The fact that $p < 2^*$ plays a decisive role in the proof that I satisfies the Palais-Smale condition, due to the compactness of the embedding of $H_0^1(\Omega)$ into $L^p(\Omega)$ when $p < 2^*$.

For the second method, we minimize I on the sphere $\mathcal{A}_p := \left\{u \in H_0^1(\Omega) : ||u||_p = 1\right\}$, since I is bounded below on $\mathcal{A}_p$. Using the method of Lagrange multipliers, we show that a positive scalar multiple of the minimizer is a solution to (1.11). Again, the fact that $p < 2^*$ is important in proving that $\mathcal{A}_p$ is sequentially closed with respect to the weak topology of $H_0^1(\Omega)$. Closure of $\mathcal{A}_p$ is needed to ensure that the limit of any minimizing sequence for I belongs to $\mathcal{A}_p$.

Semilinear Equations: The Critical Exponent

The main difficulty in solving (1.11) is when $p = 2^*$ — the Critical Sobolev Exponent. This is the subject of Chapter 4.

We first observe that the question of existence depends on the domain Ω. If Ω is *star-shaped* with respect to the origin, then (1.12) has no solutions by a result of [Pohozhaev, 1965]. However, we also find non-trivial radial solutions when $\Omega = \left\{x \in \mathbb{R}^N : a < |x| < b\right\}$ $(0 < a < b)$ is a ring. This special property of the domain allows us to transform the problem to a one-dimensional one, where the Sobolev spaces are compactly embedded into any L^p space for $p < \infty$. More generally, [Bahri and Coron, 1988] showed that non-trivial solutions of (1.11) exist when Ω has non-trivial topology.

Next we present a slightly modified version of the celebrated result by ([Brézis and Nirenberg, 1983]). Like them, we consider a linear perturbation of (1.11):

$$\begin{cases} -\Delta u = \lambda u + u^{2^*-1} & \text{in } \Omega \\ u > 0 & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{1.13}$$

We use the fact that solutions to (1.13) are positive critical points of the functional $I : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$I(u) := \frac{1}{2}\int_\Omega |\nabla u|^2 - \lambda \left(u^+\right)^2\, dx - \frac{1}{2^*}\int_\Omega \left(u^+\right)^{2^*}\, dx,\ u \in H_0^1(\Omega).$$

Due to the critical exponent 2^*, the $(PS)_c$ condition fails at most energy levels c. Define

$$S_\lambda := \inf_{u \in H_0^1(\Omega)\setminus 0} \frac{\int_\Omega |\nabla u|^2 - \lambda u^2 \, dx}{\left(\int_\Omega |u|^{2^*} \, dx\right)^{\frac{2}{2^*}}} \text{ and } S := S_0.$$

We prove that if $S_\lambda < S$, then compactness is recovered at levels $c < \frac{1}{N} S^{\frac{N}{2}}$ and a non-trivial solution exists. Also, there exists $0 \leq \lambda_* < \lambda_1$ such that $S_\lambda < S$ for any $\lambda_* < \lambda < \lambda_1$. Furthermore, $\lambda_* = 0$ if $N \geq 4$ and $\lambda_* = \frac{\lambda_1}{4}$ if Ω is an open ball in $\mathbb{R}^3$. Here λ_1 is the smallest eigenvalue of $-\Delta$.

Finally, we consider the non-homogeneous equation

$$\begin{cases} -\Delta u + cu = |u|^{2^*-2} u + f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{1.14}$$

where $f \in L^2(\Omega), f \neq 0$ and $c \geq 0$ with $c \in L^\infty(\Omega)$. It is shown in [Tarantello, 1992] that this equation has at least two solutions. We use methods similar to [Naito and Sato, 2012] involving the Implicit Function Theorem to prove the existence of a solution to (1.14).

1.5 The Structure of the Dissertation

In the next chapter (Chapter 2) we put together and recall many results that are used throughout the dissertation. These include results from calculus in $\mathbb{R}^N$, functional analysis, Sobolev spaces, and partial differential equations. The reader is assumed to be familiar with all these results and proofs are not given.

The third chapter develops the theory for solving PDEs via variational methods. We study both direct methods and critical point theory. We then apply these to linear and semilinear PDEs with sub-critical exponents.

Chapter 4 is concerned with solving semilinear elliptic PDEs with the critical Sobolev exponent by modifying the techniques developed in Chapter 3.

The last chapter is the conclusion, and gives an overview of the main results proved in the dissertation in relation with further research in the topic.

This dissertation should be accessible to a student with enough background in elementary topology and functional analysis.

Chapter 2

Preliminaries

We begin by recalling some definitions and results from basic analysis. It is assumed that the reader is familiar with these results, and they are only included for reference. Specifically, we assume that the reader is familiar with analysis in Banach and Hilbert spaces, Measure Theory, Sobolev spaces, and General Topology. We will work exclusively with real vector spaces and use 0 for both the zero vector and the number zero. Proofs of the results stated in this section will not be provided but can be found in many standard references (e.g. [Dunford et al., 1971],[Brezis, 2010]).

2.1 Differential and Integral Calculus in $\mathbb{R}^N$

Throughout this section (and chapter), we let N be a positive integer and $\Omega \subseteq \mathbb{R}^N$ be a domain (i.e. open and connected). We will denote the boundary of Ω by $\partial\Omega$ and assume that this boundary is smooth (even though some results only require less regularity, e.g. Lipschitz). We first introduce some notation and terminology that we will use throughout the dissertation. A good reference for this is [Evans, 1998].

2.1.1 Notation and Terminology

If $\alpha = (\alpha_1, \ldots, \alpha_N) \in \mathbb{N}^N$ is a *multi-index* and $u : \Omega \to \mathbb{R}$, we denote the αth *partial derivative* of u (if it exists) as

$$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \cdots \partial x_N^{\alpha_N}},$$

where $|\alpha| := \alpha_1 + \ldots + \alpha_N$ is the order of the multi-index. We write

$$\nabla u = \left(\frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_N} \right)$$

for the *gradient* and

$$\Delta u = \sum_{i=1}^{N} \frac{\partial^2 u}{\partial x_i^2}$$

for the *Laplacian* of u. If all derivatives of u of order $|\alpha| \leq k$ exist, we say that u is k-times differentiable. We also define

$$C^k(\Omega) := \{u : \Omega \to \mathbb{R} : D^\alpha u \text{ exists and is continuous for all } |\alpha| \leq k\} \text{ and } C^\infty(\Omega) := \cap_{k=1}^{\infty} C^k(\Omega).$$

We will use the convention that

$$D^{(0,0,\ldots,0)}u := u$$

and identify $C^0(\Omega)$ with $C(\Omega) := \{u : \Omega \to \mathbb{R} : u \text{ is continuous on } \Omega\}$. We will normally refer to these functions as C^k functions. The *support* of u is defined as

$$\mathrm{spt}(u) := \overline{\{x \in \Omega : u(x) \neq 0\}},$$

where $\overline{A}$ is the closure of $A \subseteq \mathbb{R}^N$ in the usual topology. The function u is said to have *compact support* in Ω if $\mathrm{spt}(u)$ is a compact subset of Ω. We denote by $C_c^k(\Omega)$ (resp. $C_c^\infty(\Omega)$) the class of C^k (resp. C^∞) functions with compact support in Ω.

2.1.2 Integration in $\mathbb{R}^N$

Let $1 \leq p \leq \infty$ and λ denote the Lebesgue measure in $\mathbb{R}^N$. We define $L^p = L^p(\Omega) = L^p(\Omega, \lambda)$ to be the space of all (equivalence classes of) Lebesgue measurable functions $u : \Omega \to \mathbb{R}$ such that $||u||_p < \infty$, where $||\cdot||_p$ is defined by

$$||u||_p := \left(\int_\Omega |u|^p \, d\lambda\right)^{\frac{1}{p}} \text{ if } p < \infty$$

and

$$||u||_\infty := \inf\left\{C > 0 : \lambda(\{x \in \Omega : |u(x)| > C\} = 0)\right\}.$$

We will use dx instead of $d\lambda$ when evaluating integrals with respect to the Lebesgue measure, i.e.

$$\int_\Omega u \, dx = \int_\Omega u(x) \, dx = \int_\Omega u \, d\lambda.$$

We define $L^p_{\mathrm{loc}}(\Omega)$ to be the set of functions $u : \Omega \to \mathbb{R}$ that are locally in $L^p(\Omega)$, in the sense that $u \in L^p(K)$ for any compact subset K of Ω. Let $1 \leq q \leq \infty$ be such that $\frac{1}{p} + \frac{1}{q} = 1$; $q = \infty$ when $p = 1$ and $q = 1$ when $p = \infty$ (q is called the *conjugate exponent* of p). If $u \in L^p(\Omega)$ and $v \in L^q(\Omega)$, then $uv \in L^1(\Omega)$ and we have *Holder's inequality*:

$$||uv||_1 \leq ||u||_p \, ||v||_q \, .$$

We will denote the $N-1$-dimensional surface integral by

$$\int_S u \, d\sigma(x),$$

where S is an $N-1$-dimensional surface. A vector field X is a function $X : \Omega \to \mathbb{R}^N$. If $X = (X_1, \ldots, X_N)$ is a C^1 vector field (component-wise), we define the divergence of X as

$$\mathrm{div}X := \sum_{i=1}^N \frac{\partial X_i}{\partial x_i}.$$

The following result is fundamental.

Theorem 2.1.1 (Divergence Theorem)**.** *Let $X : \Omega \to \mathbb{R}^d$ be a C^1 vector field. Then*

$$\int_\Omega \mathrm{div}X \, dx = \int_{\partial\Omega} X \cdot \mathbf{n} \, d\sigma(x),$$

where $\mathbf{n}$ is the outward unit normal to $\partial\Omega$.

Applying this result to $X = fg\mathbf{e}_i$, where $f, g \in C^1(\Omega)$ are real-valued and $\mathbf{e}_i = (0, 0, \ldots, 1, \ldots, 0, 0)$ is the unit vector in $\mathbb{R}^N$ with 1 on the ith position and zeros everywhere else, we get

Theorem 2.1.2 (Integration by Parts). *Let $f, g : \Omega \to \mathbb{R}$ be C^1 functions. Then for each $i = 1, 2, \ldots, N$*

$$\int_\Omega \frac{\partial f}{\partial x_i} g \, dx = -\int_\Omega \frac{\partial g}{\partial x_i} f \, dx + \int_{\partial\Omega} fg\mathbf{n}_i \, d\sigma(x),$$

where $\mathbf{n}_i$ is the ith component of the outward unit normal to $\partial\Omega$.

If f or g has compact support, the second integral on the right vanishes.

Corollary 2.1.1. *If $f \in C^1_c(\Omega)$ and $g \in C^1(\Omega)$, then*

$$\int_\Omega \frac{\partial f}{\partial x_i} g \, dx = -\int_\Omega \frac{\partial g}{\partial x_i} f \, dx.$$

This can be generalized to arbitrary derivatives.

Corollary 2.1.2. *If $f \in C^k_c(\Omega)$ and $g \in C^k(\Omega)$, then for any multi index α with $|\alpha| \leq k$,*

$$\int_\Omega (D^\alpha f) \, g \, dx = (-1)^{|\alpha|} \int_\Omega (D^\alpha g) \, f \, dx.$$

We end with an argument that we will be using extensively in later chapters.

Theorem 2.1.3 (Differentiating Under the Integral Sign). *Fix $i \in \{1, 2, \ldots, m\}$ and let $f : \Omega \times \mathbb{R}^m \to \mathbb{R}$ satisfy the following:*

1. *For each $y \in \mathbb{R}^m$, the map $x \mapsto f(x, y)$ is in $L^1(\Omega)$*
2. *For each $x \in \Omega$, $\dfrac{\partial f}{\partial y_i}$ exists*
3. *There exists $h \in L^1(\Omega)$ such that*

$$\left| \frac{\partial f}{\partial y_i} \right| \leq h(x) \text{ for every } x \in \Omega, y \in \mathbb{R}^m.$$

Define $g : \mathbb{R}^m \to \mathbb{R}$ by

$$g(y) := \int_\Omega f(x, y) \, dx \text{ for every } y \in \mathbb{R}^m.$$

Then $\dfrac{\partial g}{\partial y_i}$ exists and

$$\frac{\partial g}{\partial y_i} = \int_\Omega \frac{\partial f}{\partial y_i} \, dx.$$

2.2 Results from Functional Analysis

Let E and F be normed spaces over the real numbers and $T : E \to F$ be a linear operator between them. We say that T is *bounded* (or equivalently, *continuous*) if there exists $C > 0$ such that $||Tu|| \leq C||u||$ for every $u \in E$. We denote by $\mathcal{L}(E, F)$ the normed space of bounded linear operators with norm

$$||T|| = ||T||_{\mathcal{L}(E,F)} := \sup_{u \in E, u \neq 0} \frac{||Tu||}{||u||} = \sup_{\substack{u \in E \\ ||u||=1}} ||Tu|| \,.$$

We denote $\mathcal{L}(E, E)$ simply by $\mathcal{L}(E)$. When $F = \mathbb{R}$ we call $\mathcal{L}(E, \mathbb{R})$ the (continuous) *dual space* of E and denote it by E'. We also note that $\mathcal{L}(E, F)$ is complete if F is complete; so in particular, E' is always complete. Elements of E' are called *bounded (or continuous) linear functionals*. The space E' has a simple characterization when E is a Hilbert space. It is clear that if E is a Hilbert space with inner product $(\cdot, \cdot)$ and $u_0 \in E$, then $u^* : E \to \mathbb{R}$ defined by $u^*(u) = (u_0, u)$ for every $u \in E$ belongs E'. The converse is also true:

Theorem 2.2.1 (Riesz Representation Theorem)**.** *Let E be a Hilbert space and $u^* \in E'$. Then there exists a unique $u_0 \in E$ such that $u^*(u) = (u_0, u)$ for every $u \in E$. Furthermore, $||u^*|| = ||u_0||$.*

We now turn to the study of the normed spaces $\left(L^p(\Omega), ||\cdot||_p\right)$ for $1 \leq p \leq \infty$.

Theorem 2.2.2 (Riesz-Fischer)**.** *For any $1 \leq p \leq \infty$ the space $\left(L^p(\Omega), ||\cdot||_p\right)$ is a Banach space, and separable for $1 \leq p < \infty$. The space $L^2(\Omega)$ is a Hilbert space with the inner product $(\cdot, \cdot)_2$ defined by*

$$(u, v)_2 := \int_\Omega uv \, dx, \;\; u, v \in L^2(\Omega).$$

The following will be very useful in later chapters.

Theorem 2.2.3. *Let $(u_n) \subseteq L^p(\Omega)$ be a sequence that converges to $u \in L^p(\Omega)$ in norm (i.e. $||u_n - u||_p \to 0$ as $n \to \infty$). Then there exists a subsequence (u_{n_k}) and a function $w \in L^p(\Omega)$ such that*

$$u_{n_k}(x) \to u(x) \text{ and } |u_{n_k}(x)| \leq w(x) \text{ a.e. in } \Omega.$$

We also have a representation of the dual of L^p that resembles that of Hilbert spaces.

Theorem 2.2.4 (Riesz Representation Theorem)**.** *Let $1 \leq p < \infty$ and $1 < q \leq \infty$ be such that $\frac{1}{p} + \frac{1}{q} = 1$ ($q = \infty$ when $p = 1$). Then $L^p(\Omega)' = L^q(\Omega)$, in the sense that for each $u^* \in L^p(\Omega)'$ there exists a unique $f \in L^q(\Omega)$ such that*

$$u^*(u) = \int_\Omega fu \, dx, \text{ for every } u \in L^p(\Omega).$$

Furthermore, $||u^*|| = ||f||_q$.

2.2.1 Weak Topologies

The topology induced by the norm in a normed space E turns out to be too strong for many of our applications — see Section 3.2 for an elaboration of this point. We will thus introduce a coarser topology on E. The *weak topology* on E, $\tau(E, E')$, is the smallest (coarsest) topology that makes all elements of E' continuous (i.e. the initial topology induced by E'). In finite dimensions, this weak topology coincides with the strong topology induced by the norm.

A neighbourhood base for the weak topology at a point $u \in E$ can be constructed as follows. For each integer $k > 0$, positive number $\epsilon > 0$ and $u_1^*, \ldots, u_k^* \in E'$, define

$$U(u; u_1^*, \ldots, u_k^*; \epsilon) := \{v \in E : |u_i^*(u - v)| < \epsilon \text{ for all } i = 1, 2, \ldots, k\}.$$

Then

$$\mathcal{N}(u) = \left\{U(u; u_1^*, \ldots, u_k^*; \epsilon) : k \in \mathbb{N}^+, u_1^*, \ldots, u_k^* \in E', \epsilon > 0\right\}$$

is a neighbourhood base at u. Thus a sequence $(u_n) \subseteq E$ converges to $u \in E$ in the weak topology (or *converges weakly* to u) if

$$\lim_{n\to\infty} u^*(u_n) = u^*(u) \text{ for every } u^* \in E'.$$

We will write $u_n \rightharpoonup u$ if (u_n) converges weakly to u and $u_n \to u$ if (u_n) converges in the norm topology (i.e. strongly) to u.

Proposition 2.2.1. *Let E be a normed space and (u_n) be a sequence in E. If $u_n \rightharpoonup u$ weakly, then (u_n) is norm bounded and*

$$||u|| \leq \liminf_{n\to\infty} ||u_n||.$$

If $u_n \to u$ strongly, then $u_n \rightharpoonup u$ weakly.

In Hilbert spaces (or more generally, uniformly convex spaces), we have a partial converse.

Theorem 2.2.5. *Let E is a Hilbert space and $(u_n) \subseteq E$ be a sequence such that $u_n \rightharpoonup u$ weakly in E, and $||u_n|| \to ||u||$. Then $u_n \to u$ strongly in E.*

Recall that a vector space E equipped with a topology τ is a *Topological Vector Space* (TVS) if the vector space operations of addition and scalar multiplication are continuous with respect to τ (for addition, the domain topology is the product of τ and τ, and for scalar multiplication, it is the product of τ and the usual topology on $\mathbb{R}$). We also say that E is a *locally convex* TVS if τ is the initial topology induced by a family of seminorms.

Lemma 2.2.1. *Let E be a normed space and $\tau(E, E')$ be the weak topology. Then $(E, \tau(E, E'))$ is a locally convex, Hausdorff topological vector space.*

The weak topology on an infinite-dimensional space is not metrizable, hence a difficulty in working with the weak topology (compared to the strong (norm) topology) is that many topological notions cannot be characterized by sequences. However, compactness is an exception.

Theorem 2.2.6 (Eberlein-Smulian)**.** *Let E be a Banach space and $K \subseteq E$. Then K is weakly compact if and only if it is weakly sequentially compact.*

The weak topology is indeed weaker than the strong topology, so in general, all weakly closed sets are strongly closed, but the converse need not hold. A similar comparison can be made with respect to continuity of functions. However, we have an exception when convexity is involved.

Theorem 2.2.7. *Let E and F be Banach spaces, $T : E \to F$ be a linear map, and C be a non-empty convex subset of E.*

1. *Then C is weakly closed if and only if it is strongly closed.*

2. *T is continuous with respect to the norm topologies of E and F (i.e. continuous strong-strong) if and only if it is continuous with respect to the weak topologies of E and F (i.e. continuous weak-weak).*

Theorem 2.2.8 (Mazur)**.** *Let E be a Banach space and $(u_n) \subseteq E$ be a sequence that converges weakly to $u \in E$. Then there exists a sequence $(v_n) \subseteq E$ of convex combinations of the terms of (u_n), such that (v_n) converges strongly to u in E.*

Now let E be a normed space. The dual space E' has a natural topology induced by the dual operator norm. Thus E' is also a normed space (in fact, a Banach space), hence we can talk about its weak topology $\tau(E', E'')$ (here $E'' := (E')'$ is the *double dual*). We can even define a weaker topology on E' that is generated by special elements of E''.

For each $u \in E$, we define a map $u^{**} : E' \to \mathbb{R}$ by

$$u^{**}(u^*) := u^*(u), \ \ u^* \in E'. \tag{2.1}$$

This is a bounded linear functional on E', i.e. $u^{**} \in E''$. Thus, each element of E can be associated with a unique element of E''. We define the *weak*topology* $\tau(E', E)$ on E' as the topology generated by all the linear functionals u^{**} of the form (2.1) for $u \in E$. It is clear that $\tau(E', E) \subseteq \tau(E', E'')$, since E is in some sense 'embedded' into E''. On the other hand, if all elements of E'' are constructed as in (2.1), then the two topologies coincide.

To make this precise we first define the embedding $J : E \to E''$ by

$$J(u) := u^{**}, \ \ u \in E.$$

Then J is a linear isometry from E into E'', but is generally not onto.

Definition 2.2.1. *A Banach space E is* reflexive *if the map J defined above is an isomorphism (i.e. it is also onto).*

By far the most important property of reflexive Banach spaces to us is the following.

Theorem 2.2.9. *If E is a reflexive Banach space, then every bounded sequence (u_n) in E has a weakly convergent subsequence.*

Many commonly used Banach spaces are reflexive. In particular, all Hilbert spaces are reflexive by the Riesz representation theorem. Also, the L^p spaces are reflexive for $1 < p < \infty$, but not for $p = 1, \infty$ since $(L^\infty(\Omega))'$ is strictly bigger than $L^1(\Omega)$.

2.2.2 Differentiation in Banach Spaces

Let E and F be Banach spaces and $I : E \to F$ be a mapping.

Definition 2.2.2. *We say that I is* Gateaux differentiable *at a point $u \in E$ if there exists a unique $I'_G(u) \in \mathcal{L}(E, F)$, called the* Gateaux derivative *of I at u, such that*

$$\lim_{t \to 0} \frac{I(u + tv) - I(u)}{t} = I'_G(u)(v)$$

for every $v \in E$.
We say that I is Frechet differentiable *(or simply, differentiable) at u if there exists a unique $I'(u) \in \mathcal{L}(E, F)$, called the* Frechet derivative *of I at u, such that*

$$\lim_{||v|| \to 0} \frac{I(u+v) - I(u) - I'(u)(v)}{||v||} = 0.$$

We say that I is Gateaux differentiable *(resp.* Frechet differentiable*) in an open subset $U \subseteq E$ if I is Gateaux differentiable (resp. Frechet differentiable) at every point $u \in U$.*

If I is Frechet differentiable, then it is also Gateaux differentiable, and the two derivatives coincide. The converse does not hold in general, except in the following important special case.

Theorem 2.2.10. *If $U \subseteq E$ is open, I is Gateaux differentiable on U and $I'_G : U \to \mathcal{L}(E, F)$ defined by $u \mapsto I'_G(u)$ is continuous in the norm topologies, then I is also Frechet differentiable, and $I' = I'_G$. In this case, we say I is continuously differentiable and write $I \in C^1(U, F)$.*

For this reason, we will often use I' for both types of derivatives. Another important special case that is useful in optimization is when $F = \mathbb{R}$. So let $U \subseteq E$ be open and let $I : U \subseteq E \to \mathbb{R}$. If I is Frechet differentiable at $u \in U$, then $I'(u) \in \mathcal{L}(E, \mathbb{R}) = E'$. In particular, when E is a Hilbert space, the Riesz representation theorem tells us that there exists a unique $u_0 \in E$ such that

$$I'(u)(v) = (u_0, v) \text{ for every } v \in E.$$

We denote this unique element u_0 by $\nabla I(u)$ and call it the *gradient* of I at u.

Recall that $u \in U$ is a *local minimizer* of I if there exists a neighbourhood $V \subseteq U$ of u in E such that $I(u) \leq I(v)$ for every $v \in V$. We say that u is a *global minimizer* of I on U if $I(u) \leq I(v)$ for every $v \in U$; a global minimizer is also a local minimizer. The following result is well-known in calculus.

Theorem 2.2.11 (Fermat)**.** *Let $U \subseteq E$ be open and $u \in U$ be a local minimizer of I. If I is Gateaux differentiable at u, then $I'(u) = 0$.*

Note that the statement that $I'(u) = 0$ means that $I'(u)(v) = 0$ for every $v \in E$, since $I'(u) \in E'$. For constrained optimization, here is a generalisation of the Lagrange multiplier rule from calculus.

Theorem 2.2.12 (Lagrange Multipliers)**.** *Let $U \subseteq E$ be open, $I \in C^1(U, \mathbb{R}) =: C^1(U)$ and $J \in C^1(U, F)$, where F is a Banach space. Set $\mathcal{A} := J^{-1}(0) = \{v \in U : J(v) = 0\}$ and assume that $u \in \mathcal{A}$ is a (global) minimizer of I on $\mathcal{A}$. If $J'(u) : E \to F$ is surjective, then there exists a* Lagrange multiplier *$\lambda \in F'$ such that*

$$I'(u) = \lambda \circ J'(u);$$

i.e.

$$I'(u)(v) = \lambda(J'(u)(v)) \text{ for all } v \in E. \tag{2.2}$$

Note that when $F = \mathbb{R}^d$, then the Lagrange multiplier $\lambda \in (\mathbb{R}^d)' \cong \mathbb{R}^d$ is a vector, and equation (2.2) is simply the dot product.

2.2.3 Compact Operators and the Spectral Theorem

Throughout this section, we assume that E and F are Banach spaces and denote by $B_E(u,r) := \{v \in E : ||u-v|| < r\}$ for $r > 0$ and $u \in E$. Also let $B_E := B_E(0,1)$ be the open unit ball in E.

Definition 2.2.3. *An operator $T \in \mathcal{L}(E,F)$ is called a* compact operator *if $T(B_E)$ has compact closure in F. Equivalently, T is compact if for every norm-bounded sequence $(u_n) \subseteq E$, $(T(u_n))$ has a strongly convergent subsequence in F. We denote the set of compact operators from E to F by $\mathcal{K}(E,F)$ and write $\mathcal{K}(E) = \mathcal{K}(E,E)$.*

Compact operators send bounded sets to totally bounded sets. Let $T \in \mathcal{K}(E,F)$ and $(u_n) \subseteq E$ be a sequence that converges weakly to $u \in E$. Then (u_n) is bounded, hence (Tu_n) has a subsequence (Tu_{n_k}) that converges strongly, and hence weakly, to some $v \in F$. Since T is continuous strong-strong, it is also continuous weak-weak, thus $v = Tu$ (the weak topology is Hausdorff, so limits are unique). But it follows from the following Lemma — which is also important in its own right — that the whole sequence (Tu_n) converges strongly to Tu in F.

Lemma 2.2.2 (Uryson's Subsequence Principle)**.** *Let (X,τ) be a topological space, $(x_n) \subseteq X$ be a sequence on X and $x \in X$. The following are equivalent:*

1. *$x_n \to x$*
2. *Every subsequence of (x_n) has a further subsequence that converges to x.*

Proposition 2.2.2. *Let $T \in \mathcal{K}(E,F)$ and $u_n \rightharpoonup u$ weakly in E. Then $Tu_n \to Tu$ strongly in F.*

Definition 2.2.4. *If $T \in \mathcal{L}(E)$, then $\lambda \in \mathbb{R}$ is called an* eigenvalue *of T if there exists a non-zero vector $u \in E$ such that $Tu = \lambda u$. The corresponding u is called an* eigenvector *corresponding to λ.*

We are primarily interested in the eigenvalues and eigenvectors of a compact operator. If E and F are Hilbert spaces and $T \in \mathcal{L}(E,F)$, then we define the *adjoint operator* of T as the unique linear operator $T^* \in \mathcal{L}(F,E)$ such that

$$(Tu, v) = (u, T^*v) \text{ for every } u \in E, v \in F.$$

We say that T is *self-adjoint* if $E = F$ and $T = T^*$; i.e. $(Tu,v) = (u,Tv)$ for every $u, v \in E$.

Theorem 2.2.13. *If $T \in \mathcal{L}(E)$ is a self-adjoint operator, then the eigenvectors corresponding to distinct eigenvalues are orthogonal*

Recall that in a Hilbert space $E = (E, (\cdot,\cdot))$, a collection $\{e_\alpha : \alpha \in \Lambda\} \subseteq E$ is an *orthonormal basis* for E if

$$(e_\alpha, e_\beta) = \begin{cases} 1 & \text{if } \alpha = \beta \\ 0 & \text{if } \alpha \neq \beta \end{cases}$$

and for every $u \in E$,

$$u = \sum_{\alpha \in \Lambda} (u, e_\alpha) e_\alpha, \tag{2.3}$$

where the sum in (2.3) has countably many non-zero terms and converges in the norm topology of E.

Theorem 2.2.14 (Spectral Theorem for Compact, Self-adjoint Operators)**.** *Let E be a separable Hilbert space and $T \in \mathcal{K}(E)$ be a compact, self-adjoint operator. Then*

1. *The set* $\sigma_p(T) = \{\lambda_n : n = 1, 2, 3, \ldots\}$ *of eigenvalues of* T *is finite or countably infinite. If* $\sigma_p(T)$ *is infinite, then* 0 *is its only point of accumulation and we can arrange the sequence* (λ_n) *so that* $\lim_{n\to\infty} \lambda_n = 0$.

2. *The corresponding eigenvectors can be chosen to form an orthonormal basis for* E.

2.3 Sobolev Spaces and Partial Differential Equations

From a high-level point view, one can think of a PDE as an equation (in u, with v fixed)

$$Tu = v,$$

where $T : E \to F$ is an operator and E and F are function spaces. This point of view is of particular importance when E and F are normed spaces, since we can then use the results developed in the previous sections to study the existence of solutions to the PDE. We are going to use precisely this language when dealing with PDEs. This section introduces these ideas, and more importantly, identifies the 'right' spaces E and F to work with. We closely follow presentations from [Evans, 1998] and [Brezis, 2010].

2.3.1 Sobolev Spaces

An obvious choice for E and F would be the space of C^k functions, where k is the order of the PDE. Unfortunately though, the regularity that is required by these spaces makes it difficult to establish existence results using functional analysis. Also, these spaces lack many of the desirable properties of normed spaces like reflexivity; see [Evans, 1998] for an elaboration of this point. In this section we introduce an alternative — the *Sobolev spaces* — which turns out to be very useful in the study of PDEs. The main idea in their construction is to weaken the notion of the derivative. We only briefly discuss key results and definitions and refer the reader to the standard literature for a more detailed exposition (see e.g. [Evans, 1998], [Brezis, 2010], [Adams and Fournier, 2003]).

Throughout the section, we will assume that $\Omega \subseteq \mathbb{R}^N$ is an open and connected set (a domain). If $u \in C^1(\Omega)$, then the integration by parts formula (Theorem 2.1.2) tells us that

$$\int_\Omega u \frac{\partial \varphi}{\partial x_i}\, dx = -\int_\Omega \frac{\partial u}{\partial x_i} \varphi \, dx, \text{ for every } \varphi \in C_c^1(\Omega). \tag{2.4}$$

We want to weaken the regularity requirement on u. Note that the left hand side of (2.4) is well defined even when u is only in $L^1_{\text{loc}}(\Omega)$; but then $\frac{\partial u}{\partial x_i}$ on the right will have no meaning in that case. Thus, we will need to replace $\frac{\partial u}{\partial x_i}$ with an appropriate function $v \in L^1_{\text{loc}}(\Omega)$ such that (2.4) remains true.

Definition 2.3.1. *Let* $u \in L^1_{loc}(\Omega)$*. We say that* $v \in L^1_{loc}(\Omega)$ *is a* weak partial derivative *(or simply, a* derivative*) of* u *if*

$$\int_\Omega u \frac{\partial \varphi}{\partial x_i}\, dx = -\int_\Omega v\varphi \, dx, \text{ for every } \varphi \in C_c^\infty(\Omega). \tag{2.5}$$

If such v *exists, we say that* u *is* weakly differentiable *(or* differentiable*). More generally, if* α *is a multi-index, we say that* $v \in L^1_{loc}(\Omega)$ *is a* αth weak derivative *of* u *if*

$$\int_\Omega u D^\alpha \varphi \, dx = (-1)^{|\alpha|} \int_\Omega v\varphi \, dx, \text{ for every } \varphi \in C_c^\infty(\Omega).$$

We should be calling v "the" weak derivative, and we will do so after the following results.

Lemma 2.3.1 (Fundamental Lemma of Calculus of Variations)**.** *Let $u \in L^1_{loc}(\Omega)$ be such that*

$$\int_\Omega u\varphi \, dx = 0 \text{ for every } \varphi \in C_c^\infty(\Omega).$$

Then $u = 0$ a.e. in Ω.

The following results follow easily.

Proposition 2.3.1. *If $u \in L^1_{loc}(\Omega)$ is weakly differentiable, then its weak derivative is unique up to a set of measure zero.*

Proposition 2.3.2. *If $u \in C^k(\Omega)$, then for any $|\alpha| \leq k$, the αth weak derivative of u coincides with the classical αth derivative of u.*

We will thus use the same notation for both the weak and classical derivative, even if the function is only weakly differentiable. We are now ready to define the Sobolev spaces.

Definition 2.3.2. *Let $1 \leq p \leq \infty$ and k be a positive integer. We define the* Sobolev space $W^{k,p}(\Omega)$ *as the set of (equivalence classes of) functions in $L^p(\Omega)$ whose weak derivatives of order k or less exist and belong to $L^p(\Omega)$. That is,*

$$W^{k,p}(\Omega) := \{u \in L^p(\Omega) : \text{ for every } |\alpha| \leq k, D^\alpha u \text{ exists (weakly) and belongs to } L^p(\Omega)\}.$$

We also write $H^k(\Omega) := W^{k,2}(\Omega)$ and $W^{0,p}(\Omega) := L^p(\Omega)$.

We will equip $W^{k,p}(\Omega)$ with the norm

$$||u||_{W^{k,p}} := \begin{cases} \left(\sum_{|\alpha| \leq k} ||D^\alpha u||_p^p \right)^{1/p} & \text{when } 1 \leq p < \infty \\ \sum_{|\alpha| \leq k} ||D^\alpha u||_\infty & \text{when } p = \infty \end{cases}, \quad u \in W^{k,p}(\Omega). \tag{2.6}$$

We also note that $H^k(\Omega)$ is an inner product space with the scalar product:

$$(u, v)_{H^k(\Omega)} := \sum_{|\alpha| \leq k} (D^\alpha u, D^\alpha v)_2 = \sum_{|\alpha| \leq k} \int_\Omega D^\alpha u D^\alpha v \, dx, \;\; u, v \in H^k(\Omega).$$

We will sometimes suppress the reference to Ω to simplify notation when there is no risk of confusion. We note that for any $k \in \mathbb{N}$ and $1 \leq p \leq \infty$,

$$W^{k+1,p}(\Omega) = \left\{ u \in W^{k,p}(\Omega) : \frac{\partial u}{\partial x_i} \in W^{k,p}(\Omega) \text{ for } i = 1, \ldots, N \right\} \subseteq W^{k,p}(\Omega) \subseteq W^{1,p}(\Omega) \subseteq L^p(\Omega).$$

Thus, most analysis will focus on $W^{1,p}(\Omega)$ for $1 \leq p \leq \infty$.

Theorem 2.3.1. *The Sobolev spaces $\left(W^{k,p}(\Omega), ||\cdot||_{W^{k,p}}\right)$ are Banach spaces for each $k \in \mathbb{N}$ and $1 \leq p \leq \infty$. The spaces $\left(H^k(\Omega), (\cdot,\cdot)_{H^k}\right)$ are Hilbert spaces for each $k \in \mathbb{N}$. Furthermore, $W^{k,p}(\Omega)$ is reflexive if and only if $1 < p < \infty$ and separable if and only if $1 \leq p < \infty$.*

It is clear that $C_c^\infty(\Omega) \subseteq W^{k,p}(\Omega)$, but we have more when $\Omega = \mathbb{R}^N$.

Theorem 2.3.2. *The space $C_c^\infty(\mathbb{R}^N)$ is dense in $W^{k,p}(\mathbb{R}^N)$.*

For general $\Omega \subseteq \mathbb{R}^N$, the above result no longer holds. We denote by $W_0^{k,p}(\Omega)$ the norm closure of $C_c^\infty(\Omega)$ in $W^{k,p}(\Omega)$; that is

$$W_0^{k,p}(\Omega) := \overline{C_c^\infty(\Omega)}$$

with respect to the topology induced by the norm $||\cdot||_{W^{k,p}}$. We also define $H_0^k(\Omega) := W_0^{k,2}(\Omega)$ for $k \in \mathbb{N}$. The elements of $W_0^{k,p}(\Omega)$ are often interpreted as functions $u \in W^{k,p}(\Omega)$ such that

$$D^\alpha u = 0 \text{ on } \partial\Omega \text{ for all } |\alpha| \leq k-1.$$

However, we cannot as yet make such an interpretation, since Sobolev functions are only defined up to a Lebesgue null set, and $\partial\Omega$ has N-dimensional Lebesgue measure zero. The following result allows us to make this interpretation more precise.

Theorem 2.3.3. *Assume that Ω is bounded with a smooth boundary. Then there exists a bounded linear operator $T : W^{1,p}(\Omega) \to L^p(\partial\Omega)$, called the* trace operator*, such that $Tu = u$ for every $u \in W^{1,p}(\Omega) \cap C(\bar{\Omega})$.*

We then call Tu the *trace* of u on $\partial\Omega$. We can then make a precise statement of the above heuristic.

Theorem 2.3.4. *Assume that Ω is bounded with a smooth boundary, and $u \in W^{1,p}(\Omega)$. Then*

$$u \in W_0^{1,p}(\Omega) \iff Tu = 0 \text{ on } \partial\Omega.$$

In general, if $g \in L^p(\partial\Omega)$ and $u \in W^{k,p}(\Omega)$, then by $u = g$ on $\partial\Omega$ we mean $Tu = g$. The space $H_0^1(\Omega)$ will play a crucial role in the study of second order PDEs, so we present some useful facts about it.

Theorem 2.3.5. *Let $\Omega \subseteq \mathbb{R}^N$ be bounded (i.e. $\Omega \subseteq B_{\mathbb{R}^N}(0,R)$ for some $R > 0$) and $u \in H_0^1(\Omega)$. Then $|u| \in H_0^1(\Omega)$, and if $v : \mathbb{R}^N \to \mathbb{R}$ is defined by*

$$v(x) = \begin{cases} u(x) & \text{if } x \in \Omega \\ 0 & \text{otherwise,} \end{cases}$$

then $v \in H^1(\mathbb{R}^N)$.

We will denote by $H^{-1}(\Omega)$ the dual space of $H_0^1(\Omega)$; that is $H^{-1}(\Omega) := (H_0^1(\Omega))'$. We already know that if $u \in W^{k,p}(\Omega)$, then $u \in L^p(\Omega)$. It turns out that we can say more.

Definition 2.3.3. *If $(E, ||\cdot||_E)$ and $(F, ||\cdot||_F)$ are Banach spaces, we say that E is* continuously embedded *into F (denoted $E \hookrightarrow F$) if there exists an injective linear map $i : E \to F$ and a constant $C > 0$ such that*

$$||i(u)||_F \leq C||u||_E \text{ for every } u \in E.$$

We say that E is compactly embedded *into F if the linear map i above is a compact operator. We will identify u and $i(u)$.*

Our aim is to show that if $u \in W^{k,p}(\Omega)$, then, in addition to $u \in L^p(\Omega)$, u also belongs to other L^q-spaces for some $q > p$. Let $N \geq 3$, $1 \leq p < N$ and define the *Sobolev conjugate* of p to be

$$p^* := \frac{Np}{N-p} > p.$$

Theorem 2.3.6. *Let $\Omega \subseteq \mathbb{R}^N$ be a bounded domain in $\mathbb{R}^N$ with smooth boundary and $1 \leq p < N$. Then $W^{1,p}(\Omega) \hookrightarrow L^{p^*}(\Omega)$ and there exists a constant $C > 0$, dependent only on p, N and Ω, such that*

$$||u||_{p^*} \leq C||u||_{W^{1,p}} \text{ for every } u \in W^{1,p}(\Omega).$$

We will abuse notation and write

$$||\nabla u||_p := \left(\int_\Omega |\nabla u|^p \, dx \right)^{\frac{1}{p}}$$

for $u \in W^{1,p}(\Omega)$. In $W_0^{1,p}(\Omega)$ we have even better results.

Theorem 2.3.7 (Poincare's Inequality). *Assume that Ω is a bounded domain and $u \in W_0^{1,p}(\Omega)$ for some $1 \leq p < N$. The there exists a constant $C > 0$, depending only on p, q, N and Ω, such that*

$$||u||_q \leq C||\nabla u||_p$$

for each $q \in [1, p^*]$.

Remark 2.3.1. This theorem implies that when Ω is bounded, the norm $||\nabla u||_p$ is equivalent to the usual $W^{1,p}$ norm. In that case, we will use the following notation to denote this norm:

$$||u||_{W_0^{1,p}} := ||\nabla u||_p, \ u \in W_0^{1,p}(\Omega).$$

The next result is the most important result of this section. We will use it repeatedly in applications to PDE theory in the next chapters.

Theorem 2.3.8 (Rellich-Kondrachov Compactness Theorem). *Assume that Ω is a bounded domain with smooth boundary and $1 \leq p < N$. Then $W^{1,p}(\Omega)$ is compactly embedded into $L^q(\Omega)$ for every $1 \leq q < p^*$.*

Remark 2.3.2. It follows that $W_0^{1,p}(\Omega)$ is also compactly embedded into $L^q(\Omega)$ for every $1 \leq q < p^*$.

Remark 2.3.3. The embedding is not compact when $q = p^*$ (for $N \geq 3$).

We end with a useful characterization of Sobolev spaces in one dimension taken from [Brezis, 2010].

Theorem 2.3.9. *Let $N = 1$, $\Omega \subseteq \mathbb{R}$ be an open interval and assume that $1 \leq p \leq \infty$.*

1. *For each $u \in W^{1,p}(\Omega)$, there exists a unique absolutely continuous function $\tilde{u} \in C(\overline{\Omega})$ such that $u(x) = \tilde{u}(x)$ a.e. in Ω and*

$$\tilde{u}(x) - \tilde{u}(y) = \int_y^x u'(t)\, dt, \text{ for } x, y \in \overline{\Omega}.$$

The function $\tilde{u}$ is called the continuous representative *of u. We will often use $\tilde{u}$ in place of u when appropriate.*

2. *If Ω is bounded, then $W^{1,p}(\Omega)$ is compactly embedded into $C(\overline{\Omega})$ for all $1 < p \leq \infty$.*

3. *If Ω is bounded, then $W^{1,p}(\Omega)$ is compactly embedded into $L^q(\Omega)$ for all $1 \leq q < \infty$.*

2.3.2 Semilinear Partial Differential Equations

Here we introduce notation commonly used in the study of PDEs and discuss the variational formulation of some semilinear elliptic PDEs. Throughout this section, we assume that Ω is a fixed bounded domain in $\mathbb{R}^N$ with smooth boundary. We consider second order semilinear PDEs with Dirichlet boundary conditions of the form

$$\begin{cases} Lu = f(x,u) & \text{in } \Omega \\ u = g & \text{on } \partial\Omega, \end{cases} \tag{2.7}$$

where $f : \Omega \times \mathbb{R} \to \mathbb{R}$ and $g : \partial\Omega \to \mathbb{R}$ are functions satisfying appropriate regularity conditions, and L is a partial differential operator of the form

$$Lu := -\sum_{i,j=1}^{N} \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial u}{\partial x_j}\right) + \sum_{i=1}^{N} b_i(x)\frac{\partial u}{\partial x_i} + c(x)u,$$

for some coefficient functions $(a_{ij}), b_i, c,\ i,j = 1,2,\ldots,N$. We will assume that the matrix $A(x) = (a_{ij}(x))$ is symmetric for all $x \in \Omega$. We call a function $u \in C^2(\Omega) \cap C(\bar{\Omega})$ a *classical solution* of (2.7) if u satisfies (2.7) for each $x \in \overline{\Omega}$ (assuming that f and g are continuous). If $u \in C^2(\Omega) \cap C(\bar{\Omega})$ is a classical solution of (2.7), then multiplying (2.7) by any $v \in C_c^\infty(\Omega)$ and integrating by parts gives

$$\int_\Omega \sum_{i,j=1}^{N} a_{ij}\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j} + \sum_{i=1}^{N} b_i \frac{\partial u}{\partial x_i} v + cuv\,dx = \int_\Omega f(x,u)v\,dx. \tag{2.8}$$

Just like in the definition of the weak derivative, we again notice that (2.8) is well defined when $u \in H^1(\Omega), v \in H_0^1(\Omega)$ and f satisfies the following growth condition:

$$\exists\ C > 0 \text{ and } 2 < q \leq 2^* = \frac{2N}{N-2} \text{ such that } |f(x,u(x))| \leq C|1 + |u(x)|^{q-1}|\ \forall x \in \Omega. \tag{2.9}$$

Thus, we will define $u \in H^1(\Omega)$ to be a *weak solution* of (2.7) if (2.8) holds for every $v \in H_0^1(\Omega)$ and $u = g$ on $\partial\Omega$ in the trace sense. Note that if $g = 0$, then u is a weak solution of (2.7) if (2.8) holds for every $v \in H_0^1(\Omega)$ and $u \in H_0^1(\Omega)$. If $u \in H^1(\Omega)$ is a weak solution to (2.7), and φ is any smooth function such that $\varphi = g$ on $\partial\Omega$, then $\tilde{u} := u - \varphi$ solves the homogeneous equation

$$\begin{cases} \tilde{L}\tilde{u} = \tilde{f}(x,\tilde{u}) & \text{in } \Omega \\ \tilde{u} = 0 & \text{on } \partial\Omega, \end{cases} \tag{2.10}$$

where $\tilde{L}$ and $\tilde{f}$ are appropriately modified. So we will restrict ourselves to the case when $g = 0$ and study solutions to the equation

$$\begin{cases} Lu = f(x,u) & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{2.11}$$

Definition 2.3.4. *We say that $u \in H_0^1(\Omega)$ is a* weak solution *to (2.11) if*

$$\int_\Omega \sum_{i,j=1}^{N} a_{ij}\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j} + \sum_{i=1}^{N} b_i \frac{\partial u}{\partial x_i} v + cuv\,dx = \int_\Omega f(x,u)v\,dx \tag{2.12}$$

for every $v \in H_0^1(\Omega)$.

Using integration by parts, it is easy to show that the notion of a weak solution is indeed more general.

Proposition 2.3.3. *If $u \in H_0^1(\Omega)$ is a classical solution to (2.11), then u is also a weak solution to (2.11).*

There are a couple of reasons why it might be useful or even necessary to weaken the definition of a solution. First, physical constraints may prevent solutions to many real world PDEs from being smooth, thus a weak solution would make sense in such a context. A second reason is that it is usually easier to prove the existence of a weak solution than proving the existence of a classical solution, and under certain conditions, one can often prove that a weak solution is in fact a classical solution.

In finding solutions to PDEs, we will proceed using the following steps from [Brezis, 2010]:

1. Defining precisely what we mean by a weak solution using Sobolev spaces.
2. Prove (usually via variational methods) that a weak solution exists.
3. Prove that the weak solution found is sufficiently regular (i.e. belongs to $C^2(\Omega) \cap C(\overline{\Omega})$).
4. Show that the solution is a classical solution by proving that every weak solution that is sufficiently regular is a classical solution.

The last step is immediate from integration by parts.

Proposition 2.3.4. *If $u \in C^2(\Omega) \cap C(\overline{\Omega})$ is a weak solution to (2.11), then u is a classical solution.*

Step 3 is usually the most difficult to establish and we will avoid it. In this direction, we only state the following useful result that can be found in [Struwe, 1990].

Theorem 2.3.10. *Let $\Omega \subseteq \mathbb{R}^N$ be a bounded domain with smooth boundary and $f : \Omega \times \mathbb{R} \to \mathbb{R}$ be a function such that $x \mapsto f(x,t)$ is measurable for all $t \in \mathbb{R}$ and $t \mapsto f(x,t)$ is continuous for every $x \in \Omega$. Further assume that f satisfies the growth condition (2.9). Let $u \in H_0^1(\Omega)$ be a weak solution to*

$$\begin{cases} -\Delta u = f(x,u) & x \in \Omega \\ u = 0 & x \in \partial\Omega. \end{cases} \tag{2.13}$$

Then u is a classical solution[1]*.*

We now make further assumptions about L. We say that L is *uniformly elliptic* if there exists $\theta > 0$ such that

$$\xi^T A(x)\xi \geq \theta\, |\xi|^2 \text{ for every } \xi \in \mathbb{R}^N \text{ and } x \in \Omega.$$

A classical example of a uniformly elliptic operator is the Laplace operator $Lu := -\Delta u$. Another uniformly elliptic operator we will use very often is $Lu = -\Delta u + cu$. The following result can be found in [Gilbarg and Trudinger, 2015].

Theorem 2.3.11 (Strong Maximum Principle)**.** *Let Ω be a domain, $Lu = -\Delta u + cu$ where $c \geq 0$ and suppose that $u \in C^2(\Omega) \cap C(\overline{\Omega})$ satisfies*

$$Lu \geq 0 \text{ in } \Omega.$$

If $\inf_{x \in \overline{\Omega}} u(x) \leq 0$ and this infimum is attained at some $x_0 \in \Omega$, then u is constant on Ω.

---

[1]To be precise, there exists a classical solution $\tilde{u}$ such that $\tilde{u} = u$ a.e.

Chapter 3

Linear Elliptic PDE and Variational Methods

3.1 Linear PDE

In this section we show how the existence of weak solutions to linear PDEs can be established using elementary results from linear functional analysis. Even though the results hold for a wider class of linear elliptic operators L, we will only discuss the special case when $L = -\Delta + c$ for some $c \in L^\infty(\Omega)$. This simple case still captures all of the main functional analytic ideas used without the notational complications of a more general linear elliptic operator. The style of this presentation follows [Evans, 1998] and [Badiale and Serra, 2010], and a discussion of more general operators can be found in [Evans, 1998].

3.1.1 General Existence Results

We consider solutions to the linear elliptic PDE presented in Section 2.3.2 for the case when $L = -\Delta + c$ and $f = f(x)$ does not depend on u. We will focus our attention to the problem

$$\begin{cases} -\Delta u + cu = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{3.1}$$

where $f \in L^2(\Omega)$, $\Omega \subseteq \mathbb{R}^N$ is a bounded domain with smooth boundary $\partial\Omega$, and $c \in L^\infty(\Omega)$.

Recall that $u \in H_0^1(\Omega)$ is a weak solution of (3.1) if

$$\int_\Omega \nabla u \cdot \nabla v \, dx + \int_\Omega cuv \, dx = \int_\Omega fv \, dx \quad \text{for every } v \in H_0^1(\Omega). \tag{3.2}$$

We will show that under a certain condition on c, (3.1) has a unique weak solution. We do so by first showing that under this condition, the left hand side of (3.2) is an inner product that induces a norm that is equivalent to the norm on $H_0^1(\Omega)$, and then apply the Riesz Representation Theorem to the bounded linear functional $u \mapsto (f, u)_2$.

Lemma 3.1.1. *If $c \geq 0$, then the bilinear form $B_c : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ defined by*

$$B_c(u, v) := \int_\Omega \nabla u \cdot \nabla v \, dx + \int_\Omega cuv \, dx, \; (u, v) \in H_0^1(\Omega) \times H_0^1(\Omega)$$

is an inner product on $H_0^1(\Omega)$ that induces a norm $||\cdot||_c$ that is equivalent to $||\cdot||_{H_0^1(\Omega)}$, the usual norm on $H_0^1(\Omega)$.

Proof. The fact that B_c is a bilinear form is obvious. Also, for $u \in H_0^1(\Omega)$ we have

$$B_c(u,u) = \int_\Omega |\nabla u|^2 \, dx + \int_\Omega cu^2 \, dx \geq 0,$$

and if $B_c(u,u) = 0$ then

$$0 = B_c(u,u) = \int_\Omega |\nabla u|^2 \, dx + \int_\Omega cu^2 \, dx = ||u||^2_{H_0^1(\Omega)} + \int_\Omega cu^2 \, dx \geq ||u||^2_{H_0^1(\Omega)},$$

which implies that $u = 0$. Clearly B_c is symmetric, thus it is an inner product on $H_0^1(\Omega)$.
We now prove equivalence of the norms $||\cdot||_c$ and $||\cdot||_{H_0^1(\Omega)}$. First it is clear that $||u||_c \geq ||u||_{H_0^1(\Omega)}$ for every $u \in H_0^1(\Omega)$. On the other hand, Poincare's inequality gives (for $u \in H_0^1(\Omega)$)

$$||u||_c^2 = ||u||^2_{H_0^1(\Omega)} + \int_\Omega cu^2 \, dx \leq ||u||^2_{H_0^1(\Omega)} + ||c||_\infty \int_\Omega u^2 \, dx \leq C ||u||^2_{H_0^1(\Omega)}$$

for some $C > 0$. ∎

Remark 3.1.1. The condition that $c \geq 0$ is actually not necessary for B_c to be an equivalent inner product (and later, for the existence of a solution). It is enough to assume that c is not 'too negative', in the sense that $c \geq -\mu$ for some $\mu > 0$. However, the non-negative case is the only case of interest to us. See [Evans, 1998] for the more general case.

We can now apply the Riesz Representation theorem on the Hilbert space $(H_0^1(\Omega), B_c(\cdot,\cdot))$.

Proposition 3.1.1. *If $c \geq 0$, then for every $f \in L^2(\Omega)$ there exists a unique weak solution $u \in H_0^1(\Omega)$ to (3.1).*

Proof. First note that the functional $v \mapsto (f,v)_2$ is a bounded linear functional on $(H_0^1(\Omega), B_c)$. Indeed, for every $v \in H_0^1(\Omega)$

$$|(f,v)_2| \leq ||f||_2 ||v||_2 \leq C||f||_2 ||v||_{H_0^1(\Omega)} \leq C||f||_2 ||v||_c$$

for some constant $C > 0$. By the Riesz Representation Theorem, there exists a unique $u \in H_0^1(\Omega)$ such that

$$B_c(u,v) = (f,v)_2 \text{ for every } v \in H_0^1(\Omega).$$

But this says

$$\int_\Omega \nabla u \cdot \nabla v \, dx + \int_\Omega cuv \, dx = \int_\Omega fv \, dx \text{ for every } v \in H_0^1(\Omega).$$

∎

Remark 3.1.2. The above procedure cannot be applied to prove the existence of solutions for general linear PDEs, since we used the fact that the operator $-\Delta + c$ induces a *symmetric* bilinear form, which turned out to be an equivalent inner product on $H_0^1(\Omega)$. However, in some cases when the operator is not symmetric (but satisfies *continuity* and *coercivity* — see [Evans, 1998] and Section 3.2 below), one can still prove the existence of a unique weak solution to (3.1) using the *Lax-Milgram Theorem* ([Lax and Milgram, 2016]).

Remark 3.1.3. It is clear that the above result holds more generally for every $f \in (H_0^1(\Omega))' = H^{-1}(\Omega)$, and the proof is essentially unchanged. However, for our purposes, the L^2 case is sufficient and we thus felt the need to express it explicitly.

We now turn to the question of regularity of the solution found above. From Proposition 2.3.4, we know that all we need to show is that the solution $u \in C^2(\Omega) \cap C(\overline{\Omega})$. This is a delicate topic and we give a sketch (see [Brezis, 2010]) for the case of one dimension $N = 1$. Let $\Omega = (a, b)$ for some $a, b \in \mathbb{R}$ with $a < b$. In this case, the equation can be written as

$$\begin{cases} -u'' + cu = f & \text{in } (a,b) \\ u(a) = u(b) = 0 \end{cases} \tag{3.3}$$

We first assume that both f and c are continuous on $[a, b]$ and let $u \in H_0^1((a, b))$ be a weak solution to (3.3). That is,

$$\int_a^b u'v' + cuv\,dx = \int_a^b fv\,dx \text{ for every } v \in H_0^1((a,b)). \tag{3.4}$$

Recall that every element of $W^{1,p}((a, b))$ (in particular $H_0^1((a, b))$) has a continuous version, so we assume that u is continuous. We can write (3.4) as

$$\int_a^b u'v'\,dx = -\int_a^b (cu - f)v\,dx \text{ for every } v \in H_0^1((a,b)),$$

which implies that $u' \in H^1((a, b))$ (since $cu - f \in L^2((a, b))$), with $(u')' = cu - f \in C([a, b])$). So $u \in H^2((a, b))$, and since u' is continuous, we have that $u \in C^2([a, b])$, implying that u is a classical solution. We refer the reader to [Brezis, 2010], [Evans, 1998] and Theorem 2.3.10 for the proof of the more general result.

Proposition 3.1.2. *If f is continuous on $\bar{\Omega}$, then the solution to (3.1) is a classical solution.*

Finally, we notice that the weak solution u to (3.1) is actually a minimizer of the functional $I : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$I(v) = \frac{1}{2}\int_\Omega |\nabla v|^2\,dx + \frac{1}{2}\int_\Omega cv^2\,dx - \int_\Omega fv, \ \ v \in H_0^1(\Omega).$$

To see this, first note that u being a weak solution to (3.1) means that u is a critical point of I (i.e. $I'(u)(v) = 0$ for every $v \in H_0^1(\Omega)$) since the Frechet derivative of I is given by

$$I'(u)(v) = \int_\Omega \nabla u \cdot \nabla v\,dx + \int_\Omega cuv\,dx - \int_\Omega fv\,dx \text{ for every } v \in H_0^1(\Omega).$$

Hence for any $v \in H_0^1(\Omega)$ we have

$$I(u+v) = I(u) + \frac{1}{2}\int_\Omega |\nabla v|^2\,dx + \frac{1}{2}\int_\Omega cu^2\,dx + I'(u)(v) \geq I(u).$$

Thus, an equivalent method of establishing the existence of solutions to (3.1) is to find a minimizer or critical point of I. We do follow this approach later in the chapter.

3.1.2 Eigenvalues

Let Ω be a bounded domain with smooth boundary. We now consider the problem of finding eigenvalues of the operator $L := -\Delta + c$ for $c \geq 0, c \in L^\infty(\Omega)$ on $H_0^1(\Omega)$. That is, we find non-trivial weak solutions to the equation:

$$\begin{cases} Lu = -\Delta u + cu = \lambda u & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{3.5}$$

for some $\lambda \in \mathbb{R}$. The weak formulation of this problem is to find a non-trivial $u \in H_0^1(\Omega)$ such that

$$\int_\Omega \nabla u \cdot \nabla v + cuv \, dx = \lambda \int_\Omega uv \, dx \text{ for every } v \in H_0^1(\Omega). \tag{3.6}$$

This can also be written as

$$B_c(u, v) = \lambda\,(u, v)_2 \text{ for every } v \in H_0^1(\Omega).$$

Recall that an eigenvalue of the operator L is a real number λ such that (3.5) has a non-trivial solution, and the corresponding non-trivial solution is called an eigenvector or *eigenfunction* corresponding to the eigenvalue λ.

First note that Proposition 3.1.1 tells us that every $f \in L^2(\Omega)$ can be associated with a unique $u \in H_0^1(\Omega)$ via (3.1). This allows us to define a map $K : L^2(\Omega) \to H_0^1(\Omega) \subseteq L^2(\Omega)$ by

$$Kf = u \iff u \text{ is a weak solution to (3.1), i.e. } u \in H_0^1(\Omega) \text{ and } Lu = f \text{ weakly.} \tag{3.7}$$

The map K can be thought of as the inverse operator of L, and K turns out to have the right properties for the application of spectral theory. Following a presentation similar to [Evans, 1998], we will proceed as follows:

1. We first show that K is a bounded, linear, self-adjoint and compact operator on $L^2(\Omega)$.
2. From step 1 and the spectral theorem for compact and self-adjoint operators, it will follow that K has a countable sequence of real positive eigenvalues $(\mu_n)_{n=1}^\infty$ such that $\lim_{n\to\infty} \mu_n = 0$.
3. From step 2, we will be able to conclude that the differential operator $L = K^{-1}$ has a sequence of eigenvalues $(\lambda_n)_{n=1}^\infty = (1/\mu_n)_{n=1}^\infty$ that can be ordered as $0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq, \ldots,$ with $\lambda_n \to \infty$ as $n \to \infty$.

Lemma 3.1.2. *The operator $K : L^2(\Omega) \to L^2(\Omega)$ is a bounded, linear, self-adjoint, compact operator.*

Proof. First, K is well defined by Proposition 3.1.1. Let $f, g \in L^2(\Omega)$ and $\alpha, \beta \in \mathbb{R}$ with $Kf = u$ and $Kg = v$. So we have, in the weak sense,

$$Lu = f, \; Lv = g, \text{ and } u, v \in H_0^1(\Omega).$$

By the linearity of L we also have $L(\alpha u + \beta v) = \alpha Lu + \beta Lv = \alpha f + \beta g$, hence $K(\alpha f + \beta g) = \alpha u + \beta v = \alpha Kf + \beta Kg$. So K is linear.
Since $Kf = u$ we also have

$$||Kf||_2^2 = ||u||_2^2 \leq C||u||_c^2 = CB_c(u, u) = C(f, u)_2 \leq C||f||_2||u||_2 = C||f||_2||Kf||_2.$$

Thus

$$||Kf||_2 \leq C||f||_2.$$

Hence K is bounded. Now let $(f_n) \subseteq L^2(\Omega)$ be a sequence with $||f_n||= 1$. So we have

$$||Kf_n||_c^2 = B_c(Kf_n, Kf_n) = (f_n, Kf_n)_2 \leq C||f_n||_2||Kf_n||_2 \leq C||Kf_n||_c.$$

Hence (Kf_n) is bounded in $H_0^1(\Omega)$. Since $H_0^1(\Omega)$ is reflexive, there exists $u \in H_0^1(\Omega)$ such that, up to a subsequence, $Kf_n \rightharpoonup u$ weakly in $H_0^1(\Omega)$. Since $H_0^1(\Omega)$ is compactly embedded into $L^2(\Omega)$, we can conclude that $Kf_n \to u$ strongly in $L^2(\Omega)$. Thus K is a compact operator on $L^2(\Omega)$.
Finally we show that K is self-adjoint. For $f, g \in L^2(\Omega)$ we have

$$(f, Kg)_2 = B_c(Kf, Kg) = (Kf, g)_2.$$

■

We now apply the spectral theorem to get the following.

Theorem 3.1.1. *The operator L has an infinite sequence of real positive eigenvalues $0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq, \ldots,$. Furthermore, the corresponding eigenfunctions can be chosen to form an orthonormal basis of $L^2(\Omega)$.*

Proof. This is an application of the spectral theorem in Section 2.2. Since K is compact and self-adjoint on $L^2(\Omega)$, there exists a sequence (μ_n) of eigenvalues of K with corresponding eigenvectors $(\varphi_n) \subseteq H_0^1(\Omega) \subseteq L^2(\Omega)$ that form an orthonormal basis for $L^2(\Omega)$. Furthermore, $\lim_{n\to\infty} \mu_n = 0$. For each $n \geq 1$ we have

$$\mu_n B_c(\varphi_n, \varphi_n) = B_c(\mu_n\varphi_n, \varphi_n) = B_c(K\varphi_n, \varphi_n) = (\varphi_n, \varphi_n)_2, \text{ i.e. } \mu_n = \frac{||\varphi_n||_2^2}{||\varphi_n||_c^2},$$

hence $\mu_n > 0$. Now for each n, since

$$K\varphi_n = \mu_n\varphi_n \iff L(\mu_n\varphi_n) = \varphi_n,$$

it follows that

$$\lambda_n = \frac{1}{\mu_n}$$

is an eigenvalue of L with eigenvector φ_n. ■

Remark 3.1.4. From the regularity theory in [Evans, 1998], $\varphi_n \in C^\infty(\Omega)$.

Since $(\varphi_n) \subseteq H_0^1(\Omega)$ is an orthonormal basis for $L^2(\Omega)$ (we will always assume that the eigenvectors are normalized so that $||\varphi_n||_2 = 1$), we have, for $n \neq m$,

$$0 = \lambda_n (\varphi_m, \varphi_n)_2 = B_c(\varphi_n, \varphi_m),$$

Hence the eigenvectors are also orthogonal in $(H_0^1(\Omega), B_c(\cdot,\cdot))$. Furthermore, if $u \in H_0^1(\Omega)$ is such that $B_c(u, \varphi_n) = 0$ for $n = 1, 2, 3, \ldots$, then for each positive integer n,

$$0 = B_c(u, \varphi_n) = \lambda_n (u, \varphi_n)_2 \implies (u, \varphi_n)_2 = 0 \text{ since } \lambda_n > 0,$$

hence $u = 0$ by completeness of (φ_n) in $L^2(\Omega)$. Thus (φ_n) is also complete in $(H_0^1(\Omega), B_c(\cdot,\cdot))$. Note that $||\varphi_n||_c^2 = B_c(\varphi_n, \varphi_n) = \lambda_n ||\varphi_n||^2 = \lambda_n$, hence $\left(\varphi_n/\sqrt{\lambda_n}\right)$ is an orthonormal basis for

$(H_0^1(\Omega), B_c(\cdot,\cdot))$.

Note that from the proof above we see that λ_n and φ_n are related by

$$\lambda_n = \frac{||\varphi_n||_c^2}{||\varphi_n||_2^2} = ||\varphi_n||_c^2 . \tag{3.8}$$

This motivates the following characterisation of the eigenvalues in terms of the *Rayleigh quotient*:

Theorem 3.1.2 (Variational Characterisation of Eigenvalues, [Fischer, 1905], [Courant, 1920])**.** *We have*

$$\lambda_n = \inf_{u\in E_{n-1}^{\perp}\backslash\{0\}} \frac{||u||_c^2}{||u||_2^2} = \inf_{\substack{u\in E_{n-1}^{\perp}\\ ||u||_2^2=1}} ||u||_c^2 = ||\varphi_n||_c^2 , \tag{3.9}$$

where $E_0 = \{0\}, E_n := \mathrm{Span}\,\{\varphi_1,\ldots,\varphi_n\}$ *for* $n \geq 1$ *and* $E_n^{\perp}$ *is the orthogonal complement in* $L^2(\Omega)$.

Proof. If $u \in E_{n-1}^{\perp}\backslash\{0\}$, then

$$u = \sum_{k=n}^{\infty} (u,\varphi_k)_2\, \varphi_k = \sum_{k=n}^{\infty} \frac{1}{\lambda_k} B_c(u,\varphi_k)\varphi_k,$$

with convergence both in $L^2(\Omega)$ and $(H_0^1(\Omega), B_c(\cdot,\cdot))$. Since the terms are pairwise orthogonal in both spaces, we have

$$\frac{||u||_c^2}{||u||_2^2} = \frac{\sum_{k=n}^{\infty} |(u,\varphi_k)_2|^2 \lambda_k}{\sum_{k=n}^{\infty} |(u,\varphi_k)_2|^2} \geq \frac{\lambda_n \sum_{k=n}^{\infty} |(u,\varphi_k)_2|^2}{\sum_{k=n}^{\infty} |(u,\varphi_k)_2|^2} = \lambda_n$$

since $\lambda_n \leq \lambda_k$ for $k \geq n$. We have equality when $u = \varphi_n$. ∎

Remark 3.1.5. Note that since $||\nabla|u|||_2^2/|||u|||_2^2 = ||\nabla u||_2^2/||u||_2^2$ for every $u \in H_0^1(\Omega)\backslash\{0\}$, we can take $\varphi_1 \geq 0$ in Ω, and thus strictly positive in Ω by the strong maximum principle.

Of key interest to us is the smallest eigenvalue λ_1 of the operator $-\Delta + c$, and its corresponding positive, normalized eigenvector φ_1. It can also be shown (see [Evans, 1998]) that λ_1 is simple, in the sense that if u is any eigenvector corresponding to λ_1, then u is simply a scalar multiple of φ_1. That is why it makes sense to speak of a *canonical* eigenfunction φ_1. We will refer to this positive, normalized (so that $||\varphi_1||_2 = 1$) eigenvector whenever we talk about the eigenvector corresponding to λ_1. This also implies that the eigenvalues can be listed as $0 < \lambda_1 < \lambda_2 \leq \lambda_3 \leq \ldots$, to illustrate the fact that λ_1 is not a repeated eigenvalue.

The previous theorem implies that

$$\frac{||u||_c^2}{||u||_2^2} \geq \lambda_1 \text{ for any } u \in H_0^1(\Omega)\backslash\{0\} \implies ||u||_2 \leq \frac{1}{\sqrt{\lambda_1}} ||u||_c \text{ for every } u \in H_0^1(\Omega), \tag{3.10}$$

with equality occurring if and only if u is an eigenvector corresponding to λ_1. Thus letting $c \equiv 0$ we get a more explicit form of Poincare's inequality (Theorem 2.3.7) when $p = 2$:

$$||u||_2 \leq \frac{1}{\sqrt{\lambda_1}} ||\nabla u||_2 \text{ for every } u \in H_0^1(\Omega). \tag{3.11}$$

This explicit formulation will be important in the next sections.

3.2 Direct Methods of Calculus of Variations

In this section we gather conditions that guarantee the existence of a minimizer of a functional defined on a subset a Banach space. This is important since solutions to many PDEs often correspond to critical points of certain functionals, and minimizers (or maximizers) of differentiable functionals are always critical points by Fermat's Theorem (Theorem 2.2.11).

3.2.1 General Results

Let $X = (X, \tau_X)$ be a topological space and $I : X \to \mathbb{R}$ be a real-valued function on X. The assumption that I is real-valued is enough for our applications, but the same theory can be extended to when I takes values on the extend real numbers $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ (see [Kurdila and Zabarankin, 2006] for that extension). We want to place conditions on X and I such that a minimizer of I in X exists. That is, we want conditions that guarantee the existence of $u \in X$ such that

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

This problem has been solved using various techniques and by placing different conditions on I and X. The classical result is the theorem of Weierstrass discussed in Chapter 1, which guarantees the existence of such a minimizer when X is compact and I is continuous. Unfortunately, these conditions turn out to be too restrictive for most infinite-dimensional optimization problems. For instance, if X is a subset of a Banach space with τ_X being the subspace topology induced by the norm, then the abundance of τ_X-open sets results in fewer compact sets. Indeed, even the closed unit ball in an infinite-dimensional Banach space is not compact with respect to the norm topology!

A possible remedy to this situation would be to choose a coarser topology on X. However, this has to be done cautiously as well since fewer open sets leads to fewer continuous functions, thus again invalidating the assumptions of the theorem of Weierstrass. In applications, one picks a topology that is a compromise — coarse enough to generate many compact sets, yet fine enough to preserve continuity. In the Banach space setting, we achieve this middle ground by using the weak topology induced by the dual space.

Let $\mu \in [-\infty, \infty)$ be the infimum of I in X:

$$\mu := \inf_{v \in X} I(v).$$

We carefully examine the proof of the theorem by Weierstrass (Theorem 1.2.1) and select the following key steps:

1. We first pick a *minimizing sequence* for I; i.e., a sequence $(u_n) \subseteq X$ such that $I(u_n) = \mu + o(1)$ (recall that $o(1)$ represents a sequence that tends to 0 as $n \to \infty$). Such a sequence always exists by the definition of μ.

2. We then show that this minimizing sequence, at least up to a subsequence, converges to a point $u \in X$. This is where compactness is used.

3. Finally we use continuity of I to get

$$I(u) = \lim_{n \to \infty} I(u_n) = \mu,$$

and conclude that u is the required minimizer of I in X.

In this section we will relax some of the requirements of Theorem 1.2.1, while keeping the same direct approach of finding a minimizer via minimizing sequences. To avoid notational inconveniences, we will continue to use (u_n) for both the main sequence and any subsequence of it, if this causes no confusion. This tradition is common in the literature.

We start by relaxing the continuity assumption. The space $X = (X, \tau_X)$ will be a topological space with a Hausdorff topology τ_X, which we will progressively enrich with more structure.

Definition 3.2.1. *Let $I : X \to \mathbb{R}$ be a function and $u \in X$ be a point of accumulation of X.*

1. *We say that I is* lower semicontinuous (lsc) *at u if*

$$I(u) \leq \liminf_{v \to u} I(v).$$

That is, if for any $\epsilon > 0$, there exists a neighbourhood $\mathcal{N}$ of u such that $I(v) > I(u) - \epsilon$ for all $v \in \mathcal{N}$. We say that I is lsc on X if it is lsc at every point $u \in X$.

2. *We say that I is* sequentially lower semicontinuous (sequentially lsc) *at u if for every sequence $(u_n) \subseteq X$ such that $u_n \to u$ in τ_X, we have*

$$I(u) \leq \liminf_{n \to \infty} I(u_n).$$

We say that I is sequentially lsc on X if I is sequentially lsc at every point $u \in X$.

It is clear that lower semicontinuity implies sequential lower semicontinuity, but not the other way around. The converse is true though in a sequential space (e.g. a metric space).

Proposition 3.2.1. *If X is a first countable topological space (i.e. each point of X has a countable base of neighbourhoods), then I is lsc on X if and only if it is sequentially lsc on X.*

Proof. ($\implies$): This one is obvious.
($\impliedby$): Fix $u \in X$ to be a point of accumulation and assume that I is not lsc at u. We want to show that this implies that I is not sequentially lsc at $u \in X$. Let $(\mathcal{N}_n)_{n=1}^{\infty}$ be a countable base for the neighbourhoods at u. Since I is not lsc at u, there exists $\epsilon > 0$ such that we can find $u_1 \in \mathcal{N}_1$ such that $I(u_1) \leq I(u) - \epsilon$. Similarly, we can pick $u_2 \in \mathcal{N}_1 \cap \mathcal{N}_2$ such that $I(u_2) \leq I(u) - \epsilon$. Continuing this way we can construct a sequence $(u_n) \subseteq X$ such that

$$u_n \in \bigcap_{k=1}^{n} \mathcal{N}_k \text{ and } I(u_n) \leq I(u) - \epsilon \text{ for every } n.$$

Now let $\mathcal{N}$ be an arbitrary neighbourhood of u. There exists a basic neighbourhood $\mathcal{N}_N$ such that $u \in \mathcal{N}_N \subseteq \mathcal{N}$. Thus, if $n \geq N$, then

$$u_n \in \bigcap_{k=1}^{n} \mathcal{N}_k \subseteq \mathcal{N}_N \subseteq \mathcal{N},$$

implying that $u_n \to u$. We also have

$$\liminf_{n \to \infty} I(u_n) \leq I(u) - \epsilon < I(u).$$

Thus I is not sequentially lsc at u. ■

This result is useful since sequences are easier to deal with, and all normed spaces are first countable with respect to the norm topology. Here is another characterisation of lsc functions.

Lemma 3.2.1. *The function $I : X \to \mathbb{R}$ is lsc on X if and only if for every $a \in \mathbb{R}$, the set $I^a := \{u \in X : I(u) \leq a\}$ is closed in τ_X.*

Proof. ($\implies$): Assume that I is lsc on X and for each $a \in \mathbb{R}$ define $A^a := \{u \in X : I(u) > a\} = X \backslash I^a$. We need to show that A^a is an open set. Take $u \in A^a$ and pick $0 < \epsilon < I(u) - a$. Then there exists a neighbourhood $\mathcal{N}$ of u such that $I(v) > I(u) - \epsilon > a$ for every $v \in \mathcal{N}$. That is, there exists a neighbourhood $\mathcal{N}$ of u such that $\mathcal{N} \subseteq A^a$.
($\impliedby$): Now assume that I^a is closed for every $a \in \mathbb{R}$ and pick $u \in X$ and $\epsilon > 0$. Then $A^{I(u)-\epsilon} = \{v \in X : I(v) > I(u) - \epsilon\}$ is open and $u \in A^{I(u)-\epsilon}$, hence there exists a neighbourhood $\mathcal{N}$ of u such that $u \in \mathcal{N} \subseteq A^{I(u)-\epsilon}$; i.e. $I(v) > I(u) - \epsilon$ for all $v \in \mathcal{N}$. ■

An analogous result holds for sequentially lsc functions. Note that if a minimizer u of I exists, and (u_n) is a sequence that converges to u, then

$$I(u) \leq I(u_n)\ \forall n \in \mathbb{N}^+ \implies I(u) \leq \liminf_{n\to\infty} I(u_n);$$

that is, I is sequentially lower semicontinuous at u. So sequential lower semicontinuity at u is a necessary condition for the existence of a minimizer. We will make this assumption in our first step of improving Theorem 1.2.1.

Theorem 3.2.1. *Let X be a sequentially compact topological space and $I : X \to \mathbb{R}$ be sequentially lsc on X. Then I has a minimizer in X.*

Proof. Let $(u_n) \subseteq X$ be a minimizing sequence. Up to a subsequence, $u_n \to u \in X$ due to sequential compactness of X. By sequential lower semicontinuity of I, we get

$$I(u) \leq \liminf_{n\to\infty} I(u_n) = \mu.$$

Hence $I(u) = \mu$ as required. ■

Note that the same result is achieved by assuming compactness and semicontinuity without the 'sequential' qualification. This alternative approach is also common in the literature and heavily relies on the sub-level sets $I^a, a \in \mathbb{R}$.

3.2.2 Reflexive Banach Spaces

We now assume that $E = (E, ||\cdot||)$ is a reflexive Banach space and $X \subseteq E$.

Definition 3.2.2. *Assume that X is unbounded and let $I : X \to \mathbb{R}$ be a functional on X. We say that I is* coercive *if*

$$\lim_{\substack{||u||\to\infty \\ u\in X}} I(u) = \infty.$$

That is, if for every $K > 0$ there exists $R > 0$ such that if $u \in X$ and $||u|| > R$ then $I(u) > K$.

This condition is meant to replace the strong requirement that X be compact or sequentially compact, even though it is slightly weaker than both forms of compactness. Indeed, if I is coercive and (u_n) is a minimizing sequence, then for some $K > 0$ sufficiently large, there exists $R > 0$ such that $I(u) > K$ for all u with $||u|| > R$. Hence, (u_n) is eventually norm-bounded by R. In finite dimensions, the Bolzano-Weierstrass theorem would allow us to extract a convergent subsequence from (u_n). Unfortunately, all the Banach spaces we consider are infinite-dimensional.

The case when E is reflexive restores some hope though, since bounded sets are relatively weakly compact, by Theorem 2.2.9! Thus, if we instead equip E with the weak topology induced by its dual, then we recover relative compactness of the minimizing sequence (u_n) when I is coercive. That is, up to a subsequence, (u_n) converges weakly to some $u \in E$. The recovery of compactness is precisely due to the weak topology having fewer open sets.

However, using the weak topology also causes problems. First, we will now need I to satisfy the *stronger* condition of sequential lower semicontinuity with respect to the weak topology. We also need to ensure that $u \in X$; that is, X is weakly sequentially closed. These potential problems disappear when X is a convex subset of E and I is a convex functional on X. We will prove a theorem that shows equivalence between strong and weak notions of lower semicontinuity and closure when we assume convexity. The theorem crucially relies on the theorem of Mazur (Theorem 2.2.8).

Theorem 3.2.2. *Let $X \subseteq E$ be a convex subset. The following are equivalent:*

1. *X is closed in the norm topology (i.e. X is strongly closed)*
2. *X is closed in the weak topology (i.e. X is weakly closed)*
3. *If $(u_n) \subseteq X$ is a sequence that converges weakly in E to some $u \in E$, then $u \in X$ (i.e. X is sequentially weakly closed).*

Proof. The fact that $(1 \iff 2)$ follows from Theorem 2.2.7. Also, $(2 \implies 3)$ follows from that sequential closure is weaker than closure in any topological space. So we only need to prove that $(3 \implies 1)$. Assume that X is weakly sequentially closed and let $(u_n) \subseteq X$ be a sequence that converges strongly to $u \in E$. But then (u_n) also converges weakly to u, hence $u \in X$. ■

We have a similar result for lower semicontinuity.

Theorem 3.2.3. *Assume that $X \subseteq E$ is a convex subset of a Banach space E, and let $I : X \to \mathbb{R}$ be a convex functional. The following are equivalent:*

1. *I is sequentially lsc with respect to the weak topology (i.e. weakly sequentially lsc)*
2. *I is sequentially lsc with respect to the norm topology (i.e. strongly sequentially lsc)*
3. *I is lsc with respect to the norm topology (i.e. strongly lsc)*
4. *I is lsc with respect to the weak topology (i.e. weakly lsc)*

Proof. $(1 \implies 2)$: Assume that I is weakly sequentially lsc and let $(u_n) \subseteq X$ be a sequence that converges strongly to $u \in X$. Then (u_n) also converges weakly to u, hence

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

since I is weakly sequentially lsc.
$(2 \implies 3)$: We have shown this in Proposition 3.2.1 since the norm topology is first countable.
$(3 \implies 4)$: Assume that I is strongly lsc, and let $a \in \mathbb{R}$. The set $I^a = \{u \in X : I(u) \leq a\}$ is strongly closed and convex, hence weakly closed by the previous theorem.
$(4 \implies 1)$: This has also been shown already in Proposition 3.2.1. ■

Remark 3.2.1. In a Banach space, the norm and the square of the norm are weakly sequentially lsc.

With our conscience clear, we can finally state and prove the main result of this section.

Theorem 3.2.4. *Assume that X is weakly sequentially closed and $I : X \to \mathbb{R}$ is weakly sequentially lower semicontinuous on X. If X is unbounded, we also assume that I is coercive. Then I attains its infimum.*

Proof. Set

$$\mu := \inf_{v \in X} I(v).$$

Since I is coercive, $\mu > -\infty$. We divide the proof into steps:

1. Let $(u_n) \subseteq X$ be a minimizing sequence; i.e. $I(u_n) \to \mu$ as $n \to \infty$.
2. Since I is coercive, (u_n) is norm-bounded, and due to reflexivity of E, we can extract a subsequence, still denoted by (u_n), that converges weakly to $u \in E$. Since X is weakly sequentially closed, $u \in X$.
3. Now we apply weak sequential lower semicontinuity of I to obtain

$$\mu \leq I(u) \leq \liminf_{n \to \infty} I(u_n) = \mu.$$

Hence u is the required minimizer.

■

3.2.3 Application to Semilinear PDE

We now show how the results of the previous section can be applied to solving the PDE

$$\begin{cases} -\Delta u + cu = f(x,u) & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.12}$$

where $c \in L^\infty(\Omega), c \geq 0$, $f : \Omega \times \mathbb{R} \to \mathbb{R}$ is (jointly) continuous, and Ω is a bounded domain with a smooth boundary. We investigate the existence of solutions to (3.12) by imposing different *growth conditions* on f.

Assume that $u \in C^2(\Omega) \cap C(\overline{\Omega})$ is a classical solution of (3.12). Multiplying (3.12) through by $v \in C_c^\infty(\Omega)$ and integrating by parts we get

$$B_c(u,v) = \int_\Omega \nabla u \cdot \nabla v \, dx + \int_\Omega cuv \, dx = \int_\Omega f(x,u)v \, dx. \tag{3.13}$$

Under appropriate hypotheses on f that ensure that $f(x,u)v \in L^1(\Omega)$ for every $u, v \in H_0^1(\Omega)$, we define $u \in H_0^1(\Omega)$ to be a weak solution to (3.12) if and only if

$$B_c(u,v) = \int_\Omega f(x,u)v \, dx \text{ for every } v \in H_0^1(\Omega). \tag{3.14}$$

Define $F : \Omega \times \mathbb{R} \to \mathbb{R}$ by $F(x,t) := \int_0^t f(x,s)\,ds$ for every $x \in \Omega$ and $t \in \mathbb{R}$, and assume that $F(x,u) \in L^1(\Omega)$ for every $u \in H_0^1(\Omega)$. Then weak solutions of (3.12) correspond to points where the Gateaux derivative of $I : H_0^1(\Omega) \to \mathbb{R}$, defined by

$$\begin{aligned} I(u) &:= \frac{1}{2}\int_\Omega |\nabla u|^2\,dx + \frac{1}{2}\int_\Omega cu^2\,dx - \int_\Omega F(x,u)\,dx \\ &= \frac{1}{2}\,||u||_c^2 - \int_\Omega F(x,u)\,dx,\ \ u \in H_0^1(\Omega), \end{aligned}$$

is equal to zero. To see this, fix $u, v \in H_0^1(\Omega)$ and $t \in \mathbb{R}\backslash\{0\}$. Consider

$$g(t) := I(u+tv) = \frac{1}{2}\,||u+tv||_c^2 - \int_\Omega F(x,u+tv)\,dx = \frac{1}{2}\,||u||_c^2 + tB_c(u,v) + \frac{1}{2}t^2\,||v||_c^2 - \int_\Omega F(u+tv)\,dx.$$

Under some boundedness conditions on F, we can apply the Theorem 2.1.3 for differentiation under the integral sign to the last integral, to get the (Gateaux) derivative:

$$I_G'(u)(v) = g'(0) = B_c(u,v) - \int_\Omega f(x,u)v\,dx = \int_\Omega \nabla u \cdot \nabla v\,dx + \int_\Omega cuv\,dx - \int_\Omega f(x,u)v\,dx.$$

Placing stronger growth conditions on f, we will show that I is actually continuously Frechet differentiable on $H_0^1(\Omega)$. Thus, $u \in H_0^1(\Omega)$ is a weak solution for (3.12) if and only if $I'(u)(v) = 0$ for every $v \in H_0^1(\Omega)$; i.e. if and only if $I'(u) = 0$ in $H^{-1}(\Omega)$.

From Fermat's Theorem, we know that at a minimizer of I, the Gateaux derivative is indeed zero, thus our first attempt at finding solutions is to look for minimizers of I. Results of the previous section give us one way of achieving this — by showing that I is weakly sequentially lower semicontinuous and coercive. Closely examining the form of I, we see that the first two terms clearly satisfy these conditions (the square of a norm is always weakly sequentially lsc and coercive). So, any potential problems might be caused by the interaction of the last term with these terms. We will solve this problem by controlling the growth of f, and consequently that of F. We first assume that f satisfies the following growth condition:

$$\exists a, b > 0 \text{ and } \gamma \in (0,1) \text{ such that } |f(x,t)| \leq a + b|t|^\gamma \text{ for every } t \in \mathbb{R} \text{ and } x \in \Omega. \tag{3.15}$$

Condition (3.15) means that f grows sub-linearly, which implies that F grows sub-quadratically, since for some real constants a_1 and b_1, $|F(x,t)| \leq a_1|t| + b_1|t|^{\gamma+1}$ for every $t \in \mathbb{R}$ and $x \in \Omega$ with $\gamma + 1 < 2$. This is the condition that will guarantee coercivity of I.

Lemma 3.2.2. *Assume that (3.15) holds. Then I is weakly sequentially lsc on $H_0^1(\Omega)$.*

Proof. Let $(u_n) \subseteq H_0^1(\Omega)$ with $u_n \rightharpoonup u \in H_0^1(\Omega)$ weakly in $H_0^1(\Omega)$. We have to show that

$$I(u) \leq \liminf_{n\to\infty} I(u_n).$$

We first deal with the last term of I. By the compactness of the embedding of $H_0^1(\Omega)$ into $L^2(\Omega)$, (u_n) converges strongly to u in $L^2(\Omega)$. By Theorem 2.2.3, we can find a subsequence of (u_n) (still denoted by (u_n)) and $w \in L^2(\Omega)$ such that

$$u_n(x) \to u(x) \text{ and } |u_n(x)| \leq w(x) \text{ for almost every } x \in \Omega \text{ and } n \in \mathbb{N}^+.$$

Since F is continuous, we have that $F(x, u_n(x)) \to F(x, u(x))$ for almost every $x \in \Omega$. The growth condition on F also gives

$$|F(x, u_n(x))| \leq a_1 |u_n(x)| + b_1 |u_n(x)|^{\gamma+1} \leq a_1 |w(x)| + b_1 |w(x)|^{\gamma+1} \in L^1(\Omega)$$

since Ω has finite measure (therefore $L^p(\Omega) \subseteq L^q(\Omega)$ when $1 \leq q \leq p \leq \infty$). So we can apply the dominated convergence theorem to conclude that

$$\int_\Omega F(x, u_n)\, dx \to \int_\Omega F(x, u)\, dx \text{ as } n \to \infty.$$

Since this holds for a subsequence of (u_n), we can conclude that

$$\limsup_{n\to\infty} \int_\Omega F(x, u_n)\, dx \geq \int_\Omega F(x, u)\, dx.$$

In fact, if (u_{n_k}) is any subsequence of (u_n), then by repeating the same process as above, we see that (u_{n_k}) has a further subsequence $(u_{n_{k_j}})$ such that

$$\int_\Omega F(x, u_{n_{k_j}})\, dx \to \int_\Omega F(x, u)\, dx \text{ as } j \to \infty.$$

Thus by Lemma 2.2.2,

$$\lim_{n\to\infty} \int_\Omega F(x, u_n)\, dx = \int_\Omega F(x, u)\, dx.$$

Now since we know that $\left(u \mapsto \frac{1}{2} ||u||_c^2, u \in H_0^1(\Omega)\right)$ is weakly sequentially lsc, we can conclude that

$$\liminf_{n\to\infty} I(u_n) = \liminf_{n\to\infty} \left(\frac{1}{2} ||u_n||_c^2 - \int_\Omega F(x, u_n)\, dx\right) \geq \frac{1}{2} \liminf_{n\to\infty} ||u_n||_c^2 - \limsup_{n\to\infty} \int_\Omega F(x, u_n)\, dx$$
$$\geq \frac{1}{2} ||u||_c^2 - \int_\Omega F(x, u)\, dx = I(u).$$

■

Theorem 3.2.5. *Assume that (3.15) holds. Then (3.12) has a weak solution in $H_0^1(\Omega)$.*

Proof. We will show that I is coercive and then apply Theorem 3.2.4 to $X = (H_0^1(\Omega), B_c(\cdot, \cdot))$. Since Ω is bounded, we first note that if $1 \leq q < 2$, then $L^2(\Omega) \subseteq L^q(\Omega)$ and for any $u \in L^2(\Omega)$ we have

$$||u||_q = \left(\int_\Omega |u|^q\ dx\right)^{\frac{1}{q}} \leq \left(\left(\int_\Omega 1^{\frac{2}{2-q}}\ dx\right)^{\frac{2-q}{2}} \left(\int_\Omega |u|^2\ dx\right)^{\frac{q}{2}}\right)^{\frac{1}{q}} \leq C_q ||u||_2 .$$

Now if $u \in H_0^1(\Omega) \subseteq L^2(\Omega)$, then

$$\int_\Omega F(x, u)\, dx \leq \int_\Omega |F(x, u)|\ dx \leq a_1 \int_\Omega |u|\ dx + b_1 \int_\Omega |u|^{\gamma+1}\ dx \leq C_1 ||u||_2 + C_2 ||u||_2^{\gamma+1} .$$

Also, since $||u||_2 \leq C ||u||_c$, we have

$$\int_\Omega F(x, u)\, dx \leq C_1 ||u||_c + C_2 ||u||_c^{\gamma+1}$$

for some positive constants C_1 and C_2. Hence for any $u \in H_0^1(\Omega)$ we have

$$I(u) \geq \frac{1}{2} ||u||_c^2 - C_1 ||u||_c - C_2 ||u||_c^{\gamma+1}.$$

Since $\gamma + 1 < 2$, the first quadratic term dominates the other terms, hence

$$\lim_{||u||_c \to \infty} I(u) \geq \lim_{||u||_c \to \infty} \frac{1}{2} ||u||_c^2 - C_1 ||u||_c - C_2 ||u||_c^{\gamma+1} = \infty.$$

Thus I is coercive. Since $H_0^1(\Omega)$ is weakly closed, we can apply Theorem 3.2.4 to conclude that there exists a minimizer $u \in H_0^1(\Omega)$ for I in $H_0^1(\Omega)$. We further conclude from Fermat's Theorem that the derivative of I vanishes at u; that is

$$B_c(u, v) = \int_\Omega f(x, u) v \, dx \text{ for every } v \in H_0^1(\Omega).$$

So u is the required weak solution. ■

The growth condition (3.15) is crucial in proving that I is coercive. The coercivity condition nolonger holds when f grows super-linearly (i.e. when $\gamma > 1$), since the positive quadratic term $\frac{1}{2} ||u||_c^2$ is dominated by a negative term that grows super-quadratically. The interesting case is when the growth of f is dominated by a linear term, i.e. when $\gamma = 1$. In this case we expect the coefficients of the positive and negative terms to play a crucial role in determining whether or not I is coercive. This is pointed out in [Badiale and Serra, 2010], and we investigate it here by making the following assumption:

$$\exists a, b > 0 \text{ such that } |f(x,t)| \leq a + b|t| \text{ for every } t \in \mathbb{R} \text{ and } x \in \Omega. \tag{3.16}$$

This condition implies that

$$|F(x,t)| \leq a\,|t| + \frac{b}{2} t^2 \text{ for every } t \in \mathbb{R} \text{ and } x \in \Omega.$$

We show that coercivity is regained if we control the size of b; this consequently establishes the existence of a minimizer. In what follows, λ_1 is the smallest eigenvalue of the operator $-\Delta + c$ discussed in Section 3.1.2.

Theorem 3.2.6. *Assume that (3.16) holds with $b < \lambda_1$. Then (3.12) has a weak solution in $H_0^1(\Omega)$.*

Proof. The proof that I is weakly sequentially lsc is the same as above, thus we only show that I is coercive. From equation (3.10) we have that

$$||u||_2^2 \leq \frac{1}{\lambda_1} ||u||_c^2 \text{ for every } u \in H_0^1(\Omega).$$

Thus for any $u \in H_0^1(\Omega)$,

$$\int_\Omega F(x,u)\, dx \leq \int_\Omega |F(x,u)|\, dx \leq a\, ||u||_1 + \frac{b}{2} ||u||_2^2 \leq C\, ||u||_c + \frac{b}{2\lambda_1} ||u||_c^2.$$

Hence

$$I(u) \geq \frac{1}{2} ||u||_c^2 - C\, ||u||_c - \frac{b}{2\lambda_1} ||u||_c^2 = \frac{1}{2}\left(1 - \frac{b}{\lambda_1}\right) ||u||_c^2 - C\, ||u||_c,$$

and since $b < \lambda_1$, the coefficient of $||u||_c^2$ is positive, again implying that I is coercive. ■

The case of superlinear growth ($\gamma > 1$) is covered in the next two sections.

3.3 Critical Point Theory

In the previous section we managed to find solutions to some semilinear PDEs through minimization of an appropriate energy functional I. We achieved this by proving that I is coercive and weakly sequentially lsc. We further observed that whether or not I was coercive depended crucially on the exponent of the growth of the negative term γ. When $\gamma \in (0,1)$, I is always coercive, while coercivity is achieved in some cases when $\gamma = 1$. However, coercivity is completely lost when $\gamma > 1$, and in this case we cannot use the direct methods of the previous section.

In fact, solutions to a number of PDEs can be characterised as critical points of functionals that are neither bounded above nor below, thus also making the direct methods of the previous section inappropriate. In this section we overcome this difficulty by searching directly for critical points of I, which may now correspond to local minima or local maxima or be of saddle point type.

3.3.1 General Results

We begin with some definitions. Throughout this section, X is a Banach space and $I : X \to \mathbb{R}$ is a functional that is continuously differentiable on X, i.e. $I \in C^1(X)$. Recall that the *sublevel sets* of I are defined as

$$I^a := \{u \in X : I(u) \leq a\}.$$

We say that $c \in \mathbb{R}$ is a *critical level* or *critical value* of I if there exists $u \in X$ such that $I(u) = c$ and $I'(u) = 0$; the corresponding u is called a *critical point* of I. We are interested in determining whether or not I has a critical point. We will mimic the work of [Ambrosetti and Rabinowitz, 1973] on critical point theory as presented in [Badiale and Serra, 2010].

We first motivate the results by considering the case when $X = \mathbb{R}$ and $I(u) = \frac{u}{1+u^2}$, $u \in \mathbb{R}$. The sublevel sets of I are as follows:

$$I^a = \begin{cases} \mathbb{R} & \text{if } a > \frac{1}{2} \\ \left(-\infty, \frac{1-\sqrt{1-4a^2}}{2a}\right) \cup \left(\frac{1+\sqrt{1-4a^2}}{2a}, \infty\right) & \text{if } 0 < a \leq \frac{1}{2} \\ (-\infty, 0] & \text{if } a = 0 \\ \left(\frac{1+\sqrt{1-4a^2}}{2a}, \frac{1-\sqrt{1-4a^2}}{2a}\right) & \text{if } -\frac{1}{2} \leq a < 0 \\ \emptyset & \text{if } a < -\frac{1}{2}. \end{cases}$$

From this we observe that the form of the sublevel sets changes at the points $a = \frac{1}{2}, 0$ and $-\frac{1}{2}$, and this change of form can be described as a change in connectedness (topology). Since

$$I'(u) = \frac{1-u^2}{(1+u^2)^2}, \ u \in \mathbb{R},$$

we see that $u = \pm 1$ are critical points of I, and the corresponding critical values are $\pm\frac{1}{2}$. Thus, with the exception of 0, the change in topology of the sublevel sets occurs at the critical levels of I. However, 0 is also close to being a critical value, as there exists a sequence (u_n) (e.g. take $u_n = n$) such that $I(u_n) \to 0$ and $I'(u_n) \to 0$. In fact, if (u_n) has a subsequence (still denoted by (u_n)) that converges to some $u \in \mathbb{R}$, then since $I \in C^1(\mathbb{R})$, we get

$$I'(u) = \lim_{n\to\infty} I'(u_n) = 0 \text{ and } I(u) = 0,$$

implying that u is a critical point of I. The fact that 0 is not a critical value for I implies that there is no such subsequence for (u_n). That is, (u_n) is not relatively compact.

Definition 3.3.1. *A sequence $(u_n) \subseteq X$ such that $(I(u_n))$ is bounded in $\mathbb{R}$ and $I'(u_n) \to 0$ in X' is called a* Palais-Smale (PS) sequence.
If $c \in \mathbb{R}$, $I(u_n) \to c$ and $I'(u_n) \to 0$, then (u_n) is called a Palais-Smale sequence for I at level c *($(PS)_c$), and c is said to be a* Palais-Smale level *for I.*
We say that I satisfies the Palais-Smale condition *(PS) if every PS sequence has a strongly convergent subsequence in X. Similarly, we say that I satisfies the* Palais-Smale condition at level c *($(PS)_c$) if every $(PS)_c$ sequence has a strongly convergent subsequence in X.*

It is clear that if $u \in X$ is a critical point of I with critical level c, then c is also a Palais-Smale level for I (simply choose $u_n = u$ as a $(PS)_c$ sequence). Thus, the example above suggests that the change in the topology of the sub-level sets occurs at a Palais-Smale level of I. Using the same language, we also observe that I defined above does not satisfy $(PS)_0$, as the sequence $(u_n) = (n)$ is a $(PS)_0$ sequence with no convergent subsequence. So the level 0 is a Palais-Smale level that is not a critical level, and this can be attributed to the failure of I to satisfy $(PS)_0$.

Proposition 3.3.1. *If I satisfies $(PS)_c$, and c is a Palais-Smale level for I, then c is a critical value for I.*

Proof. Since c is a PS level for I, we can find a $(PS)_c$ sequence $(u_n) \subseteq X$. Since I satisfies $(PS)_c$, $u_n \to u \in X$ up to a subsequence. From the continuity of I and I' we get

$$I'(u) = \lim_{n\to\infty} I'(u_n) = 0 \text{ and } I(u) = \lim_{n\to\infty} I(u_n) = c.$$

Hence u is a critical point with critical value c. ■

Note that if we instead assume that I satisfies (PS), the Bolzano-Weierstrass theorem in $\mathbb{R}$ still ensures the existence of a critical point for I. Thus, if $I \in C^1(X)$ satisfies (PS), then a change of topology can only result from passing through a critical point. Let us make these ideas more formal.

Definition 3.3.2. *Let $A \subseteq B \subseteq X$ be subsets. We say that B is* deformable *in A if there exists a jointly continuous function $\eta : [0, 1] \times B \to B$ such that*

1. *$\eta(0, u) = u$ for all $u \in B$*
2. *$\eta(t, u) \in A$ for all $u \in A$ and $t \in [0, 1]$*
3. *$\eta(1, u) \in A$ for all $u \in B$.*

The mapping η is called a deformation *of B in A.*

A deformation $\eta : [0, 1] \times B \to B$ of B in A can also be viewed as a family of continuous maps $(\eta_t : B \to B)_{t\in[0,1]}$, where $\eta_t = \eta(t, \cdot)$. The continuity requirement on η implies a kind of topological similarity between A and B for such a deformation to exist. We will apply this to the sublevel sets I^a. The idea is that deforming I^b in I^a for $a < b$ will not be possible if I^b and I^a are topologically different. The next result expresses this using the correct language.

First we need some terminology. We define $C^{1,1}(X)$ to be the space of functionals in $C^1(X)$ such that the function $(X \ni u \mapsto I'(u) \in X')$ is Lipschitz continuous on bounded sets; that is, the derivative is *locally Lipschitz*. We state the following fundamental result in the setting of Hilbert spaces, but we note that it can be generalized to Banach spaces by the use of the *pseudo-gradient vector field* (see [Rabinowitz et al., 1986]). Also recall that in Hilbert spaces, the derivative $I'(u) \in X'$ can be identified with the gradient $\nabla I(u) \in X$.

Theorem 3.3.1 (Deformation Lemma). *Assume that X is a Hilbert space, $a, b \in \mathbb{R}$ with $a < b$ and let $I \in C^{1,1}(X)$ be a functional. Assume that for every $c \in [a, b]$ there is no PS sequence for I at level c; that is, for any $c \in [a, b]$ there is no $(PS)_c$ level for I. Then I^b is deformable in I^a.*

Remark 3.3.1. In other words, the only obstruction to deformations is the presence of PS sequences.

Proof. First note that there exists $\sigma > 0$ such that $||\nabla I(u)|| \geq \sigma$ for all $u \in I^b \backslash I^a$. If not, we can otherwise form a sequence $(u_n) \subseteq I^b \backslash I^a$ such that $||\nabla I(u_n)|| \leq \frac{1}{n}$ for every $n \in \mathbb{N}^+$. Since $a < I(u_n) \leq b$ for every $n \in \mathbb{N}^+$, we can find a subsequence (u_{n_k}) and $c \in [a, b]$ such that

$$||\nabla I(u_{n_k})|| \to 0 \text{ and } I(u_{n_k}) \to c.$$

This is a contradiction since we assumed that $[a, b]$ contains no PS levels for I.
Next we define $g : [0, \infty) \to \mathbb{R}$ by

$$g(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq 1 \\ \frac{1}{t} & \text{if } t > 1. \end{cases}$$

For each $u \in X$ consider the initial value problem

$$\begin{cases} \eta'(t, u) := \frac{d}{dt}\eta(t, u) = -\gamma g(||\nabla I(\eta(t, u))||)\nabla I(\eta(t, u)) & \text{if } t > 0 \\ \eta(0, u) = u \end{cases} \tag{3.17}$$

where $\gamma > 0$ is a constant to be determined. Since the right hand side of (3.17) is locally Lipschitz, the problem has a unique solution η (see [Brezis, 2010]). We now check that η does indeed satisfy the conditions of a deformation.
First note that η is continuous with $\eta(0, u) = u$ for each $u \in X$, and

$$\frac{d}{dt} I(\eta(t, u)) = (\nabla I(\eta(t, u)), \eta'(t, u)) = -\gamma g(||\nabla I(\eta(t, u))||)||\nabla I(\eta(t, u))||^2 \leq 0.$$

Hence if $u \in I^a$, then $\eta(t, u) \in I^a$ for every $t \in [0, 1] \subseteq [0, \infty)$. Now

$$I(\eta(1, u)) - I(\eta(0, u)) = \int_0^1 \frac{d}{dt} I(\eta(t, u))\, dt = \int_0^1 -\gamma g(||\nabla I(\eta(t, u))||)||\nabla I(\eta(t, u))||^2\, dt \leq -\gamma\sigma^2.$$

So if $u \in I^b \backslash I^a$, then

$$I(\eta(1, u)) \leq I(u) - \gamma\sigma^2 \leq b - \gamma\sigma^2.$$

Thus if we pick γ so that $\gamma > \frac{b-a}{\sigma^2} > 0$ we get

$$I(\eta(1, u)) \leq a \implies \eta(1, u) \in I^a$$

as required. ∎

For applications, it is usually better to use deformations with additional properties.

Definition 3.3.3. *Let $A \subseteq B \subseteq X$ be subsets and η be a deformation of B in A. We say that η* fixes *a subset $C \subseteq A$ if $\eta(t, u) = u$ for every $t \in [0, 1]$ and $u \in C$.*

Corollary 3.3.1. *If X is a Hilbert space and c is not a PS level for I, then there exists a sufficiently small $\epsilon_0 > 0$ such that for every positive $\epsilon < \epsilon_0$, $I^{c+\epsilon}$ is deformable in $I^{c-\epsilon}$ by a deformation that fixes $I^{c-2\epsilon}$.*

Proof. The idea of the proof is to first show that there exists $\epsilon_0 > 0$ such that $[c-\epsilon_0, c+\epsilon_0]$ contains no PS levels for I, and then apply a similar but modified argument to the above proof to get the desired deformation.
Assume, for a contradiction, that for every $\epsilon > 0$, $[c-\epsilon, c+\epsilon]$ contains a PS level for I. This implies that for each positive integer n, we can find $c_n \in [c - \frac{1}{2^n}, c + \frac{1}{2^n}]$ and a sequence $(u^n_m)_{m=1}^\infty$ such that

$$I(u^n_m) \to c_n \text{ and } I'(u^n_m) \to 0 \text{ as } m \to \infty.$$

Now pick the sequence (v_n) as follows. Let $N_1 \in \mathbb{N}^+$ be such that $m \geq N_1 \implies |I(u^1_m) - c_1| < \frac{1}{2}$ and $||I'(u^1_m)|| < 1$. Set $v_1 = u^1_{N_1}$. Again let $N_2 \geq N_1 + 1$ be such that $m \geq N_2 \implies |I(u^2_m) - c_2| < \frac{1}{4}$ and $||I'(u^2_m)|| < \frac{1}{2}$. Set $v_2 = u^2_{N_2}$. Continuing this way, for each positive $n \geq 3$, we inductively choose $N_n \geq N_{n-1} + 1$ such that $m \geq N_n \implies |I(u^n_m) - c_n| < \frac{1}{2^n}$ and $||I'(u^n_m)|| < \frac{1}{2^{n-1}}$, and set $v_n = u^n_{N_n}$. So for every positive integer n we have

$$|I(v_n) - c| \leq |I(v_n) - c_n| + |c_n - c| \leq \frac{1}{2^n} + \frac{1}{2^n} = \frac{1}{2^{n-1}} \text{ and } ||I'(v_n)|| < \frac{1}{2^{n-1}}.$$

Hence (v_n) is a PS sequence at level c, a contradiction. Thus there exists $\epsilon_0 > 0$ such that $[c-\epsilon_0, c+\epsilon_0]$ contains no PS levels for I. By the argument of the previous theorem, we can find $\sigma > 0$ such that $||I'(u)|| \geq \sigma$ for every $u \in I^{c+\epsilon_0} \backslash I^{c-\epsilon_0}$. Now let $\epsilon \in (0, \epsilon_0)$ and define the cut-off function $h : \mathbb{R} \to \mathbb{R}$ by

$$h(t) := \frac{\text{dist}(t, (-\infty, c-2\epsilon])}{\text{dist}(t, (-\infty, c-2\epsilon]) + \text{dist}(t, [c-\epsilon, c+\epsilon])}, \quad t \in \mathbb{R},$$

where $\text{dist}(t, A) := \inf \{|t - x| : x \in A\}$ for $A \subseteq \mathbb{R}$. Note that h is continuous and satisfies the following:

$$0 \leq h \leq 1, \; h((-\infty, c-2\epsilon]) = \{0\} \text{ and } h([c-\epsilon, c+\epsilon]) = \{1\}.$$

Next we consider the following initial value problem:

$$\begin{cases} \eta'(t,u) = -\gamma h(I(\eta(t,u)))g(||\nabla I(\eta(t,u))||)\nabla I(\eta(t,u)) & \text{if } t > 0 \\ \eta(0,u) = u \end{cases} \tag{3.18}$$

where $\gamma > 0$ is a constant to be determined. We again use the fact that the right hand side is locally Lipschitz to conclude that there exists a unique solution to the ODE. If $u \in I^{c-2\epsilon}$, then $h(I(u)) = 0$, hence $\eta(t,u) = u$ for all $t > 0$. Thus this deformation fixes $I^{c-2\epsilon}$. The rest of the properties follow from the previous theorem since $h(I(u)) = 1$ for $u \in I^{c+\epsilon} \backslash I^{c-\epsilon}$. ■

We can finally show a direct application to critical points.

Theorem 3.3.2 (Mountain Pass Theorem, [Ambrosetti and Rabinowitz, 1973])**.** *Let X be a Hilbert space and $I \in C^{1,1}(X)$ with $I(0) = 0$. Assume that there exist positive numbers α and ρ such that*

MP1 *$I(u) \geq \alpha$ if $||u|| = \rho$*

MP2 *There exists $v \in X$ such that $||v|| > \rho$ and $I(v) \leq 0$.*

Define

$$c := \inf_{\gamma \in \Gamma} \sup_{t \in [0,1]} I(\gamma(t))$$

where $\Gamma := \{\gamma \in C([0,1], X) : \gamma(0) = 0, \gamma(1) = v\}$ is the set of continuous paths connecting 0 to v. Then

1. *$c \geq \alpha$*

2. *c is a PS level for I*

3. *If I satisfies (PS) or $(PS)_c$, then c is a critical level for I.*

Proof. To prove assertion 1, note that for any $\gamma \in \Gamma$, the function $||\cdot|| \circ \gamma : [0,1] \to [0,\infty)$ is continuous with $||\gamma(0)|| = ||0|| = 0$ and $||\gamma(1)|| = ||v|| > \rho$, hence by the intermediate value theorem, there exists $t_0^\gamma \in (0,1)$ such that $||\gamma(t_0^\gamma)|| = \rho$, which implies that $I(\gamma(t_0^\gamma)) \geq \alpha$. Hence

$$c = \inf_{\gamma\in\Gamma} \sup_{t\in[0,1]} I(\gamma(t)) \geq \inf_{\gamma\in\Gamma} I(\gamma(t_0^\gamma)) \geq \alpha.$$

Assertion 3 will easily follow from Proposition 3.3.1 and assertion 2, so we focus on proving assertion 2. Assume that c is not a PS level for I. Then by Corollary 3.3.1 we can find a sufficiently small $\epsilon_0 > 0$ such that for every positive $\epsilon < \epsilon_0$, $I^{c+\epsilon}$ is deformable in $I^{c-\epsilon}$ by a deformation that fixes $I^{c-2\epsilon}$. Choose $\epsilon > 0$ such that $0 < \epsilon < \min\left\{\epsilon_0, \frac{\alpha}{2}\right\}$ and denote the corresponding deformation by η. By definition of c, we can find $\gamma^\epsilon \in \Gamma$ such that

$$c \leq \sup_{t\in[0,1]} I(\gamma^\epsilon(t)) < c + \epsilon. \tag{3.19}$$

Now define $\tilde{\gamma} := \eta(1,\cdot) \circ \gamma^\epsilon$. Note that $\tilde{\gamma}(0) = \eta(1, \gamma^\epsilon(0)) = \eta(1,0) = 0$ since $I(0) = 0 \leq c - \alpha < c - 2\epsilon$ (i.e. $0 \in I^{c-2\epsilon}$) and η fixes $I^{c-2\epsilon}$. Also $\tilde{\gamma}(1) = \eta(1, \gamma^\epsilon(1)) = \eta(1, v) = v$ for the same reason. Since $\tilde{\gamma}$ is continuous, we therefore have that $\tilde{\gamma} \in \Gamma$. Finally, (3.19) implies that $\gamma^\epsilon(t) \in I^{c+\epsilon}$ for every $t \in [0,1]$, hence $\tilde{\gamma}(t) = \eta(1, \gamma^\epsilon(t)) \in I^{c-\epsilon}$ for all $0 \leq t \leq 1$. This implies that

$$\sup_{t\in[0,1]} I(\gamma(\tilde{t})) \leq c - \epsilon,$$

a contradiction since c is the greatest lower bound. Thus c is a PS level for I.

■

3.3.2 Application

We now apply the results of the previous section to prove the existence of non-trivial solutions to the equation

$$\begin{cases} -\Delta u = |u|^{p-2}u & x \in \Omega \\ u = 0 & x \in \partial\Omega \end{cases} \tag{3.20}$$

where Ω is a bounded domain with smooth boundary $\partial\Omega$ and $2 < p < 2^*$. Solutions to (3.20) are critical points of the functional $I : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$I(u) := \frac{1}{2}\int_\Omega |\nabla u|^2\,dx - \frac{1}{p}\int_\Omega |u|^p\,dx,\ \ u \in H_0^1(\Omega).$$

We will apply the Mountain Pass Theorem (MPT) to I on the Hilbert space $H_0^1(\Omega)$ with inner product and norm given by

$$(u,v) = \int_\Omega \nabla u \cdot \nabla v\,dx \text{ and } ||u||^2 = \int_\Omega |\nabla u|^2\,dx,\ \ u,v \in H_0^1(\Omega).$$

The next results verify that I satisfies the conditions of MPT. First we derive some inequalities that will be useful here. Let $h(r) = |r|^{p-2} r$ for $r \in \mathbb{R}$. Note that since $p > 2$, $h \in C^1(\mathbb{R}, \mathbb{R})$ with $h'(r) = (p-1)\,|r|^{p-2}$. If $s < t$ then there exists $r_0 \in (t, s)$ such that

$$h(t) - h(s) = h'(r_0)(t - s)$$

by the mean value theorem. Hence we have

$$\left||t|^{p-2}\,t - |s|^{p-2}\,s\right| = \left|(p-1)\,|r_0|^{p-2}\,(t-s)\right| \leq (p-1)\max\left\{|t|^{p-2}, |s|^{p-2}\right\}|t-s|$$

$$\leq C\left(|t|^{p-2} + |s|^{p-2}\right)|t-s|$$

for some constant $C > 0$. The same holds when $s > t$, so we have

$$\left||t|^{p-2}\,t - |s|^{p-2}\,s\right| \leq C\left(|t|^{p-2} + |s|^{p-2}\right)|t-s| \text{ for all } s, t \in \mathbb{R},$$

where C is a constant dependent only on p.
Next let $q > 0$ and a, b be positive real numbers. If $a < b$ then

$$(a+b)^q = b^q\left(1 + \frac{a}{b}\right)^q \leq b^q 2^q$$

since $0 < a/b < 1$. Arguing similarly when $a > b$, we get

$$(a+b)^q \leq 2^q \max\{a^q, b^q\} \leq 2^q\,(a^q + b^q) \text{ for any } a, b \in [0, \infty) \text{ and } q > 0. \tag{3.21}$$

Lemma 3.3.1. *The functional I is continuously differentiable on $H_0^1(\Omega)$ and its derivative is locally Lipschitz continuous. That is, $I \in C^{1,1}(H_0^1(\Omega), \mathbb{R})$.*

Proof. We first write $I = I_1 - I_2$, where

$$I_1(u) = \frac{1}{2}\int_\Omega |\nabla u|^2\,dx = \frac{1}{2}\,||u||^2 \text{ and } I_2(u) = \frac{1}{p}\int_\Omega |u|^p\,dx,\ u \in H_0^1(\Omega).$$

Note that

$$I_1'(u)(w) = \int_\Omega \nabla u \cdot \nabla w\,dx \text{ and } I_2'(u)(w) = \int_\Omega |u|^{p-2} uw\,dx, \text{ for every } u, w \in H_0^1(\Omega).$$

We will be done if we prove that both I_1 and I_2 are continuously differentiable with locally Lipschitz continuous derivatives.
We first show that $I_1 \in C^{1,1}(H_0^1(\Omega), \mathbb{R})$. Let $u, v, w \in H_0^1(\Omega)$ with $||w|| \leq 1$. Then

$$|I_1'(u)(w) - I_1'(v)(w)| = \left|\int_\Omega (\nabla u - \nabla v)\cdot\nabla w\,dx\right| \leq ||u - v||\,||w|| \leq ||u - v||$$

by Holder's inequality. Thus

$$||I_1'(u) - I_1'(v)|| = \sup_{||w|| \leq 1} |I_1'(u)(w) - I_1'(v)(w)| \leq ||u - v||$$

for every $u, v \in H_0^1(\Omega)$, implying that I_1' is Lipschitz — and therefore locally Lipschitz. This, of course, also implies that I_1 is continuously differentiable on $H_0^1(\Omega)$.

Now we show that $I_2 \in C^{1,1}(H_0^1(\Omega), \mathbb{R})$. Let $R > 0$ and $u, v, w \in H_0^1(\Omega)$ with $||u||, ||v|| \leq R$ and $||w|| \leq 1$. Then applying Holder's inequality with exponents p and $\frac{p}{p-1}$ gives

$$|I_2'(u)(w) - I_2'(v)(w)| = \left| \int_\Omega \left(|u|^{p-2} u - |v|^{p-2} v\right) w \, dx \right|$$

$$\leq \left(\int_\Omega \left(|u|^{p-2} u - |v|^{p-2} v\right)^{\frac{p}{p-1}} dx \right)^{\frac{p-1}{p}} \left(\int_\Omega |w|^p \, dx \right)^{\frac{1}{p}}$$

$$\leq C \left(\int_\Omega \left(|u|^{p-2} u - |v|^{p-2} v\right)^{\frac{p}{p-1}} dx \right)^{\frac{p-1}{p}}$$

by the Sobolev embedding of $H_0^1(\Omega)$ into $L^p(\Omega)$. We now apply the inequality above to get

$$\left(\int_\Omega \left(|u|^{p-2} u - |v|^{p-2} v\right)^{\frac{p}{p-1}} dx \right)^{\frac{p-1}{p}} \leq C \left(\int_\Omega \left(|u|^{p-2} + |v|^{p-2}\right)^{\frac{p}{p-1}} |u - v|^{\frac{p}{p-1}} dx \right)^{\frac{p-1}{p}}$$

which upon the application of Holder's inequality with exponents $p-1$ and $\frac{p-1}{p-2}$ gives

$$\left(\int_\Omega \left(|u|^{p-2} u - |v|^{p-2} v\right)^{\frac{p}{p-1}} dx \right)^{\frac{p-1}{p}} \leq C \left(\left(\int_\Omega \left(|u|^{p-2} + |v|^{p-2}\right)^{\frac{p}{p-2}} dx \right)^{\frac{p-2}{p-1}} \left(\int_\Omega |u - v|^p \, dx \right)^{\frac{1}{p-1}} \right)^{\frac{p-1}{p}}$$

$$= C \left(\int_\Omega \left(|u|^{p-2} + |v|^{p-2}\right)^{\frac{p}{p-2}} dx \right)^{\frac{p-2}{p}} \left(\int_\Omega |u - v|^p \, dx \right)^{\frac{1}{p}}$$

$$\leq C \left(\int_\Omega \left(|u|^p + |v|^p\right) dx \right)^{\frac{p-2}{p}} \left(\int_\Omega |u - v|^p \, dx \right)^{\frac{1}{p}}.$$

Applying the Sobolev inequality and the fact that the norms of u and v are bounded above by R we get

$$||I_2'(u) - I_2'(v)|| = \sup_{||w|| \leq 1} |I_2'(u)(w) - I_2'(v)(w)| \leq \left(\int_\Omega \left(|u|^{p-2} u - |v|^{p-2} v\right)^{\frac{p}{p-1}} dx \right)^{\frac{p-1}{p}} \leq C \, ||u - v||,$$

where C depends on R. ■

Lemma 3.3.2. *I satisfies PS.*

Proof. Let (u_n) be a PS sequence in $H_0^1(\Omega)$. That is

$$I'(u_n) \to 0 \text{ and } |I(u_n)| \leq C \ \forall n \in \mathbb{N}^+ \text{ for some constant } C > 0.$$

This implies that, in particular, for some constant $C > 0$, $I(u_n) \leq C$ and $|I'(u_n)(u_n)| \leq C||u_n||$ for every n. Hence

$$C(1 + ||u_n||) \geq I(u_n) - \frac{1}{p} I'(u_n)(u_n) = \left(\frac{1}{2} - \frac{1}{p} \right) ||u_n||^2 .$$

So (u_n) is (norm) bounded in $H_0^1(\Omega)$. Thus, there exists $u \in H_0^1(\Omega)$ such that, up to a subsequence (still denoted by (u_n)), $u_n \rightharpoonup u$ weakly in $H_0^1(\Omega)$ and $u_n \to u$ strongly in $L^p(\Omega)$ ($p < 2^*$, so the embedding is compact). Now

$$|I'(u_n)(u_n - u)| \leq ||I'(u_n)|| \, ||u_n - u|| \leq C \, ||I'(u_n)|| \to 0,$$

hence

$$o(1) = I'(u_n)(u_n - u) = \int_\Omega \nabla u_n \cdot \nabla u_n - \nabla u_n \cdot \nabla u \, dx - \int_\Omega |u_n|^{p-2} u_n(u_n - u) \, dx$$
$$= ||u_n||^2 - \int_\Omega \nabla u_n \cdot \nabla u - \int_\Omega |u_n|^p \, dx + \int_\Omega |u_n|^{p-2} u_n u \, dx.$$

Now by Theorem 2.2.3 and considering a further subsequence, there exists $w \in L^p(\Omega)$ such that

$$u_n(x) \to u(x) \text{ and } |u_n(x)| \leq w(x) \text{ a.e. } x \in \Omega.$$

Thus we have $\left||u_n|^{p-2} u_n u\right| \leq w^{p-1} |u| \in L^1(\Omega)$, since by Holder's inequality

$$\int_\Omega w^{p-1} |u| \, dx \leq \left(\int_\Omega w^p \, dx\right)^{\frac{p-1}{p}} \left(\int_\Omega |u|^p\right)^{\frac{1}{p}} < \infty.$$

Thus we can apply the dominated convergence theorem to get

$$\lim_{n\to\infty} \int_\Omega |u_n|^{p-2} u_n u \, dx = \int_\Omega |u|^p \, dx.$$

Since $u_n \rightharpoonup u$ weakly in $H_0^1(\Omega)$ and $u_n \to u$ strongly in $L^p(\Omega)$ we have

$$\lim_{n\to\infty} \int_\Omega \nabla u_n \cdot \nabla u \, dx = \int_\Omega |\nabla u|^2 \, dx = ||u||^2 \text{ and } \lim_{n\to\infty} \int_\Omega |u_n|^p \, dx = \int_\Omega |u|^p \, dx.$$

Hence $||u_n||^2 \to ||u||^2$, which implies that, by Theorem 2.2.5, $u_n \to u$ strongly in $H_0^1(\Omega)$ as required. ■

We can finally apply the Mountain Pass theorem to establish the existence of a non-trivial critical point of I.

Theorem 3.3.3. *The functional I has at least one non-trivial critical point, which is also a weak solution to (3.20).*

Proof. We apply the Mountain Pass Theorem. We have shown that $I \in C^{1,1}(H_0^1(\Omega), \mathbb{R})$ and that I satisfies (PS). Clearly $I(0) = 0$ and

$$I(u) = \frac{1}{2} ||u||^2 - \frac{1}{p} ||u||_p^p \geq \frac{1}{2} ||u||^2 - C ||u||^p$$

for some $C > 0$. Now if $||u|| = \rho$, then $I(u) \geq \frac{1}{2}\rho^2 - C\rho^p$, which is positive for sufficiently small ρ since $p > 2$. On the other hand, if we fix $u \neq 0$, then $I(tu) = \frac{t^2}{2} ||u||^2 - \frac{t^p}{p} ||u||_p^p \to -\infty$ as $t \to \infty$. Thus, there exists $||v|| > \rho$ such that $I(v) \leq 0$. Hence $c > 0$ is a PS level for I, where c is defined by

$$c := \inf_{\gamma\in\Gamma} \sup_{t\in[0,1]} I(\gamma(t))$$

and $\Gamma := \{\gamma \in C([0,1], X) : \gamma(0) = 0, \gamma(1) = v\}$ is the set of continuous paths connecting 0 to v. Also, since I satisfies (PS), c is a critical value of I with critical point $u \neq 0$ ($I(u) = c > 0$). ■

Remark 3.3.2. The proof that I satisfies (PS) relies crucially on the fact that the embedding of $H_0^1(\Omega)$ into $L^p(\Omega)$ is compact since $p < 2^*$. In the next chapter, we will see that I fails to satisfy the (PS) condition when $p = 2^*$, and that there are cases where the PDE has no (non-trivial) solutions.

3.4 Constrained Minimization on Spheres

As a final application of variational methods, we consider constrained minimization. We provide an alternative to the previous section for dealing with cases when the functional I is not coercive, by introducing artificial constraints that ensure that I is coercive on the constrained subset.

Together with the method of *the Nehari manifold* (see [Nehari, 1960]), constrained minimization on spheres is one of the most widely used constrained optimization techniques to solve non-coercive PDE problems. We have chosen this method mainly for its vast applications to solving more complicated problems that we study in the next chapter.

Here we illustrate the method by solving problem (3.1) in the special case when $f(x,u) = |u|^{p-2}u$. We have seen that we cannot let p be arbitrary, so we again make the following assumption:

$$2 < p < 2^* = \frac{2N}{N-2}. \tag{3.22}$$

Note that the case when $1 < p \leq 2$ has been covered by results from Section 3.2.3. We want to find non-trivial weak solutions to the following:

$$\begin{cases} -\Delta u + cu = |u|^{p-2}u & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.23}$$

under assumption (3.22). We still assume that $c \geq 0, c \in L^\infty(\Omega)$ and that Ω is a bounded domain with smooth boundary. A fucntion $u \in H_0^1(\Omega)$ is a weak solution to (3.23) if

$$B_c(u,v) = \int_\Omega \nabla u \cdot \nabla v + cuv\, dx = \int_\Omega |u|^{p-2}\, uv\, dx \text{ for every } v \in H_0^1(\Omega).$$

We have also shown that weak solutions to (3.23) correspond to critical points of the functional $I : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$\begin{aligned} I(u) &:= \frac{1}{2}\int_\Omega |\nabla u|^2\, dx + \frac{1}{2}\int_\Omega cu^2\, dx - \frac{1}{p}\int_\Omega |u|^p\, dx,\ u \in H_0^1(\Omega). \\ &= \frac{1}{2}\,||u||_c^2 - \frac{1}{p}\,||u||_p^p. \end{aligned}$$

However, as pointed out before, this functional is not coercive due to the growth of $f(x,u) = |u|^{p-2}u$ being super-linear (since $p > 2$). In the previous section we overcame this problem by applying MPT. In this section we will instead prove that if we constrain I on an appropriate sphere, we recover coercivity, and a positive scalar multiple of the constrained minimizer of I on that sphere is a solution.

Define

$$\mathcal{A}_p := \left\{u \in H_0^1(\Omega) : ||u||_p = 1\right\}.$$

Theorem 3.4.1. *If (3.22) holds, then*

$$\mu_p := \inf_{v \in \mathcal{A}_p} I(v)$$

is attained by some $u \in \mathcal{A}_p$ such that $u \geq 0$ on Ω.

Proof. First note that if $v \in \mathcal{A}_p$, then $I(v) = \frac{1}{2}||v||_c^2 - \frac{1}{p}$. Now let $(u_n) \subseteq \mathcal{A}_p$ be a minimizing sequence; that is,

$$\frac{1}{2}||u_n||_c^2 - \frac{1}{p} = \mu_p + o(1).$$

Clearly (u_n) is bounded in norm, hence due to reflexivity, we can assume that $u_n \rightharpoonup u$ weakly in $H_0^1(\Omega)$ for some $u \in H_0^1(\Omega)$. Since $p < 2^*$, $H_0^1(\Omega)$ is compactly embedded into $L^p(\Omega)$, thus $u \to u$ strongly in $L^p(\Omega)$. Hence due to continuity of the norm, we have $||u||_p = \lim_{n\to\infty} ||u_n||_p = 1$, which implies that $u \in \mathcal{A}_p$; i.e. $\mathcal{A}_p$ is weakly sequentially closed. We finally apply lower semicontinuity of $||\cdot||_c^2$ to get

$$I(u) = \frac{1}{2}||u||_c^2 - \frac{1}{p} \leq \liminf_{n\to\infty} \left(\frac{1}{2}||u_n||_c^2 - \frac{1}{p}\right) = \mu_p.$$

Hence u is a minimizer of I on $\mathcal{A}_p$. Noting that $u \in \mathcal{A}_p \implies |u| \in \mathcal{A}_p$ and $I(u) = I(|u|)$, we can assume that $u \geq 0$ (otherwise we can replace u with $|u|$). Also, u is non-trivial since $||u||_p = 1$. ■

Unlike before, the minimizer u found above is not a critical point of I — and therefore not a solution to (3.23)— since $\mathcal{A}_p$ is not a vector space (so the derivative of I does not vanish at u). However, a scalar multiple of u provides a solution to (3.23).

Theorem 3.4.2. *Under assumption (3.22), the PDE (3.23) has a non-trivial solution $w \in H_0^1(\Omega)$ such that $w > 0$ in Ω.*

Proof. Define $J : H_0^1(\Omega) \to \mathbb{R}$ by $J(u) = ||v||_p^p - 1$ for any $v \in H_0^1(\Omega)$, so that $\mathcal{A}_p = \{v \in H_0^1(\Omega) : J(v) = 0\}$. We know from a slight modification of Lemma 3.3.1 that both I and J belong to $C^1(H_0^1(\Omega))$, hence by Theorem 2.2.12, there exists a Lagrange multiplier $\lambda \in \mathbb{R}$ such that

$$I'(u)(v) = \lambda J'(u)(v) \text{ for all } v \in H_0^1(\Omega), \tag{3.24}$$

where u is the non-trivial minimizer from the above theorem. Equation (3.24) reads

$$B_c(u, v) = \lambda \int_\Omega |u|^{p-2} uv \, dx \text{ for all } v \in H_0^1(\Omega).$$

Letting $v = u$, we see that $\lambda = \frac{||u||_c^2}{||u||_p^p} = ||u||_c^2$. If we let $w = (\lambda)^{\frac{1}{p-2}} u = (||u||_c)^{\frac{2}{p-2}} u$, then w satisfies

$$B_c(w, v) = B_c((||u||_c)^{\frac{2}{p-2}} u, v) = (||u||_c)^{\frac{2}{p-2}} B_c(u, v) = (||u||_c)^{\frac{2}{p-2}} ||u||_c^2 \int_\Omega |u|^{p-2} uv \, dx$$

$$= \int_\Omega \left|(||u||_c)^{\frac{2}{p-2}} u\right|^{p-2} (||u||_c)^{\frac{2}{p-2}} uv \, dx = \int_\Omega |w|^{p-2} wv \, dx \ \forall v \in H_0^1(\Omega).$$

Hence w, which is a positive scalar multiple of u, is a non-negative solution to (3.23). By the strong maximum principle, $w > 0$ on Ω.

■

Remark 3.4.1. This result implies that w satisfies

$$\begin{cases} -\Delta w + cw = w^{p-1} & \text{in } \Omega \\ w > 0 & \text{in } \Omega \\ w = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.25}$$

Positive solutions to this equation will be investigated again in the next section for $p = 2^*$.

Remark 3.4.2. Again, the previous result highlights the importance of the exponent p. In this case, we used the compactness of the embedding of $H_0^1(\Omega)$ into $L^p(\Omega)$ when $p < 2^*$ to prove that the set $\mathcal{A}_p$ is weakly sequentially closed in $H_0^1(\Omega)$. The failure of this technique when $p = 2^*$ will form part of the subject of the next chapter.

Chapter 4

The Critical Sobolev Exponent

In this chapter we tackle the problem of solving equations of the form

$$\begin{cases} -\Delta u = u^{p-1} & x \in \Omega \\ u > 0 & x \in \Omega \\ u = 0 & x \in \partial\Omega \end{cases} \tag{4.1}$$

when $p = 2^* = \frac{2N}{N-2}$ is the *critical Sobolev exponent* of the embedding of $H_0^1(\Omega)$ into $L^p(\Omega)$. We will also consider variants of (4.1) of the form:

$$\begin{cases} -\Delta u = |u|^{p-2}u + f(x,u) & x \in \Omega \\ u = 0 & x \in \partial\Omega \end{cases} \tag{4.2}$$

for some appropriate function $f : \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}$. In Chapter 1, we remarked that solutions to (4.1) are related to the famous Yamabe problem, and [Brézis and Nirenberg, 1983] elaborate on this.

In the previous chapter, we managed to solve (4.1) and its variants when $p < 2^*$, and the proof crucially depends on this fact. Solutions to (4.1) are *positive* critical points of the functional $I : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$I(u) := \frac{1}{2}\int_\Omega |\nabla u|^2 \, dx - \frac{1}{p}\int_\Omega |u|^p \, dx, \; u \in H_0^1(\Omega).$$

A first attempt at proving that I has a positive critical point would be to apply the MPT. It can be shown just like in Lemma 3.3.1 and Theorem 3.3.3 that I satisfies all the conditions of MPT except (PS). Indeed, we cannot simply follow the proof of Lemma 3.3.2 in proving that I satisfies (PS), since one of the key steps in that proof uses the fact that any (PS) sequence has a weakly convergent subsequence in $H_0^1(\Omega)$, and therefore strongly convergent in $L^p(\Omega)$. The strong convergence of this subsequence in $L^p(\Omega)$ is a direct application of Proposition 2.2.2 and Theorem 2.3.8, which only holds when $p < 2^*$. So this technique cannot be used as yet to solve (4.1) for $p = 2^*$.

Another approach that is also used in Section 3.4 is that of constrained minimization. That is, we define the sphere

$$\mathcal{A}_p := \left\{ u \in H_0^1(\Omega) : ||u||_p = 1 \right\}$$

and then minimize I on $\mathcal{A}_p$. We would then conclude as in Theorem 3.4.2 that a positive multiple of this minimizer is a solution. Since I is coercive on $\mathcal{A}_p$, any minimizing sequence $(u_n) \subseteq \mathcal{A}_p$ is norm-bounded in $H_0^1(\Omega)$, and hence weakly converges, up to a subsequence, to some $u \in H_0^1(\Omega)$. Also, I being sequentially weakly lsc gives

$$I(u) \leq \liminf_{n\to\infty} I(u_n) = \inf_{v\in\mathcal{A}_p} I(v).$$

In Theorem 3.4.1 the proof is completed by using the fact that (u_n) converges strongly to u in $L^p(\Omega)$, thus giving

$$||u||_p = \lim_{n\to\infty} ||u_n||_p = 1 \implies u \in \mathcal{A}_p.$$

Thus, the conclusion that $\mathcal{A}_p$ is sequentially weakly closed also uses the compactness of the embedding of $H_0^1(\Omega)$ into $L^p(\Omega)$ when $p < 2^*$. This again makes the method inadequate to solving (4.1) when $p = 2^*$.

In this chapter we will discuss how these methods can be modified to solve (4.1) and its variants in some special cases. We also emphasize that the difficulties in applying variational techniques to solving (4.1) in this special case are not just due to deficiencies of variational techniques, but that there are cases where no solution exists.

4.1 Dependence on the Domain

We first show that the existence or non-existence of solutions to (4.1) is dependent on the domain Ω. That is, we show that (4.1) has a solution for some domains Ω and no solutions on other domains. We begin with the negative result.

4.1.1 Pohozaev's Non-existence Result

We present a well-known result due to [Pohozhaev, 1965] about the non-existence of a smooth positive solution to (4.1) for a bounded *star-shaped* domain Ω. We will assume that $0 \in \Omega$ throughout this subsection.

Definition 4.1.1. *An open set Ω is called* star-shaped *with respect to $0 \in \mathbb{R}^N$ if for each $x \in \bar{\Omega}$, the line segment*

$$\{\alpha x : \alpha \in [0, 1]\}$$

is contained in $\bar{\Omega}$.

The following well known result highlights a useful property of star-shaped domains. Its proof can be found in [Evans, 1998].

Lemma 4.1.1. *If $\partial\Omega$ is C^1 and Ω is star-shaped with respect to 0, then*

$$x \cdot \mathbf{n}(x) \geq 0 \text{ for all } x \in \partial\Omega,$$

where $\mathbf{n}(x)$ is the outward unit normal at x.

Another useful result is by [Pohozhaev, 1965], which we state below.

Theorem 4.1.1 (Pohozaev's Identity)**.** *Let Ω be a bounded smooth domain and $u \in C^2(\Omega)$ be a positive solution of*

$$\begin{cases} -\Delta u = g(u) & x \in \Omega \\ u = 0 & x \in \partial\Omega, \end{cases} \tag{4.3}$$

where $g \in C^1(\mathbb{R})$ *with* $g(0) = 0$*. Then*

$$(2-N)\int_\Omega u\,g(u)\,dx + 2N\int_\Omega G(u)\,dx = \int_{\partial\Omega} |\nabla u|^2 (x\cdot \mathbf{n}(x))\,d\sigma(x) \tag{4.4}$$

where

$$G(t) = \int_0^t g(s)\,ds.$$

Equation (4.4) is known as the *Pohozaev Identity.* Applying this identity to $g(u) = u^{p-1}$ we get:

Theorem 4.1.2. *If* Ω *is a bounded smooth star-shaped domain (with respect to 0), then (4.1) has no solutions for* $p \geq 2^*$*.*

Proof. Assume that u is a solution of (4.1) and let $g(u) = u^{p-1}$. Then (4.4) can be written as

$$\int_\Omega \left((2-N) + \frac{2N}{p}\right) u^p\,dx = \int_{\partial\Omega} |\nabla u|^2 (x\cdot \mathbf{n}(x))\,d\sigma(x) \geq 0.$$

Thus if

$$(2-N) + \frac{2N}{p} \leq 0 \iff p \geq \frac{2N}{N-2} = 2^*,$$

then (4.1) has no solution. ■

From the regularity theory of Theorem 2.3.10, every weak solution of (4.1) is also a classical solution. Thus, the previous result tells us that when Ω is star-shaped, (4.1) has no weak solutions. This is one example of a case where the lack of compactness of the embedding of $H_0^1(\Omega)$ onto $L^{2^*}(\Omega)$ is serious, in the sense that it leads to the non-existence of solutions. The result also highlights the important fact that it is not just the failure of variational methods to identify a solution, but no solution exists.

What about domains that are not star-shaped? In general, this problem is still open (a more detailed discussion can be found in [Brézis and Nirenberg, 1983]). A considerable amount of research has been done on this problem and both positive and negative results have been obtained under different conditions imposed on Ω. One important result is by [Bahri and Coron, 1988] who proved that if, unlike in Theorem 1.1.2, Ω has non-trivial topology, then (4.1) has a solution. Their proof uses results from homology theory. We will present what can be considered a special case of their result, that can be handled by the variational methods we developed in the previous chapter.

4.1.2 A Solution on the Annulus

Here we prove that a solution of the equation

$$\begin{cases} -\Delta u = u^{2^*-1} & \text{in } \Omega \\ u > 0 & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{4.5}$$

exists when Ω is the annulus:

$$\Omega = \left\{x \in \mathbb{R}^N : a < |x| < b\right\} \text{ for } 0 < a < b.$$

This result is stated in [Kazdan and Warner, 1975], but we give a more detailed account here. Solutions to (4.5) correspond to positive critical points of the functional $I : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$I(u) := \frac{1}{2}\int_\Omega |\nabla u|^2 \, dx - \frac{1}{2^*}\int_\Omega |u|^{2^*} \, dx, \; u \in H_0^1(\Omega).$$

Since I is not bounded below on $H_0^1(\Omega)$, one attempt at solving (4.5) is, as in Section 3.2.1, to minimize I on the sphere:

$$\mathcal{A}_{2^*} := \left\{u \in H_0^1(\Omega) : ||u||_{2^*} = 1\right\}. \tag{4.6}$$

However, as discussed at the beginning of this chapter, the non-compactness of the embedding of $H_0^1(\Omega)$ into $L^{2^*}(\Omega)$ means that we cannot conclude that $\mathcal{A}_{2^*}$ is weakly sequentially closed in $H_0^1(\Omega)$. Thus, if a minimizing sequence $(u_n) \subseteq \mathcal{A}_{2^*}$ converges weakly to $u \in H_0^1(\Omega)$, we cannot conclude that $u \in \mathcal{A}$, and thereby getting that u is the constrained minimizer of I on $\mathcal{A}_{2^*}$.

In this section we show that the symmetric structure of Ω allows us to reduce this to a one-dimensional problem — where the embedding is compact — by restricting the search for minimizers to *radially symmetric* functions.

Definition 4.1.2. *A function $u : \Omega \to \mathbb{R}$ is* radially symmetric *or* radial *if there exists a function $g : \mathbb{R} \to \mathbb{R}$ such that $u(x) = g(|x|)$ for every $x \in \Omega$.*

We will use the notation $H_r(\Omega)$ for the class of radially symmetric functions in $H_0^1(\Omega)$. That is,

$$H_r(\Omega) := \left\{u \in H_0^1(\Omega) : u \text{ is radially symmetric}\right\}.$$

The advantage of restricting the search for minimizers to $H_r(\Omega)$ is that there is a one to one correspondence between $H_r(\Omega)$ and the one-dimensional space $H_0^1((a,b))$, and from Theorem 2.3.9, $H_0^1((a,b))$ is compactly embedded into $L^q((a,b))$ for any $q \geq 1$.

If $u \in H_r(\Omega)$, then there exists a (almost everywhere) unique $g : (a,b) \to \mathbb{R}$ such that $u(x) = g(|x|)$ for a.e. $x \in \Omega$. To see this, note that if $g, h : (a,b) \to \mathbb{R}$ are such that $u(x) = g(|x|) = h(|x|)$ λ-a.e., then the null set $\{x \in \Omega : g(|x|) \neq h(|x|)\}$ is mapped to the null set $\{r \in (a,b) : g(r) \neq h(r)\}$ by the Lipschitz continuous function $(x \mapsto |x|, x \in \Omega)$.

Now recall that if $g : \mathbb{R} \to \mathbb{R}$ is integrable, then by polar coordinates we have

$$\int_{\mathbb{R}^N} g(|x|)\, dx = \omega_{N-1}\int_0^\infty g(r) r^{N-1}\, dr,$$

where ω_{N-1} is the area of the $N-1$-dimensional unit sphere. In particular,

$$\int_\Omega g(|x|)\, dx = \omega_{N-1}\int_a^b g(r) r^{N-1}\, dr.$$

Now let $g \in C^2((a,b))$ and set $u(x) = g(|x|)$ for every $x \in \Omega$. Then

$$\frac{\partial u}{\partial x_i} = g'(|x|)\frac{x_i}{|x|} \implies \nabla u(x) = g'(|x|)\frac{x}{|x|} \text{ for } x \in \Omega.$$

Hence $u \in C^1(\Omega)$. In fact, it can be shown that $u \in C^2(\Omega)$ and

$$\Delta u(x) = g''(|x|) + \frac{1}{|x|}g'(|x|), \text{ for } x \in \Omega$$

Hence we will have solved (4.5) if we can solve the one-dimensional equation

$$\begin{cases} -\left(g''(r) + \frac{1}{r}g'(r)\right) = g(r)^{2^*-1} & \text{in } (a,b) \\ g > 0 & \text{in } (a,b) \\ g(a) = g(b) = 0. \end{cases} \tag{4.7}$$

So we now turn our attention to solving the one-dimensional equation (4.7), and we do so by changing the Sobolev norm.

Lemma 4.1.2. *In one dimension, the function* $||\cdot||_{p,w}$: $L^p((a,b)) \to \mathbb{R}$ *($1 \le p < \infty$), defined by*

$$||g||_{p,w} := \left(\omega_{N-1}\int_a^b |g(r)|^p r^{N-1}\,dr\right)^{1/p}, \quad g \in L^p((a,b)),$$

is a norm on $L^p((a,b))$ *that is equivalent to the usual* L^p *norm. Furthermore, when* $p = 2$,

$$(g,h)_{2,w} := \omega_{N-1}\int_a^b g(r)h(r)r^{N-1}\,dr, g,h \in L^2((a,b))$$

defines an inner product on $L^2((a,b))$*, that induces a norm that is equivalent to the usual norm on* $L^2((a,b))$*.*

Proof. The properties of a norm and inner product follow easily, so we only prove equivalence. Note that if $g \in L^p((a,b))$, then

$$a^{N-1}|g(r)|^p \le |g(r)|^p r^{N-1} \le b^{N-1}|g(r)|^p \text{ for every } r \in (a,b)$$

$$\implies \omega_{N-1}a^{N-1}||g||_p^p = \omega_{N-1}\int_a^b a^{N-1}|g(r)|^p\,dr \le \omega_{N-1}\int_a^b |g(r)|^p r^{N-1}\,dr$$

$$\le \omega_{N-1}\int_a^b b^{N-1}|g(r)|^p\,dr = \omega_{N-1}b^{N-1}||g||_p^p.$$

The result follows from taking the pth root. ■

From the above the following result immediately follows.

Corollary 4.1.1. *In one dimension, the function* $||\cdot||_{W^{1,p},w}$: $W^{1,p}((a,b)) \to \mathbb{R}$ *($1 \le p < \infty$), defined by*

$$||g||_{W^{1,p},w} := \left(||g||_{p,w}^p + ||g'||_{p,w}^p\right)^{\frac{1}{p}}, \quad g \in W^{1,p}((a,b)),$$

is a norm on $W^{1,p}((a,b))$*, that is equivalent to the usual* $W^{1,p}$ *norm. Furthermore, when* $p = 2$,

$$(g,h)_{H^1,w} := (g,h)_{2,w} + (g',h')_{2,w}, \;\; g,h \in H^1((a,b))$$

defines an inner product on $H^1((a,b))$ *that induces a norm that is equivalent to the usual norm on* $H^1((a,b))$*.*

Remark 4.1.1. Note that on $H_0^1((a,b))$, by Poincare's inequality,

$$(g,h) \mapsto (g',h')_{2,w}, \;\; g,h \in H_0^1((a,b))$$

is an equivalent inner product.

Now we note that a *positive* function $g \in H_0^1((a.b))$ is a weak solution to (4.7) if

$$(g', h')_{2,w} = \omega_{N-1} \int_a^b g^{2^*-1} h r^{N-1}\, dr, \text{ for all } h \in H_0^1((a,b)).$$

Thus, solutions to (4.7) correspond to positive critical points of $I : H_0^1((a,b)) \to \mathbb{R}$ defined by

$$I(g) := \frac{1}{2}\, ||g'||_{2,w}^2 - \frac{1}{2^*}\, ||g||_{2^*,w}^{2^*},\ g \in H_0^1((a,b)).$$

Since I is not bounded below on $H_0^1((a,b))$, we will instead find a minimizer on the sphere:

$$\mathcal{A}_{2^*} := \left\{ g \in H_0^1((a,b)) : ||g||_{2^*,w} = 1 \right\}.$$

Theorem 4.1.3. *The infimum*

$$\mu_{2^*} := \inf_{g \in \mathcal{A}_{2^*}} I(g)$$

is attained by some $g \in \mathcal{A}_{2^*}$ with $g \geq 0$ on (a,b).

Proof. First note that for $g \in \mathcal{A}_{2^*}$, $I(g) = \frac{1}{2}\, ||g'||_{2,w}^2 - \frac{1}{2^*}$. Now let $(g_n) \subseteq \mathcal{A}_{2^*}$ be a minimizing sequence for I. It follows that (g_n) is bounded on $H_0^1((a,b))$, hence, up to a subsequence, there exists $g \in H_0^1((a,b))$ such that $g_n \rightharpoonup g$ weakly in $H_0^1((a,b))$. Also, since I is weakly sequentially lsc, we have

$$I(g) \leq \liminf_{n \to \infty} I(g_n) = \mu_{2^*}.$$

So all we need to show is that $g \in \mathcal{A}_{2^*}$, i.e. $\mathcal{A}_{2^*}$ is weakly sequentially closed in $H_0^1((a,b))$. Here is the crux of the proof: since we are working in one dimension, we know from Theorem 2.3.9 that $H_0^1((a,b))$ is compactly embedded into $L^q((a,b))$ for any $q \geq 1$. Thus, we also have that $g_n \to g$ strongly in $L^{2^*}((a,b))$. Hence $||g||_{2^*,w} = 1$, i.e. $g \in \mathcal{A}_{2^*}$. Now since $I(g) = I(|g|)$, we can assume that $g \geq 0$. ■

Theorem 4.1.4. *There exists a classical solution $f \in C^2([a,b])$ to (4.7).*

Proof. Let g be the (non-zero, since $g \in \mathcal{A}_{2^*}$) non-negative minimizer found in the previous result. As usual we will find a positive scalar multiple of g that solves (4.7). By the Lagrange multiplier rule, g satisfies the equation

$$(g', h')_{2,w} = \lambda \omega_{N-1} \int_a^b g^{2^*-1} h r^{N-1}\, dr \text{ for every } h \in H_0^1((a,b))$$

for the Lagrange multiplier $\lambda = \frac{||g'||_{2,w}^2}{||g||_{2^*,w}^{2^*}} = ||g'||_{2,w}^2$. Again, for some $c > 0$, $f = cg$ satisfies

$$(f', h')_{2,w} = \omega_{N-1} \int_a^b f^{2^*-1} h r^{N-1}\, dr \text{ for every } h \in H_0^1((a,b)),$$

that is

$$\int_a^b f' h' r^{N-1}\, dr = \int_a^b f^{2^*-1} h r^{N-1}\, dr \text{ for every } h \in H_0^1((a,b)).$$

This implies that

$$\int_a^b \left(f' r^{N-1}\right) h'\, dr = \int_a^b f^{2^*-1} h r^{N-1}\, dr \text{ for every } h \in H_0^1((a,b)) \tag{4.8}$$

Now since all elements of $H_0^1((a,b))$ have continuous representatives (see Theorem 2.3.9), we will always assume that we are working with these continuous representatives. We also note that both r^{N-1} and $\frac{1}{r^{N-1}}$ belong to $C^\infty((a,b))$ since $0 < a < b$. Thus, from (4.8) we see that $f'r^{N-1} \in H^1((a,b))$ with $\left(f'r^{N-1}\right)' = -f^{2^*-1}r^{N-1}$. Thus $f'r^{N-1}$ is continuous, $\implies$ f' is also continuous. That is, $f \in C^1([a,b])$. Also, $f' = \left(f'r^{N-1}\right)\frac{1}{r^{N-1}} \in H^1((a,b)) \implies f \in H^2((a,b))$. So the derivative of $f'r^{N-1}$ is continuous; hence $f'r^{N-1} \in C^1([a,b])$. Again we get that $f' = \left(f'r^{N-1}\right)\frac{1}{r^{N-1}} \in C^1([a,b]) \implies f \in C^2([a,b])$. Hence

$$-f^{2^*-1}r^{N-1} = \left(f'r^{N-1}\right)' = f''r^{N-1} + (N-1)r^{N-2}f'$$
$$\implies -\left(f'' + \frac{N-1}{r}f'\right) = f^{2^*-1}.$$

Again by the (general) maximum principle, $f > 0$ on (a,b). ■

Corollary 4.1.2. *There exists a classical solution to (4.5).*

Proof. Let f be the positive solution to (4.7). Then set $u(x) = f(|x|)$ for every $x \in \Omega$. The chain rule easily shows that $u \in C^2(\bar{\Omega})$ and satisfies (4.5). Note that u is also a weak solution to (4.5). ■

4.2 Brezis-Nirenberg

In the previous section we considered two important special cases of problem (4.1): one for the existence of a solution in an annulus and another one for non-existence in a star-shaped domain. The general problem is still open, and enormous research has been done on the problem and its variants. One of the earliest and arguably most influential papers on the quest for more general existence results was by H. Brezis and L. Nirenberg in [Brézis and Nirenberg, 1983]. In this remarkable paper, they overcame the lack of compactness and, among other results, established the existence of solutions to the perturbed problem

$$-\Delta u = \lambda u + u^{2^*-1}.$$

Their ideas were motivated by the works of [Brézis and Lieb, 1983], [Aubin, 1976] and [Trudinger, 1968]. This section, with some modifications, is dedicated to discussing these results. We study the existence of solutions to the perturbed problem:

$$\begin{cases} -\Delta u = \lambda u + u^{2^*-1} & x \in \Omega \\ u > 0 & x \in \Omega \\ u = 0 & x \in \partial\Omega \end{cases} \tag{4.9}$$

on a bounded domain Ω with smooth boundary. Solutions to (4.9) are positive critical points of the functional $\tilde{I} : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$\tilde{I}(u) := \frac{1}{2}\int_\Omega |\nabla u|^2\,dx - \frac{\lambda}{2}\int_\Omega u^2\,dx - \frac{1}{2^*}\int_\Omega |u|^{2^*}\,dx,\ \ u \in H_0^1(\Omega).$$

Since $\tilde{I}$ is neither bounded above nor below, we can either proceed by applying the MPT to establish the existence of a positive critical point of $\tilde{I}$ (which would correspond to a solution to (4.9)), or minimize $\tilde{I}$ over an appropriate sphere (the solution to (4.9) would then be a scalar multiple of this minimizer). We will go with the former option and follow the presentations in [Brézis and Nirenberg, 1983] and [Naito and Sato, 2012].

Let us first show that if (4.9) is to have a solution, then λ cannot be arbitrary.

Proposition 4.2.1. *If $\lambda \leq 0$ and Ω is smooth and star-shaped with respect to 0, then (4.9) has no solutions.*

Proof. This is another application of the Pohozaev identity to $g(u) = \lambda u + u^{2^*-1}$. ■

At the same time, λ cannot be too large. Throughout this section, λ_1 is the smallest eigenvalue of the Laplacian operator $-\Delta$ on $H_0^1(\Omega)$.

Proposition 4.2.2. *If $\lambda \geq \lambda_1$, then (4.9) has no solutions.*

Proof. Assume that $u \in H_0^1(\Omega)$ is a solution to (4.9) and let φ_1 be the positive normalized ($||\varphi_1||_2 = 1$) eigenfunction corresponding to λ_1. Then

$$\lambda_1 \int_\Omega u\varphi_1 \, dx = \int_\Omega \nabla u \cdot \nabla \varphi_1 \, dx = \lambda \int_\Omega u\varphi_1 \, dx + \int_\Omega u^{2^*-1}\varphi_1 \, dx,$$

hence

$$(\lambda_1 - \lambda) \int_\Omega u\varphi_1 \, dx = \int_\Omega u^{2^*-1}\varphi_1 \, dx > 0$$

since both u and φ_1 are positive. Thus, $\lambda_1 > \lambda$. ■

So, a necessary condition for the existence of a solution to (4.9) is that

$$0 < \lambda < \lambda_1.$$

We assume that this is the case for the remainder of this section.

A decisive role in establishing the existence of a solution to (4.9) is played by the so-called *best constant* in the Sobolev embedding. For $\alpha \in \mathbb{R}$ we define

$$S_\alpha := \inf_{\substack{u \in H_0^1(\Omega) \\ u \neq 0}} \frac{||\nabla u||_2^2 - \alpha \, ||u||_2^2}{||u||_{2^*}^2} \text{ and } S := S_0. \tag{4.10}$$

Here S is the best constant in the Sobolev embedding of $H_0^1(\Omega)$ into $L^{2^*}(\Omega)$, and has the following properties (see [Talenti, 1976] and [Brézis and Nirenberg, 1983] for more details on these properties):

1. S is independent of Ω
2. When $\Omega = \mathbb{R}^N$ the infimum S is attained by functions $u(x) = \gamma\psi(\alpha(x-y))$ for constants $\alpha, \gamma > 0$ and $y \in \mathbb{R}^N$, where ψ is the fundamental solution given by

$$\psi(x) = \frac{1}{\left(a + b\,|x|^2\right)^{\frac{N-2}{2}}}, \quad a, b > 0.$$

3. The infimum S is never attained on a bounded domain Ω.

The third property rules out the possibility of using constrained minimization to solve (4.9) when $\lambda = 0$. It is clear that $S_\lambda \leq S$ since $\lambda > 0$. The main result of this section is the following.

Theorem 4.2.1 ([Brézis and Nirenberg, 1983])**.** *If $S_\lambda < S$, then there exists a solution to (4.9).*

We will prove this result by using the MPT to the slightly modified functional $I : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$I(u) := \frac{1}{2}\int_\Omega |\nabla u|^2 \, dx - \frac{\lambda}{2}\int_\Omega \left(u^+\right)^2 \, dx - \frac{1}{2^*}\int_\Omega \left(u^+\right)^{2^*} \, dx, \; u \in H_0^1(\Omega).$$

As discussed above, the main challenge lies in showing that I satisfies (PS) or $(PS)_c$ for some c. This is due to the exponent 2^* in the third term, which prevents weakly convergent sequences in $H_0^1(\Omega)$ from converging strongly in $L^{2^*}(\Omega)$ (since the embedding of $H_0^1(\Omega)$ into $L^{2^*}(\Omega)$ is not compact). However, we will prove that a version of $(PS)_c$ is recovered at some sufficiently small energy levels c.

We first note that even though we have changed the functional, positive critical points of I still correspond to solutions to (4.9). Indeed, if $u, v \in H_0^1(\Omega)$, then

$$I'(u)(v) = \int_\Omega \nabla u \cdot \nabla v \, dx - \lambda \int_\Omega u^+ v \, dx - \int_\Omega \left(u^+\right)^{2^*-1} v \, dx.$$

Thus, if $u > 0$ is a critical point of I, then

$$\int_\Omega \nabla u \cdot \nabla v \, dx - \lambda \int_\Omega uv \, dx - \int_\Omega u^{2^*-1} v \, dx = 0 \text{ for all } v \in H_0^1(\Omega).$$

We will also take for granted the fact that $I \in C^{1,1}(H_0^1(\Omega))$ since its proof is just a modification of the proof of Lemma 3.3.1. We will simply verify the other conditions of the MPT, applied to I on the Hilbert space $H_0^1(\Omega)$ with norm $||u|| = ||\nabla u||_2$ for every $u \in H_0^1(\Omega)$. We do so using the following results.

Lemma 4.2.1. *Every (PS) (or $(PS)_c$) sequence for I is bounded.*

Proof. Let $(u_n) \subseteq H_0^1(\Omega)$ be a sequence such that for some constant $C > 0$

$$|I(u_n)| \leq C, \; ||I'(u_n)|| \leq C \text{ for all } n \in \mathbb{N}^+ \text{ and } I'(u_n) \to 0 \text{ as } n \to \infty.$$

Then

$$-C \leq I(u_n) \leq C \text{ and } -C\,||u_n|| \leq I'(u_n)(u_n) \leq C\,||u_n|| \text{ for all } n \in \mathbb{N}^+$$

$$\implies -C \leq I(u_n) \leq C \text{ and } -C\,||u_n|| \leq -\frac{1}{2^*} I'(u_n)(u_n) \leq C\,||u_n|| \text{ for all } n \in \mathbb{N}^+$$

for some appropriate $C > 0$. Adding the two equations we get

$$C(1 + ||u||) \geq I(u_n) - \frac{1}{2^*} I'(u_n)(u_n) = \left(\frac{1}{2} - \frac{1}{2^*}\right) ||u_n||^2 - \left(\frac{\lambda}{2} - \frac{\lambda}{2^*}\right) ||u_n^+||_2^2$$

$$\geq \left(\frac{1}{2} - \frac{1}{2^*}\right) ||u_n||^2 - \frac{\lambda}{\lambda_1}\left(\frac{1}{2} - \frac{1}{2^*}\right) ||u_n^+||^2 \geq D\,||u_n||^2$$

for some positive constant D. Thus (u_n) is norm-bounded in $H_0^1(\Omega)$. ■

Now let us first recall *Young's inequality.* If a, b are positive and $p, q > 1$ are conjugate exponents ($\frac{1}{p} + \frac{1}{q} = 1$), then by convexity of the exponential function, we have

$$ab = \exp(\ln a + \ln b) = \exp\left(\frac{1}{p}\ln a^p + \frac{1}{q}\ln b^q\right) \leq \frac{1}{p}a^p + \frac{1}{q}b^q.$$

Now if $\epsilon > 0$, then

$$ab = \left(a\epsilon^{\frac{1}{p}}\right)\left(b\epsilon^{\frac{-1}{p}}\right) \leq \frac{\epsilon}{p}a^p + C_\epsilon b^q \leq \epsilon a^p + C_\epsilon b^q,$$

where C_ϵ is a constant that depends on q, p and ϵ.

Lemma 4.2.2. *Let $0 < c < \frac{1}{N}S^{\frac{N}{2}}$. Then every $(PS)_c$ sequence has a subsequence that converges weakly in $H_0^1(\Omega)$ to a non-trivial critical point u of I.*

Proof. Let $(u_n) \subseteq H_0^1(\Omega)$ be a $(PS)_c$ sequence, with $0 < c < \frac{1}{N}S^{\frac{N}{2}}$. From above, we know that (u_n) is bounded in $H_0^1(\Omega)$, thus there exists $u \in H_0^1(\Omega)$ such that, up to a subsequence,

$$u_n \rightharpoonup u \quad \text{weakly in } H_0^1(\Omega) \tag{4.11}$$

$$u_n \to u \quad \text{strongly in } L^2(\Omega) \tag{4.12}$$

$$u_n \to u \quad \text{a.e. in } \Omega. \tag{4.13}$$

Also, $|u_n^+ - u^+|^2 \leq |u_n - u|^2 \implies u_n^+ \to u^+$ strongly in $L^2(\Omega)$. We now show that $(u_n^+)^{2^*-1} \rightharpoonup (u^+)^{2^*-1}$ weakly in $\left(L^{2^*}(\Omega)\right)' \cong L^{\frac{2^*}{2^*-1}}(\Omega)$. To this end, let $v \in L^{2^*}(\Omega)$ and $\epsilon > 0$. Then by Young's inequality (with $p = 2^*/(2^*-1)$ and $q = 2^*$), there exists a constant $C_\epsilon > 0$ such that

$$\left|(u_n^+)^{2^*-1} v\right| \leq \epsilon \left(u_n^+\right)^{2^*} + C_\epsilon |v|^{2^*} \text{ and } \left|(u^+)^{2^*-1} v\right| \leq \epsilon \left(u^+\right)^{2^*} + C_\epsilon |v|^{2^*}.$$

Hence we have

$$\left|(u_n^+)^{2^*-1} v - (u^+)^{2^*-1} v\right| \leq \left|(u_n^+)^{2^*-1} v\right| + \left|(u^+)^{2^*-1} v\right| \leq \epsilon \left(u_n^+\right)^{2^*} + \epsilon \left(u^+\right)^{2^*} + C_\epsilon |v|^{2^*}$$

for some constant $C_\epsilon > 0$. Next define

$$W_{n,\epsilon} := \left(\left|(u_n^+)^{2^*-1} v - (u^+)^{2^*-1} v\right| - \epsilon \left(u_n^+\right)^{2^*}\right)^+,$$

and note that

$$0 \leq W_{n,\epsilon} \leq \epsilon \left(u^+\right)^{2^*} + C_\epsilon |v|^{2^*} \in L^1(\Omega) \text{ for each } n \in \mathbb{N}^+ \text{ and } W_{n,\epsilon} \to 0 \text{ a.e.}$$

Hence by the dominated convergence theorem,

$$\int_\Omega W_{n,\epsilon}\, dx \to 0 \text{ as } n \to \infty.$$

Also, since (u_n^+) is bounded in $L^{2^*}(\Omega)$ (as it is also bounded in $H_0^1(\Omega)$ and $H_0^1(\Omega)$ is continuously embedded into $L^{2^*}(\Omega)$), we can find a constant $C > 0$ (independent of the chosen ϵ) such that

$$\int_\Omega \left(u_n^+\right)^{2^*} dx \leq C \text{ for every } n \in \mathbb{N}^+.$$

Thus, from the fact that $\left|(u_n^+)^{2^*-1} v - (u^+)^{2^*-1} v\right| \leq W_{n,\epsilon} + \epsilon \left(u_n^+\right)^{2^*}$ for every $n \in \mathbb{N}^+$, we can conclude that

$$0 \leq \limsup_{n\to\infty} \int_\Omega \left|(u_n^+)^{2^*-1} v - (u^+)^{2^*-1} v\right| dx \leq \epsilon C.$$

Upon sending $\epsilon \to 0^+$, we conclude that $(u_n^+)^{2^*-1} \rightharpoonup (u^+)^{2^*-1}$ weakly in $L^{\frac{2^*}{2^*-1}}(\Omega)$.
Thus, if $v \in H_0^1(\Omega) \subseteq L^{2^*}(\Omega)$ then

$$\begin{aligned} 0 = \lim_{n\to\infty} I'(u_n)(v) &= \lim_{n\to\infty} \int_\Omega \nabla u_n \cdot \nabla v\, dx - \lambda \int_\Omega u_n^+ v\, dx - \int_\Omega \left(u_n^+\right)^{2^*-1} v\, dx \\ &= \int_\Omega \nabla u \cdot \nabla v\, dx - \lambda \int_\Omega u^+ v\, dx - \int_\Omega \left(u^+\right)^{2^*-1} v\, dx, \end{aligned}$$

implying that u is a critical point of I. We now show that $u \neq 0$. Suppose that $u = 0$. Since (u_n) is norm-bounded in $H_0^1(\Omega)$, we have (for some $C > 0$)

$$0 \leq |I'(u_n)(u_n)| \leq ||I'(u_n)|| \, ||u_n|| \leq C \, ||I'(u_n)|| \to 0 \text{ as } n \to \infty.$$

Now note that $u_n^+ u_n = (u_n^+)^2$ and $(u_n^+)^{2^*-1} u_n = (u_n^+)^{2^*}$, hence

$$0 = \lim_{n\to\infty} I'(u_n)(u_n) = \lim_{n\to\infty} \int_\Omega |\nabla u_n|^2 \, dx - \lambda \int_\Omega u_n^+ u_n \, dx - \int_\Omega \left(u_n^+\right)^{2^*-1} u_n \, dx$$

$$= \lim_{n\to\infty} ||\nabla u_n||_2^2 - \lambda \left|\left|u_n^+\right|\right|_2^2 - \left|\left|u_n^+\right|\right|_{2^*}^{2^*},$$

and since $u_n^+ \to u^+ = 0$ in $L^2(\Omega)$ by assumption, it follows that

$$0 = \lim_{n\to\infty} ||\nabla u_n||_2^2 - \lambda \left|\left|u_n^+\right|\right|_2^2 - \left|\left|u_n^+\right|\right|_{2^*}^{2^*} = \lim_{n\to\infty} ||\nabla u_n||_2^2 - \left|\left|u_n^+\right|\right|_{2^*}^{2^*}.$$

Since $||\nabla u_n||_2$ is bounded, we may assume that, up to a subsequence, $\gamma = \lim_{n\to\infty} ||\nabla u_n||_2^2 = \lim_{n\to\infty} ||u_n^+||_{2^*}^{2^*}$. Combining this with the fact that $\frac{1}{2} ||\nabla u_n||_2^2 - \frac{\lambda}{2} ||u_n^+||_2^2 - \frac{1}{2^*} ||u_n^+||_{2^*}^{2^*} \to c$, we get that

$$c = \frac{1}{2}\gamma - \frac{1}{2^*}\gamma = \left(\frac{1}{2} - \frac{1}{2^*}\right)\gamma = \frac{1}{N}\gamma.$$

Now from the Sobolev inequality, we have for each $n \in \mathbb{N}^+$

$$||\nabla u_n||_2^2 \geq S \, ||u_n||_{2^*}^2 \geq S \left|\left|u_n^+\right|\right|_{2^*}^2,$$

which implies (by sending $n \to \infty$) that

$$\gamma \geq S\gamma^{\frac{2}{2^*}} \implies \gamma \geq S^{\frac{N}{2}} \implies c \geq \frac{1}{N} S^{\frac{N}{2}},$$

a contradiction. Thus $u \neq 0$.

■

Lemma 4.2.3. *If $S_\lambda < S$, then there exists $v_0 \in H_0^1(\Omega), v_0 \geq 0$ such that*

$$\sup_{t\geq 0} I(tv_0) < \frac{1}{N} S^{\frac{N}{2}}.$$

Proof. Since $S_\lambda < S$ we can find $v_0 \in H_0^1(\Omega)$ with $v_0 \geq 0$ and $||v_0||_{2^*} = 1$ such that

$$0 < ||\nabla v_0||_2^2 - \lambda \, ||v_0||_2^2 < S.$$

Now let $A := ||\nabla v_0||_2^2 - \lambda \, ||v_0||_2^2 > 0$ and define $g : [0, \infty) \to \mathbb{R}$ by $g(t) := I(tv_0) = \frac{1}{2}At^2 - \frac{1}{2^*}t^{2^*}$ for each $t \geq 0$. Since $2^* > 2$, $g(t) > 0$ for sufficiently small $t > 0$, and we conclude that g has a point of absolute maximum $t_0 \in (0, \infty)$ with $g'(t_0) = 0$. Solving this equation we get

$$t_0 = A^{\frac{N-2}{4}} \text{ and } g(t_0) = \frac{1}{N} A^{\frac{N}{2}} < \frac{1}{N} S^{\frac{N}{2}}.$$

Hence

$$\sup_{t\geq 0} I(tv_0) = g(t_0) < \frac{1}{N} S^{\frac{N}{2}}.$$

■

Lemma 4.2.4. *If $S_\lambda < S$, then there exists PS level $c \in \left(0, \frac{1}{N}S^{\frac{N}{2}}\right)$ for I.*

Proof. We apply the MPT to $H_0^1(\Omega)$ with the norm $||u|| = ||\nabla u||_2$ for $u \in H_0^1(\Omega)$. The proof that $I \in C^{1,1}(H_0^1(\Omega))$ is similar to the proof of Lemma 3.3.1 and uses the convergence results derived in Lemma 4.2.2. We also have $I(0) = 0$, and by the Sobolev inequality

$$I(u) \geq \left(\frac{1}{2} - \frac{1}{2^*}\right) ||u||^2 - C\,||u||^{2^*},$$

which implies (because $2^* > 2$) that there exists $\rho > 0$ and $\alpha > 0$ such that $I(u) \geq \alpha$ when $||u|| = \rho$. Since $S_\lambda < S$, we choose v_0 from Lemma 4.2.3. Then since $I(tv_0) = \frac{1}{2}At^2 - \frac{1}{2^*}t^{2^*} \to -\infty$ as $t \to \infty$, there exists $t_1 > 0$ such that $v = t_1 v_0$ satisfies $I(v) \leq 0$. Let $\Gamma := \{\gamma \in C([0,1]\,, H_0^1(\Omega)) : \gamma(0) = 0, \gamma(1) = v\}$ be the set of continuous paths joining 0 to v. By the MPT, c defined by

$$c := \inf_{\gamma\in\Gamma} \sup_{t\in[0,1]} I(\gamma(t))$$

is a PS level for I. Since $(t \mapsto tv; t \in [0,1])$ belongs to Γ, we have

$$c = \inf_{\gamma\in\Gamma} \sup_{t\in[0,1]} I(\gamma(t)) \leq \sup_{t\in[0,1]} I(tv) \leq \sup_{t\geq 0} I(tv_0) < \frac{1}{N}S^{\frac{N}{2}}.$$

We also conclude from the MPT that $c \geq \alpha > 0$, hence $c \in \left(0, \frac{1}{N}S^{\frac{N}{2}}\right)$. ■

Combining these results we obtain a solution when $S_\lambda < S$.

Theorem 4.2.2. *If $S_\lambda < S$, then there exists a solution to (4.9).*

Proof. From the previous result, there exists PS level $c \in \left(0, \frac{1}{N}S^{\frac{N}{2}}\right)$ of I. If $(u_n) \subseteq H_0^1(\Omega)$ is a $(PS)_c$ sequence, then by Lemma 4.2.2, (u_n) has a subsequence that converges weakly in $H_0^1(\Omega)$ to a non-trivial critical point u of I. Then u weakly solves the equation

$$\begin{cases} -\Delta u = \lambda u^+ + (u^+)^{2^*-1} & x \in \Omega \\ u = 0 & x \in \partial\Omega. \end{cases}$$

From the weak and strong maximum principles we can conclude that $u > 0$ in Ω. Thus u solves (4.9). ■

We are therefore left with answering the question of when is $S_\lambda < S$? The answer depends on the dimension N and is given below.

Theorem 4.2.3. *There exists $0 \leq \lambda_* < \lambda_1$ such that $S_\lambda < S$ for any $\lambda_* < \lambda < \lambda_1$. Furthermore, $\lambda_* = 0$ if $N \geq 4$ and $\lambda_* = \frac{\lambda_1}{4}$ if Ω is an open ball in $\mathbb{R}^3$.*

The proof of this result is technical and is therefore omitted. It can be found in ([Struwe, 1990]) and ([Brézis and Nirenberg, 1983]). The idea is to estimate the quotient $(||\nabla u||_2^2 - \lambda\,||u||_2^2)/||u||_{2^*}^2$ with functions $u_\epsilon = \xi\psi_\epsilon$, where ξ is a cut-off function that is equal to 1 in a neighbourhood of 0 (assuming $0 \in \Omega$), and

$$\psi_\epsilon(x) = \frac{C_\epsilon}{\left(\epsilon^2 + |x|^2\right)^{\frac{N-2}{2}}},\ \epsilon > 0,$$

where $C_\epsilon > 0$ depends only on ϵ and N.

4.3 The Non-homogeneous Equation

We now show that if we add a non-homogeneous term, the semilinear equation discussed in the previous section is solvable, as long as the non-homogeneous term is not too large. We investigate the existence of a solution to the equation

$$\begin{cases} -\Delta u + cu = |u|^{2^*-2} u + f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{4.14}$$

where $f \in L^2(\Omega), f \neq 0$ and $c \geq 0$ with $c \in L^\infty(\Omega)$. Variants of this problem have been solved by [Tarantello, 1992] and [Naito and Sato, 2012], among others, and their approaches are fundamentally different. The method by [Tarantello, 1992] is based on variational techniques, and we briefly outline it below for $c = 0$.

First note that solutions to (4.14) (when $c = 0$) correspond to critical points of the functional $I : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$I(u) := \frac{1}{2} \int_\Omega |\nabla u|^2 \, dx - \frac{1}{2^*} \int_\Omega |u|^{2^*} \, dx - \int_\Omega fu \, dx, \ u \in H_0^1(\Omega).$$

An attempt to directly apply the methods of Chapter 3 to find critical points fails due to the non-compactness of the embedding of $H_0^1(\Omega)$ into $L^{2^*}(\Omega)$. It is shown in [Tarantello, 1992] that if f satisfies the following condition:

$$\int_\Omega fu \, dx \leq \frac{4}{N-2} \left(\frac{N-2}{N+2} \right)^{\frac{N+2}{4}} (||\nabla u||_2)^{\frac{N+2}{2}} \text{ for all } u \in \mathcal{A}_{2^*} := \left\{ v \in H_0^1(\Omega) : ||v||_{2^*} = 1 \right\}, \tag{4.15}$$

then I has a critical point. This is achieved through minimizing I on the so-called *Nehari Manifold* $\mathcal{N}$ defined by

$$\mathcal{N} := \left\{ u \in H_0^1(\Omega) : I'(u)(u) = 0 \right\}.$$

Note that $\mathcal{N}$ contains all critical points of I, and if we assume (4.15) then I is bounded below on $\mathcal{N}$. The proof of this result uses *Ekeland's Variational Principle* (see [Ekeland, 1974] for more details) to find a minimizing sequence that is also a $(PS)_\mu$ sequence, where μ is the infimum of I on $\mathcal{N}$.

In contrast, we are going to use the method by [Naito and Sato, 2012] which applies the implicit function theorem. First define $F : L^2(\Omega) \times H_0^1(\Omega) \to H^{-1}(\Omega)$ by

$$F(f, u) := -\Delta u + cu - |u|^{2^*-2} u - f, \ u \in H_0^1(\Omega) \text{ and } f \in L^2(\Omega).$$

Since $F(f, u) \in H^{-1}(\Omega)$, this definition means that $F(f, u)$ is a functional that maps each $v \in H_0^1(\Omega)$ to the real number $F(f, u)(v)$ given by

$$F(f, u)(v) = \int_\Omega \nabla u \cdot \nabla v + cuv - |u|^{p-2} uv - fv \, dx.$$

Note that $F(0, 0) = 0$, which is in line with the fact that (4.14) has a trivial solution when $f \equiv 0$. The idea is to apply the implicit function theorem to F at the point $(0, 0)$. We first recall this theorem, taken from [Zeidler, 1995].

Theorem 4.3.1 (Implicit Function Theorem). *Let X, Y and Z be Banach spaces, $(x_0, y_0) \in X \times Y$ and $F : U \subseteq X \times Y \to Z$ be defined in an open neighbourhood U of (x_0, y_0) (in the product of the norm topologies of X and Y). Further suppose that F satisfies the following:*

1. *$F((x_0, y_0)) = 0$.*
2. *$F \in C^1(U, Z)$*
3. *$F_y(x_0, y_0) : Y \to Z$, the partial derivative with respect to y at (x_0, y_0), is invertible; i.e. the inverse operator $F_y(x_0, y_0)^{-1} : Z \to Y$ exists as a continuous linear operator.*

Then there exist $r_0 > 0, r_1 > 0$ and $H \in C^1(B_X(x_0, r_0), B_Y(y_0, r_1))$ such that $F(x, H(x)) = 0$ for all $x \in B_X(x_0, r_0)$.

Remark 4.3.1. It is worth noting that if $||\cdot||_X$ and $||\cdot||_Y$ are norms on X and Y respectively, then $||(x, y)||_{X\times Y} := ||x||_X + ||y||_Y$ is a norm on $X \times Y$ that induces the product of the norm topologies. We will always use this norm on the product.

Theorem 4.3.2. *There exists $r > 0$ such that for each $f \in L^2(\Omega)$ with $||f||_2 < r$, (4.14) has a solution $u_f \in H_0^1(\Omega)$.*

Remark 4.3.2. [Tarantello, 1992] shows that (4.14) has at least two solutions when $||f||_2$ is sufficiently small. Our theorem only guarantees the existence of a solution in this case.

Proof. We apply the implicit function theorem to $Y = H_0^1(\Omega), X = L^2(\Omega), Z = H^{-1}(\Omega)$ and $F(f, u) := -\Delta u + cu - |u|^{2^*-2} u - f,\ u \in H_0^1(\Omega)$ and $f \in L^2(\Omega)$ as defined above. Note that F is well-defined for all $f \in L^2(\Omega)$ and $u \in H_0^1(\Omega)$, so we take $(x_0, y_0) = (0, 0)$ and $U = X \times Y$ be the neighbourhood of $(0, 0)$. We now show that F satisfies the required conditions.

1. $F((0, 0)) = 0$ is clear.
2. We now show that $F \in C^1(L^2(\Omega) \times H_0^1(\Omega), H^{-1}(\Omega))$. If $(f, u), (g, v) \in L^2(\Omega) \times H_0^1(\Omega)$ and $t \in \mathbb{R}$ is sufficiently small, then $F((f, u) + t(g, v)) = -\Delta u + t(-\Delta v) + cu + tcv - |u + tv|^{2^*-2} (u + tv) - (f + tg)$. Hence the (Gateaux for now) derivative is

$$F'((f, u))(g, v) = F'_G((f, u))(g, v) = -\Delta v + cv - (2^* - 1) |u|^{2^*-2} v - g \in H^{-1}(\Omega).$$

Now let $((f_n, u_n)) \subseteq L^2(\Omega) \times H_0^1(\Omega)$ be a sequence such that $((f_n, u_n)) \to ((f, u))$ in $L^2(\Omega) \times H_0^1(\Omega)$. Then

$$\begin{aligned} &||F'((f_n, u_n)) - F'((f, u))||_{\mathcal{L}\left(L^2(\Omega)\times H_0^1(\Omega), H^{-1}(\Omega)\right)} \\ &= \sup_{\substack{(g,v)\in L^2(\Omega)\times H_0^1(\Omega) \\ ||(g,v)||_{L^2(\Omega)\times H_0^1(\Omega)}\le 1}} ||F'((f_n, u_n))(g, v) - F'((f, u))(g, v)||_{H^{-1}(\Omega)} \\ &= \sup_{\substack{(g,v)\in L^2(\Omega)\times H_0^1(\Omega) \\ ||(g,v)||_{L^2(\Omega)\times H_0^1(\Omega)}\le 1}} \sup_{\substack{w\in H_0^1(\Omega) \\ ||w||_{H_0^1(\Omega)}\le 1}} |F'((f_n, u_n))(g, v)(w) - F'((f, u))(g, v)(w)| \\ &\le (2^* - 1) \sup_{\substack{(g,v)\in L^2(\Omega)\times H_0^1(\Omega) \\ ||(g,v)||_{L^2(\Omega)\times H_0^1(\Omega)}\le 1}} \sup_{\substack{w\in H_0^1(\Omega) \\ ||w||_{H_0^1(\Omega)}\le 1}} \int_\Omega \left| |u_n|^{2^*-2} - |u|^{2^*-2} \right| |v|\, |w|\ dx \\ &\le C \left(\int_\Omega \left| |u_n|^{2^*-2} - |u|^{2^*-2} \right|^{\frac{2^*}{2^*-2}} dx \right)^{\frac{2^*-2}{2^*}} \to 0 \end{aligned}$$

by the generalized Holder's inequality. Thus $F \in C^1(L^2(\Omega) \times H_0^1(\Omega), H^{-1}(\Omega))$.

3. Finally, $F_u(0,0)(v) = -\Delta v + cv$. From Remark 3.1.3, we know that for each $h \in H^{-1}(\Omega)$, there exists a unique $v \in H_0^1(\Omega)$ such that $F_u(0,0)(v) = -\Delta v + cv = h$. Thus $F_u(0,0) : H_0^1(\Omega) \to H^{-1}(\Omega)$ is invertible.

So we can apply the Implicit Function Theorem to find $r_0 > 0, r_1 > 0$ and $H \in C^1(B_X(0,r_0), B_Y(0,r_1))$ such that $F(f, H(f)) = 0$ for all $f \in L^2(\Omega)$ with $||f||_2 < r_0$. ■

Chapter 5

Conclusion and Further Research

In the last three chapters we have discussed and applied the variational approach to solving PDEs. In this approach we classified solutions to certain PDEs as minimizers or critical points of certain functionals in an appropriate Banach space. This process has lead us to consider generalised solutions in Sobolev spaces. We observed that the Sobolev spaces are well-suited to the methods of functional analysis since they share similar properties with the ubiquitous L^p spaces. This approach to PDEs has been around for many years and significant progress was made during the last century. This can in part be attributed to David Hilbert including variational and PDE problems in his list of twenty three problems during his famous lecture in 1900.

We first considered the linear problem

$$\begin{cases} Lu = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.1}$$

where $L = -\Delta + c, c \in L^\infty(\Omega)$ is a prototype for linear elliptic operators. The linear case presents little difficulty and the question of existence is settled by applying the Riesz representation theorem and using the symmetry of L. In cases where the linear elliptic operator L does not generate a symmetric bilinear form, the *Lax-Milgram* theorem also gives existence results when the bilinear form generated by L is continuous and coercive (see [Evans, 1998]). The same symmetry also played a crucial role in applying the spectral theorem for compact self-adjoint operators to find the eigenvalues of L. These eigenvalues also have a variational characterization as constrained minimizers of a certain functional.

In moving from linear to non-linear PDEs, the first tool we used was characterizing weak solutions as minimizers, and then establishing directly via minimizing sequences that such minimizers exist. In proving that a functional $I : E \to \mathbb{R}$ on a Banach space E has a minimum, we observed that compactness of E and continuity of I play decisive roles, and this lead to the generalisation of Weierstrass' Extreme Value Theorem (EVT). Indeed, we first observed that in applications, E is usually not compact in the strong (norm) topology. However, when E is reflexive, the bounded sets are relatively compact in the weak topology induced by the dual E', so when I is coercive (which ensures that minimizing sequences are bounded) and sequentially lower-semicontinuous with respect to this weak topology, then we recover a general version of the EVT.

We also observed that the generalised EVT does not apply in may situations of importance in PDEs, particularly due to lack of coercivity of I (which would ensure boundedness of minimizing sequences). In this situation we considered two alternatives: directly searching for critical points using the MPT and minimizing I on spheres. The main problem we considered is a variant of

$$\begin{cases} -\Delta u = |u|^{p-2} u & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.2}$$

for some exponent $p \geq 2$. The exponent p plays a crucial role in the existence of solutions to (5.2). When $p < 2^* = \frac{2N}{N-2}$ — the *sub-critical exponent*, then (5.2) has a non-trivial solution. For application of the MPT, the sub-critical exponent ensures that the associated energy functional satisfies (PS), while for constrained minimization on spheres, it ensures that the sphere (to which the functional is constrained) is sequentially closed with respect to the weak topology. Both these facts are important in the proof of existence of non-trivial solutions to (5.2). Another technique used to deal with non-coercive functionals is to perform constrained minimization the the Nehari Manifold.

The case when $p = 2^*$ — the *critical exponent* — is much more difficult. We first observed that the question of existence is dependent on the domain Ω in the following way. First we proved that if Ω is star-shaped with respect to the origin, then (5.2) does not have positive solutions. This result highlights the fact that the inability to prove the existence of solutions in the critical case is not just a deficiency of the variational technique, but that no solutions exist. We then showed that the opposite is true when the domain is a ring ($\Omega = \left\{x \in \mathbb{R}^N : 0 < a < |x| < b\right\}$). This result is a special case of a very important result by [Bahri and Coron, 1988] that states that positive non-trivial solutions to (5.2) exist when the topology of the domain is non-trivial.

The case of a general bounded domain is still open. Fundamental work in this direction was conducted by [Brézis and Nirenberg, 1983], where they considered the perturbed problem:

$$\begin{cases} -\Delta u = u^{p-1} + f(x, u) & \text{in } \Omega \\ u > 0 & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.3}$$

where $p = 2^*$ and f is a lower order homogeneous ($f(x, 0) = 0$) perturbation of u^{p-1}, in the sense that

$$\lim_{u\to\infty} \frac{f(x, u)}{u^{p-1}} = 0.$$

Considering the special case $f(x, u) = \lambda u$, we showed that there exists $\lambda_* \in (0, \lambda_1)$ such that (5.3) has a solution when $\lambda_* < \lambda < \lambda_1$. Here λ_1 is the smallest eigenvalue of the Dirichlet Laplacian $-\Delta$. The general perturbation f is considered in [Brézis and Nirenberg, 1983], but our proof and theirs are similar in approach.

Finally, we considered the non-homogeneous problem. The main takeaway is that a solution exists as long as the non-homogeneous term is 'not too large' in the L^2 norm. This is consistent with similar results obtained by, among others, ([Tarantello, 1992]) and ([Naito and Sato, 2012]).

Bibliography

[Adams and Fournier, 2003] Adams, R. A. and Fournier, J. J. (2003). *Sobolev spaces*, volume 140. Academic press.

[Ambrosetti and Rabinowitz, 1973] Ambrosetti, A. and Rabinowitz, P. H. (1973). Dual variational methods in critical point theory and applications. *Journal of functional Analysis*, 14(4):349–381.

[Aubin, 1976] Aubin, T. (1976). Equation différentielles non linéaires et problèmes de yamabe concernant la courbure scalaire. *J. Math. pures appl.*, 55:269–296.

[Badiale and Serra, 2010] Badiale, M. and Serra, E. (2010). *Semilinear Elliptic Equations for Beginners: Existence Results via the Variational Approach*. Springer Science & Business Media.

[Bahri and Coron, 1988] Bahri, A. and Coron, J. (1988). On a nonlinear elliptic equation involving the critical sobolev exponent: the effect of the topology of the domain. *Communications on pure and applied mathematics*, 41(3):253–294.

[Ball and Mizel, 1987] Ball, J. M. and Mizel, V. J. (1987). One-dimensional variational problems whose minimizers do not satisfy the euler-lagrange equation. In *Analysis and Thermomechanics*, pages 285–348. Springer.

[Brezis, 2010] Brezis, H. (2010). *Functional analysis, Sobolev spaces and partial differential equations*. Springer Science & Business Media.

[Brezis and Browder, 1998] Brezis, H. and Browder, F. (1998). Partial differential equations in the 20th century. *Advances in Mathematics*, 135(1):76–144.

[Brézis and Lieb, 1983] Brézis, H. and Lieb, E. (1983). A relation between pointwise convergence of functions and convergence of functionals. *Proceedings of the American Mathematical Society*, 88(3):486–490.

[Brézis and Nirenberg, 1983] Brézis, H. and Nirenberg, L. (1983). Positive solutions of nonlinear elliptic equations involving critical sobolev exponents. *Communications on Pure and Applied Mathematics*, 36(4):437–477.

[Buttazzo et al., 1998] Buttazzo, G., Giaquinta, M., and Hildebrandt, S. (1998). *One-dimensional variational problems: an introduction*. Number 15. Oxford University Press.

[Ciarlet, 2002] Ciarlet, P. G. (2002). *The finite element method for elliptic problems*. SIAM.

[Courant, 1920] Courant, R. (1920). Über die eigenwerte bei den differentialgleichungen der mathematischen physik. *Mathematische Zeitschrift*, 7(1-4):1–57.

[Dunford et al., 1971] Dunford, N., Schwartz, J. T., Bade, W. G., and Bartle, R. G. (1971). *Linear operators*. Wiley-interscience New York.

[Ekeland, 1974] Ekeland, I. (1974). On the variational principle. *Journal of Mathematical Analysis and Applications*, 47(2):324–353.

[Evans, 1998] Evans, L. (1998). *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society.

[Fischer, 1905] Fischer, E. (1905). Über quadratische formen mit reellen koeffizienten. *Monatshefte für Mathematik und Physik*, 16(1):234–249.

[Gelfand and Fomin, 1963] Gelfand, I. and Fomin, S. (1963). Calculus of variations. 1963.

[Gilbarg and Trudinger, 2015] Gilbarg, D. and Trudinger, N. S. (2015). *Elliptic partial differential equations of second order*. springer.

[Jannelli, 1999] Jannelli, E. (1999). The role played by space dimension in elliptic critical problems. *Journal of Differential Equations*, 156(2):407–426.

[Kazdan and Warner, 1975] Kazdan, J. L. and Warner, F. (1975). Remarks on some quasilinear elliptic equations. *Communications on Pure and Applied Mathematics*, 28(5):567–597.

[Kurdila and Zabarankin, 2006] Kurdila, A. J. and Zabarankin, M. (2006). *Convex functional analysis*. Springer Science & Business Media.

[Lax and Milgram, 2016] Lax, P. and Milgram, A. (2016). Ix. parabolic equations. *Contributions to the Theory of Partial Differential Equations.(AM-33)*, 33:167.

[Lee and Parker, 1987] Lee, J. M. and Parker, T. H. (1987). The yamabe problem. *Bull. Amer. Math. Soc. (N.S.)*, 17(1):37–91.

[Lions, 1982] Lions, P.-L. (1982). On the existence of positive solutions of semilinear elliptic equations. *SIAM review*, 24(4):441–467.

[Naito and Sato, 2012] Naito, Y. and Sato, T. (2012). Non-homogeneous semilinear elliptic equations involving critical sobolev exponent. *Annali di Matematica Pura ed Applicata*, 191(1):25–51.

[Nehari, 1960] Nehari, Z. (1960). On a class of nonlinear second-order differential equations. *Transactions of the American Mathematical Society*, 95(1):101–123.

[Nirenberg, 1953] Nirenberg, L. (1953). A strong maximum principle for parabolic equations. *Communications on Pure and Applied Mathematics*, 6(2):167–177.

[Palais and Smale, 1964] Palais, R. S. and Smale, S. (1964). A generalized morse theory. *Bull. Amer. Math. Soc.*, 70(1):165–172.

[Pohozhaev, 1965] Pohozhaev, S. (1965). Eigenfunctions of the equation $\Delta u+\lambda f(u) = 0$. *Sov. Math., Dokl.*, 6:1408–1411.

[Rabinowitz et al., 1986] Rabinowitz, P. H. et al. (1986). *Minimax methods in critical point theory with applications to differential equations*. Number 65. American Mathematical Soc.

[Riemann, 1851] Riemann, B. (1851). *Grundlagen für eine allgemeine Theorie der Functionen einer veränderlichen complexen Grösse: Eine Abh.(pro grad. Philos.).*

[Riemann, 1857] Riemann, B. (1857). *Beiträge zur Theorie der durch die Gauss' sche Reihe F (alpha, beta, gamma, x) darstellbaren Functionen.* Dieterich.

[Struwe, 1990] Struwe, M. (1990). *Variational methods*, volume 31999. Springer.

[Talenti, 1976] Talenti, G. (1976). Best constant in sobolev inequality. *Annali di Matematica pura ed Applicata*, 110(1):353–372.

[Tarantello, 1992] Tarantello, G. (1992). On nonhomogeneous elliptic equations involving critical sobolev exponent. In *Annales de l'IHP Analyse non linéaire*, volume 9, pages 281–304.

[Trudinger, 1968] Trudinger, N. S. (1968). Remarks concerning the conformal deformation of riemannian structures on compact manifolds. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 22(2):265–274.

[Zeidler, 1995] Zeidler, E. (1995). Applied functional analysis, main principles and their application.