Math 951 Lecture Notes
Chapter 5 – Variational Methods

Mathew A. Johnson [1]
Department of Mathematics, University of Kansas
matjohn@ku.edu

Contents

1 Introduction

In this chapter, we consider a class of techniques known as "variational methods" to study the existence of solutions to linear and nonlinear PDE. The method has a relatively simple idea: if one wants to show a function $f(x) = 0$ has a solution, it is enough to show that an antiderivative of f has a critical point. While this may not seem like it gives you much, the advantage is that the tool box for proving differentiable functions have critical points is different from (and sometimes easier to use than) the tool box of showing some continuous function has a root. In our PDE applications, of course, it is not immediatley clear how we can think of a PDE as the "critical point equation" for its "antiderivative". We start with an illustrating example to demonstrate the overall method, and develop a more general methodology that we will apply to both linear and nonlinear PDE problems.

1.1 A Motivating Example

Let $\Omega \subset \mathbb{R}^n$ be open and bounded, and consider the functional $F : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$F(u) \int_\Omega |Du|^2 dx.$$

Our question is the following:

What is the minimum value that F can attain on $H_0^1(\Omega)$?

Well, clearly the minimum value of F over all of $H_0^1(\Omega)$ is zero, which is achieved precisely by the constant function $u = 0$. Since that wasn't very interesting, lets change the question slightly to the following:

What is the minimum value that F can attain on $H_0^1(\Omega)$ subject to the constraint $\|u\|_{L^2(\Omega)} = 1$?

Note that the enforcement of the constraint now precludes the constant function $u = 0$, and hence the question becomes now becomes more interesting as we can not answer the question trivially! In fact, the answer is give by the following.

CLAIM: If $\phi \in H_0^1(\Omega)$ minimizes $F(u)$ over $H_0^1(\Omega)$ subject to the constraint $\|u\|_{L^2(\Omega)} = 1$, then ϕ is a non-trivial weak solution of the BVP

$$\begin{cases} -\Delta u = \lambda_1 u, & \text{in } \Omega \\ \quad u = 0, & \text{on } \partial\Omega, \end{cases}$$

where λ_1 is the principle eigenvalue for $-\Delta$ on $H_0^1(\Omega)$.

According to the above claim, the minimizer of F subject to $\|u\|_{L^2(\Omega)} = 1$ is precisely the principle eigenfunction ϕ_1, normalized so that $\|\phi_1\|_{L^2(\Omega)} = 1$, of $-\Delta$ acting on $H_0^1(\Omega)$. While this is an interesting result in its own right (and we will prove this rigorously later in Section 3, the main take away should be that, somehow, proving the existence of (in this case, constrained) minimizers of functionals can be equivalent to proving the existence of non-trivial solutions to PDE!

So, how can we attempt to justify the above claim? Well, this honestly seems a LOT like a constrained minimization problem that you might encounter in multi-variable calculus.

So, why don't we try to treat this like a calculus problem and pretend like we can apply the method of Lagrange multipliers here! Noting that the constraint $\|u\|_{L^2(\Omega)} = 1$ can be written as the zero set of the functional

$$G(u) = \int_\Omega |u|^2 dx - 1$$

the method of Lagrange multipliers from multivariable calculus suggests that if ϕ is a minimizer of $F(u)$ subject to the constraint $G(u) = 1$ then there must exist a $\lambda \in \mathbb{R}$ such that

$$\text{``}\frac{dF}{du}(\phi) = \lambda \frac{dG}{du}(\phi)\text{''}. \tag{1}$$

While we will introduce the rigorous theory to make this precise later, here we can at least formally calculate the above derivatives by attempting to do a Taylor expansion of the functionals involved. Indeed, notice that if ϕ satisfies (1) and if $v \in H_0^1(\Omega)$ is arbitrary, then for all $0 < |\epsilon| \ll 1$ we have

$$F(\phi + \epsilon v) = \int_\Omega |D\phi + \epsilon Dv|^2 \, dx = F(\phi) + 2\epsilon \int_\Omega D\phi \cdot Dv \; dx + \mathcal{O}(\epsilon^2)$$

and, similarly,

$$G(\phi + \epsilon v) = G(\phi) + 2\epsilon \int_\Omega \phi v \; dx + \mathcal{O}(\epsilon^2).$$

In particular, following our multivariable calculus intuition we should now have

$$\frac{dF}{du}(\phi) = \lim_{\epsilon \to 0} \frac{F(\phi + \epsilon v) - F(\phi)}{\epsilon} = \int_\Omega D\phi \cdot Dv \; dx$$

and, similarly

$$\frac{dG}{du}(\phi) = \lim_{\epsilon \to 0} \frac{G(\phi + \epsilon v) - G(\phi)}{\epsilon} = \int_\Omega \phi v \; dx$$

and hence the Lagrange multiplier equation (1) is equivalent with saying that λ is an eigenvalue of $-\Delta$ on $H_0^1(\Omega)$ with weak eigenfunction ϕ. Of course, while this is a step towards the our main claim above, it still remains to show that, in fact the λ and ϕ in (1) actually correspond to the principle eigenvalue/eigenfunction pair for $-\Delta$ on $H_0^1(\Omega)$ (we will take care of this later in Section 3. Nevertheless, my hope is that this at least motivates a connection between proving the existence of (in this case, constrained) minimizes of functionals and proving the existence of solutions to PDE.

Of course, in order for the above to be useful in developing existence theories for PDE, we have to introduce the appropriate abstract "mumbo-jumbo" needed to make sense of how to differentiate and minimize a functional. Also, if we want to use the above ideas to help prove the existence of solutions to PDE we have to actually decide when we can guarantee that a given functional has constrained or unconstrained critical points. Both of these issues will be discussed in detail in this chapter. Once we have the appropriate framework in place, we will return to the above problem in Section 3.

1.2 Advantages of the Variational Approach

Before continuing with our mathematical analysis, I want to point out some important advantages to the "variational" approach in PDE.

Advantage 1: Prove the existence of critical points for functionals is often easier than proving the existence of solutions to nonlinaer PDE by direct functional analytic methods. This is not always true, of course, but there are special classes of nonlinear PDE where this is a huge advantage.

Advantage 2: In time dependent problems, the classification of a critical point as a max, min, or saddle often yields important information about the "dynamics" (i.e. behavior) of solutions as time progresses. For an elementary example, consider the "nonlinear oscillator" equation

$$\frac{d^2x}{dt^2} = -\frac{dV}{dx}, \tag{2}$$

where here $x(t) \in \mathbb{R}$ and $V(x)$ represents the potential energy of the system. In this case, so-called "equilibrium solutions", i.e. solutions that don't depend on time, arise as critical points of the potential energy. Furthermore, in this case the flow is subject to conservation of energy: if $x(t)$ is a solution of (2) then the total energy for the system

$$E(t) = \frac{1}{2}\left(\frac{dx}{dt}\right)^2 + V(x(t))$$

is a constant function of time. By analysis of the phase plane it is clear that local minimia of V correspond to"stable" equilibrium solutions of (2), in the sense that solutions starting near them stay near them for all time, while equilibria which are local maxima or saddles are necessarialy "unstable", in the sense that some solutions that start near them move away over time.

Note that while the above motivation is given in the ODE setting, its corresponding application in the PDE setting is incredibly powerful and gives important results in applications such as water wave dynamics, fiber optical communication, and Bose-Einstein condensation. Take Math 851 for more details!

Advantage 3: It turns out that many fundamental laws of physics (and other sciences) arise more naturally as variational principles rather than PDE/ODE. For example, suppose one has a collection of objects whose total kinetic energy is T and whose total potential energy is $\mathcal{U}$. If this hypothetical system is "conservative", i.e. there is no energy dissipation due to internal or external friction, then the dynamics (i.e. motion) of the object follows from Hamilton's Principle of Least Action, which states the objects must move from starting position at time t_1 to ending position at time t_2 along trajectories which correspond to critical points of the functional

$$J = \int_{t_0}^{t_1} (T(s) - \mathcal{U}(s))\, ds$$

Here, the function $\mathcal{L} := T - \mathcal{U}$ is known as the Lagrangian of the system, and the above formalism leads to the Lagrangian formulation of classical mechanics. In this sense, all of Lagrangian mechanics can be viewed under the heading of a "variational method".

Similarly, even some of the most fundamental PDE studied in elementary PDE classes arise most naturally through variational methods. For example, suppose you want to model the motion of a vibrating elastic string of length L with fixed end points. Letting x denote the physical position along the string, t denote time, and $u(x,t)$ denote the displacement from equilibrium of the string at position x at time t, then the total kinetic energy of the string at time t is given by

$$T(t) = \frac{1}{2}\int_0^L \rho(x,t)(u_t)^2 dx,$$

where $\rho > 0$ denotes the (variable) density of the string[2]. Furthermore, Hooke's law implies the potential energy of the sting is proportional to the change in arclength from equilibrium, i.e. for some constant $k > 0$ we have

$$\mathcal{U}(t) = k\int_0^L \left(\sqrt{1+u_x^2} - 1\right) dx$$

By Hamilton's principle the motion of the strong from time t_1 to time t_2 is described by the critical points of the action functional

$$J(u) := \int_{t_1}^{t_2}\int_0^L \left[\frac{1}{2}\rho u_t^2 - k\left(\sqrt{1+u_x^2} - 1\right)\right] dx\ dt.$$

Using a similar Taylor expansion method as before, the critical points u of J satisfy

$$\lim_{\epsilon\to 0}\frac{J(u+\epsilon v) - J(u)}{\epsilon} = \int_{t_1}^{t_2}\int_0^L \left[-(\rho u_t)_+ \partial_x\left(\frac{ku_x}{\sqrt{1+(u_x)^2}}\right)\right] v\ dx\ dt = 0$$

for all smooth v with $v(0,t) = v(L,t)$ (enforcing the fixed endpoint condition!). Since this should hold for all such v, critical points of the action satisfy the PDE

$$(\rho u_t)_t - \partial_x\left(\frac{ku_x}{\sqrt{1+(u_x)^2}}\right) = 0,$$

which is a nonlinear wave equation allowing for the possibility of "large" deviations from equilibrium[3] Note the traditional linear wave equation is recovered by assuming $|u| \ll 1$ and $|u_x| \ll 1$ and that the density ρ is constant, in which case the above reduces (at leading order) to

$$u_{tt} = \frac{k}{\rho}u_{xx},$$

[2]Note that, morally, this is just saying $T = \frac{1}{2}mv^2$

[3]At least, large deviations within the range of validity of Hooke's law...

i.e. to the linear wave equation on $[0, L]$.

Above, the main point is that the natural description of the motion of the spring is in terms of a variational problem, and the PDE only follows *after* one does some analysis of the variational problem. In this way, in many applications the use of variational methods is more directly linked to the underlying modeling and physics.

2 Calculus in Banach Spaces

From our formal, motivating example in Section 1.1 above, it is clear that we need to develop a differential calculus for real-valued maps with infinite dimensional domains. This is the goal of this section.

To begin, let X and Y be Banach spaces (over either $\mathbb{R}$ or $\mathbb{C}$) and recall the space

$$\mathcal{L}(X,Y) := \{A : X \to Y \;\; : \;\; A \text{ is linear and bounded}\}$$

is a Banach space when equipped with the natural norm

$$\|L\| := \sup_{\|f\|_X = 1} \|L(f)\|_Y.$$

Throughout this section, let $U \subset X$ be open and let $F : U \to Y$ be a (possibly nonlinear) map. The next definitions extend the notions of directional derivatives and total derivatives to this infinite dimensional setting.

Definition 1. *The map $F : U \subset X \to Y$ has Gâteaux derivative at $u \in U$ in the direction $x \in X$ if*

$$\lim_{\epsilon \to 0} \frac{F(u + \epsilon x) - F(u)}{\epsilon}$$

exists and is finite, in which case we denote the limit by $dF[u]x$. If $dF[u]x$ exists for every direction $x \in X$, we say that F is Gâteaux differentiable at u. Further, we say F is Gâteaux differentiable on the set U if F is Gâteaux differentiable at every point $u \in U$.

Definition 2. *The map $F : U \subset X \to Y$ is Fréchet differentiable at $u \in U$ if there exists $A \in \mathcal{L}(X, Y)$ such that*

$$\lim_{\|x\|_X \to 0} \frac{\|F(u + x) - f(u) - Ax\|_Y}{\|x\|_X} = 0.$$

In this case, we denote $A = DF(u)$ and call it the Fréchet derivative of F at u. If F is Fréchet differentiable at every $u \in U$, then we say F is Fréchet differentiable on U.

Our first result should be familiar from multi-variable calculus, and simply states that if a function is differentiable then its directional derivatives all exist.

Theorem 1. *If $F : U \subset X \to Y$ is Fréchet differentiable at a point $u \in U$, then it is Gâteaux differentiable at u and*

$$dF[u]x = DF(u)x$$

for every $x \in X$.

Proof. If F is Fréchet differentiable at $u \in U$, then for all $x \in X$ we have

$$\frac{F(u+\epsilon x) - F(u)}{\epsilon} = \operatorname{sgn}(\epsilon)\|x\|_X \cdot \frac{F(u+\epsilon x) - F(u) - DF(u)\epsilon x}{\|\epsilon x\|_X} + DF(u)x$$

so that

$$\lim_{\epsilon \to 0} \frac{F(u+\epsilon x) - F(u)}{\epsilon} = DF(u)x.$$

Thus, F is Gâteaux differentiable at u with $dF[u]x = DF(u)x$ for all $x \in X$, as claimed. $\square$

We now list several important facts, all of which should be familiar from classical analysis.

Important Differentiation Facts:

(1) If F is Fréchet differentiable at $u \in U$, then F is continuous at u.

(2) In general Gâteaux differentiability does not imply Fréchet differentiability. For example, consider the function $F : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$F(x,y) = \begin{cases} \dfrac{x^3 y}{x^6 + y^2}, & \text{if } (x,y) \neq (0,0) \\ 0, & \text{if } (x,y) = (0,0). \end{cases}$$

Then you can easily check that $dF[(0,0)](a,b) = 0$ for all $(a,b) \in \mathbb{R}^2$, and F is Gâteaux differentiable and its Gâteaux derivative is linear and bounded at $(0,0)$. However, F is not Fréchet differentiable at $(0,0)$ since F is not continuous there.

(3) If $dF[v](\cdot) \in \mathcal{L}(X,Y)$ for all v in a neighborhood V of $u \in X$, and if

$$dF : V \to \mathcal{L}(X,Y)$$

is continuous, then F is Fréchet differentiable at u.

(4) Both Fréchet and Gâteaux differentiation are linear processes. Further, if F and G are Fréchet differentiable, so is the composition[4] $F \circ G$ with

$$D(F \circ G)(u) = DF(G(u)) \circ DG(u).$$

Note, the last part (i.e. the "Chain Rule") is not true if F and G are only Gâteaux differentiable.

[4]Clearly, I only meant this to be true when the composition is well-defined.

(5) The Gâteaux derivative of F at u need not be linaer, i.e. it may be that

$$dF[u](x+y) \neq dF[u]x + dF[u]y.$$

For example, consider the function $F : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$F(x,y) = \begin{cases} \dfrac{x^2 y}{x^4 + y^2}, & \text{if } (x,y) \neq (0,0) \\ 0, & \text{if } (x,y) = (0,0). \end{cases}$$

Now, we address how to find critical points of real-valued functionals $F : X \to \mathbb{R}$.

Theorem 2. *Suppose $F : U \subset X \to \mathbb{R}$ is Gâteaux differentiable at $u \in X$ and that u is a local maximum or minimum of F. Then*

$$dF[u]x = 0 \quad \forall x \in X.$$

Proof. The proof is exactly as for the finite dimensional case. Suppose u is a local minimum of F so that for all $x \in X$ we have

$$F(u + \epsilon x) \geq F(u)$$

for all $\epsilon \neq 0$ sufficiently small. In particular, we have

$$dF[u]x = \lim_{\epsilon \to 0^+} \frac{F(u+\epsilon x) - F(u)}{\epsilon} \geq 0$$

and, similarly,

$$dF[u]x = \lim_{\epsilon \to 0^-} \frac{F(u+\epsilon x) - F(u)}{\epsilon} \leq 0,$$

which completes the proof. $\square$

Thus, as expected, critical points of maps between Banach spaces can be found by seeking zeroes of the derivative. However, oftentimes in applications we are interested in critical points subject to particular constraints (for example, conservation of energy or mass). To identify such constrained critical points, let $G : X \to \mathbb{R}$ be C^1 on X, i.e. assume the map

$$DG : X \to \mathcal{L}(X, \mathbb{R})$$

is continuous, and define the admissible set

$$\mathcal{A} = \{u \in X : G(u) = 0\}$$

and assume that $\mathcal{A} \neq \emptyset$. The next result is an infinite dimensinoal generalization of what you know from multi-variable calculus: in fact, the rigorous proof is essentially the same. Nevertheless, for completeness I conclude a full proof.

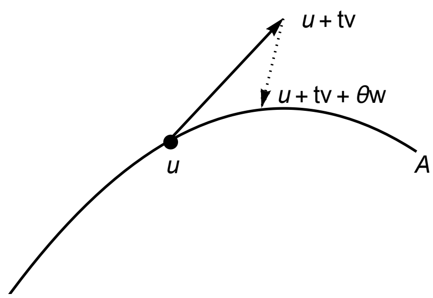

Figure 1: A schematic drawing of the constrution of the local curve through u within the constraint set $\mathcal{A}$. Note that since $DG(u)w \neq 0$, the vector w has a non-trivial normal component to the surface $\mathcal{A}$ near u.

Theorem 3 (Lagrange Multiplier Theorem)**.** *Let F and $\mathcal{A}$ be as above. If $u \in \mathcal{A}$ is a local minimum or maximum of $F|_{\mathcal{A}}$, and if the Gâteaux derivative of F at u exists and is linear, and if $DG(u) \neq 0$, then there exists a $\lambda \in \mathbb{R}$ such that*

$$dF[u]x = \lambda DG(u)x$$

for all $x \in X$. The constant λ is known as a Lagrange multiplier.

Before we present the rigorous proof, note that the big idea is construct first construct a local, one-dimensional curve $\gamma(t) \in X$, defined for $|t| \ll 1$, say for $t \in (-\epsilon, \epsilon)$, such that

(i) $\gamma(0) = u$, so that γ passes through u, and

(ii) $\gamma(t) \in \mathcal{A}$ for all $|t| < \epsilon$, so the curve is contained within the constraint set $\mathcal{A}$.

It then follows by construction that the composition $F \circ \gamma : (-\epsilon, \epsilon) \to \mathcal{A}$ has a local max or min at $t = 0$, which implies that

$$\frac{d}{dt}\bigg|_{t=0} F(\gamma(t)) = dF[u]\gamma'(0) = 0.$$

As we will see in the rigorous proof, the result will follow from above once we have an appropriate description of the curve γ and the tangent vector $\gamma'(0)$.

Proof. First, recall from Differential Geometry that the tangent space of $\mathcal{A}$ at the point $u \in \mathcal{A}$ is exactly the kernel of the derivative $DG(u)$. Thus, the assumption that $DG(u) \neq 0$ implies that there is some vector $w \in X$ that is *not tangent to $\mathcal{A}$ at u*. Now, fix a $v \in X$ such that $dF[u]v$ exists, we want to prove that for each $t \neq 0$ sufficiently small there exists a constant $\theta = \theta(t) > 0$ such that

$$u + tv + \theta(t)w \in \mathcal{A},$$

i.e. for each point $u + tv$ near $u \in \mathcal{A}$, there is a small multiple of w which puts the point back in $\mathcal{A}$. See Figure 1 for a schematic. To prove the existence of such a function, define $j : \mathbb{R}^2 \to \mathbb{R}$ by

$$j(t, \theta) = G(u + tv + \theta w)$$

and note, by assumption, that $j \in C^1(\mathbb{R}^2)$ with

$$j(0, 0) = 0 \quad \text{and} \quad \frac{dj}{d\theta} = DG(u)w \neq 0.$$

The Implicit Function Theorem then implies there exists a C^1-function $\theta : \mathbb{R} \to \mathbb{R}$ such that

$$\theta(0) = 0 \quad \text{and} \quad j(t, \theta(t)) = 0 \;\; \forall |t| \ll 1,$$

as desired[5].

Now, since $u \in \mathcal{A}$ was assumed to be a local maximum or minimum of $F|_{\mathcal{A}}$, it follows that

$$0 = \frac{d}{dt} F(u + tv + \theta(t)w) = dF[u] \left(v + \theta'(0)w \right).$$

The result now follows by observing that since $j(t, \theta(t)) = 0$ for $|t| \ll 1$ that

$$DG(u)v + \theta'(0) DG(u)w = 0$$

which, combined with above, gives

$$dF[u]v = \underbrace{\left(\frac{dF[u]w}{DG(u)w} \right)}_{=:\lambda} DG(u)v,$$

as claimed. $\square$

3 Application: A Linear Example

With the above abstract machinery in place, we now return to the motivating example we originally considered in Section 1.1. At the end of this section, we will have proven the main claim in Section 1.1.

3.1 Setup of the Problem

To begin, let $\Omega \subset \mathbb{R}^n$ be open and bounded and define the functional $F : H_0^1(\Omega) \to \mathbb{R}$ by

$$F(u) = \int_\Omega |Du|^2 dx.$$

[5] Note that the map

$$\mathbb{R} \ni t \mapsto u + tv + \theta(t)w \in \mathcal{A}$$

is precisely the rigorous description of the curve $\gamma(t)$ used in the above "big idea" discussion!

Clearly the minimum value of F over all of $H_0^1(\Omega)$ is zero, which is achieved precisely by the constant function $u = 0$. To make the problem more interesting, lets consider the minimum of this functional subject to the additional constraint that $\|u\|_{L^2(\Omega)} = 1$. That is, if we define $G : H_0^1(\Omega) \to \mathbb{R}$ by

$$G(u) = \int_\Omega |u|^2 dx - 1$$

and define the admissible set[6]

$$\mathcal{A} := \left\{ u \in H_0^1(\Omega) : G(u) = 0 \right\},$$

then we want to consider the constrained minimization problem of proving that

$$\min_{u \in \mathcal{A}} F(u)$$

exists as a real number. To begin our study of this problem, we start with the following observation.

Lemma 1. *If $u \in \mathcal{A}$ is a local minimum of $F|_{\mathcal{A}}$, then there exits a $\lambda \in \mathbb{R}$ such that*

$$-\Delta u = \lambda u$$

weakly in $H_0^1(\Omega)$.

Proof. By the Lagrange Multiplier Theorem (i.e. Theorem 3), if $u \in H_0^1(\Omega)$ is a local minimum of $F|_{\mathcal{A}}$ then there exists a $\lambda \in \mathbb{R}$ such that

$$dF[u]v = \lambda DG(u)v$$

for all $v \in H_0^1(\Omega)$. To calculate these derivatives, note by definition that if $v \in H_0^1(\Omega)$ then

$$\begin{aligned} dF[u]v &= \lim_{\epsilon \to 0} \frac{F(u + \epsilon v) - F(u)}{\epsilon} \\ &= \lim_{\epsilon \to 0} \epsilon^{-1} \int_\Omega \left(|D(u + \epsilon v|^2 - |Du|^2 \right) dx \\ &= \lim_{\epsilon \to 0} \epsilon^{-1} \int_\Omega \left(2\epsilon Du \cdot Dv + \epsilon^2 |Dv|^2 \right) dx \\ &= 2 \int_\Omega Du \cdot Dv \; dx. \end{aligned}$$

while, similarly,

$$dG[u]v = 2 \int_\Omega uv \; dx.$$

Note that since G is Gâteaux differentiable, it follows that $DG(u)v = dG[u]v$ for all $v \in H_0^1(\Omega)$, from which the result now follows directly. $\square$

[6] Note that since $G \in C^1(H_0^1(\Omega; \mathbb{R}))$ it follows that $\mathcal{A}$ is closed.

Naturally, it it remains to actually *prove* that $F|_{\mathcal{A}}$ has a local minimum. To this end, set

$$\lambda := \inf_{u \in \mathcal{A}} F(u)$$

and note, since $F(u) \geq 0$, we know that $\lambda \geq 0$ is clearly finite. By definition of λ, it follows that there exists a "minimizing sequence", i.e. there exists a sequence $\{u_k\}_{k=1}^{\infty}$ in $\mathcal{A}$ such that

$$F(u_k) \to \lambda \text{ as } k \to \infty.$$

Note that **IF** we could extract a subsequence $\{u_{k_j}\}$ that converges in $H_0^1(\Omega)$, to some $u_0 \in H_0^1(\Omega)$, then since F is clearly continuous with respect to convergence in $H_0^1(\Omega)$ we would have

$$F(u_{k_j}) \to F(u_0) = \lambda$$

so that $u_0 \in \mathcal{A}$ would be a global minimizer of $F|_{\mathcal{A}}$. However, by the definition of F we can only conclude that the sequence $\{u_k\}$ is *bounded* in $H_0^1(\Omega)$, which is an infinite dimensional Banach space. Hence, such a convergent subsequence may not exist. Previously when we have ran into this sort of problem, we have relied on the Rellich-Kondrachov Compactness Theorem, which guarantees the existence of a subsequence $\{u_{k_j}\}$ that converges in $L^2(\Omega)$. In the present case, however, this is insufficient since F is not continuous with respect to convergence in $L^2(\Omega)$. It follows that to finish the above program, we need a new compactness result. This is developed in the next section.

3.2 A Segue into Weak Convergence

To compensate for the above lack of compactness discussed above, we consider the weak topology on $H_0^1(\Omega)$.

Definition 3. *Let X be a Banach space. We say a sequence $\{u_k\}$ in X converges weakly to $u \in X$, written*

$$u_k \rightharpoonup u$$

if

$$T(u_k) \to T(u)$$

in $\mathbb{R}$ for every bounded linear functional $T \in X^*$.

Before we continue, we note a few important facts concerning weak convergence.

(1) Strong convergence implies weak convergence, i.e. if $u_k \to u$ in X, then $u_k \rightharpoonup u$ in X. This is clear since if $T \in X^*$ then we have the obvious bound

$$|T(u_k) - T(u)| = |T(u_k - u)| \leq \|T\| \, \|u_k - u\|_X.$$

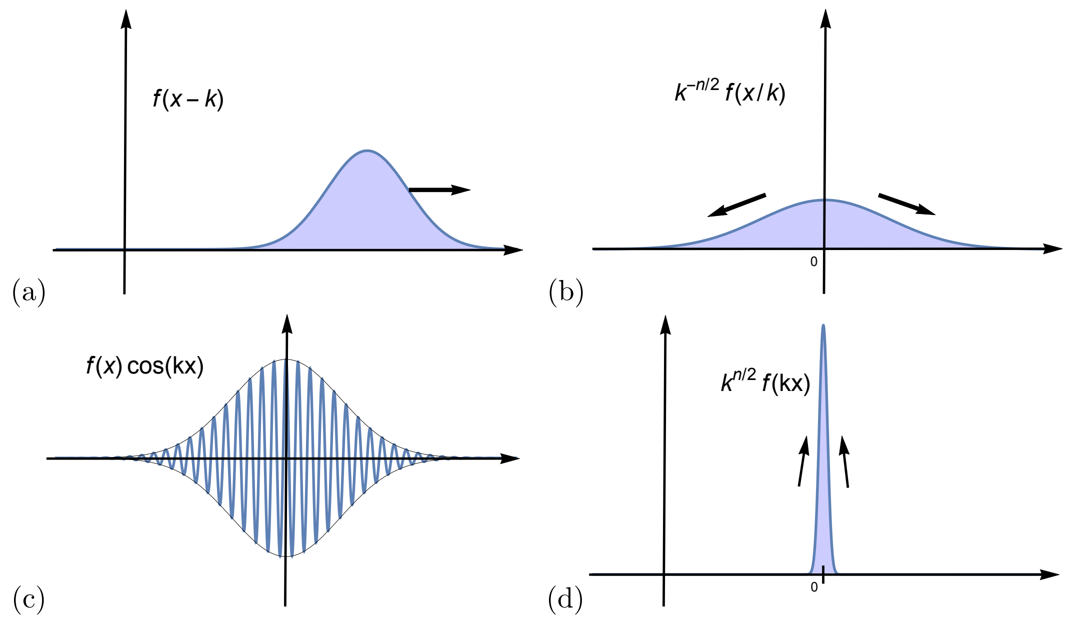

Figure 2: Four typical mechanisms for a bounded sequence in $L^2(\mathbb{R}^n)$ to converge weakly to zero, but not strongly. The mechanisms correspond to (a) "wondering off to infinity", (b) "vanishing", (c) "oscillating to death", and (d) "concentrating".

(2) Weak limits, if they exist, are unique. This is easy to see in the case where X is a Hilbert space. In that case, if both $u \in X$ and $\phi \in X$ are weak limits of $\{u_k\}$, then clearly $T(u-\phi) = 0$ for every $T \in X^*$, which implies $u = \phi$ since

$$X \ni v \mapsto \langle u - \phi, v \rangle_X \in \mathbb{R}$$

is clearly a bounded linear functional on X. In the more general Banach space setting, one must use a Corollary of the Hahn-Banach Theorem that guarantees that linear functionals on Banach spaces *separate points*, i.e. given $x, y \in X$ with $x \neq y$ then there exists a $T \in X^*$ such that $T(x) \neq T(y)$. Equipped with this result, the proof follows by the above considerations.

Example 1. *It is natural to ask how different,* really *is weak convergence from strong convergence. It turns out that, at least in L^2, there are generally four basic mechanisms by which a sequence can converge weakly but not strongly. We illustrate these mechanisms below.*

(a) Sequences may "wonder off to infinity": see Figure 2(a). As an example, fix $\phi \in C_c^\infty(\mathbb{R}^n)$ and note that, since Lebesgue measure is invariant with respect to translations, we clearly have

$$\|\phi(\cdot - k)\|_{L^2(\mathbb{R}^n)} = \|\phi\|_{L^2(\mathbb{R}^n)}$$

for all $k \in \mathbb{N}$ so that, in particular, $\phi(\cdot - k)$ does not converge to zero in $L^2(\mathbb{R}^n)$. Nevertheless, I claim that

$$\phi(\cdot - k) \rightharpoonup 0$$

in $L^2(\mathbb{R}^n)$ as $k \to \infty$. To see this, recall that since $L^2(\mathbb{R})$ is a Hilbert space we know that every bounded linear functional T on $L^2(\mathbb{R}^n)$ is of the form

$$T(v) = \int_{\mathbb{R}^n} fv \; dx$$

for some $f \in L^2(\mathbb{R}^n)$. Now, notice that if $v \in C_c^\infty(\mathbb{R}^n)$ is fixed then there exists a $K \in \mathbb{N}$ such that

$$\mathrm{spt}(\phi(\cdot - k)) \cap \mathrm{spt}(v) = \emptyset$$

for all $k \geq K$ and hence

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} \phi(x-k)v(x) \; dx = 0$$

for all $v \in C_c^\infty(\mathbb{R}^n)$. Weak convergence in $L^2(\mathbb{R}^n)$ now follows by the density of $C_c^\infty(\mathbb{R}^n)$ in $L^2(\mathbb{R}^n)$. Indeed, given $v \in L^2(\mathbb{R}^n)$ let $\varepsilon > 0$ be given and note we can find a function $g \in C_c^\infty(\mathbb{R}^n)$ such that $\|v - g\|_{L^2(\mathbb{R}^n)} < \varepsilon$. Thus, for all $k \in \mathbb{N}$ we have

$$\left| \int_{\mathbb{R}^n} \phi(x-k)v(x) \; dx \right| \leq \|\phi(\cdot - k)\|_{L^2(\mathbb{R}^n)} \|v - g\|_{L^2(\mathbb{R}^n)} + \left| \int_{\mathbb{R}^n} \phi(x-k)g \; dx \right|$$

which, by above, immediately implies that

$$\limsup_{k\to\infty} \left| \int_{\mathbb{R}^n} \phi(x-k)v(x) \; dx \right| \leq \|\phi\|_{L^2(\mathbb{R}^n)} \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, it follows that $\int_{\mathbb{R}^n} \phi(x-k)v(x) \; dx = 0$ for all $v \in L^2(\mathbb{R}^n)$, i.e. $\phi(\cdot - k) \rightharpoonup 0$ in $L^2(\mathbb{R}^n)$, as claimed.

(b) Sequences may "vanish": see Figure 2(b) As an example, fix $f \in C_c^\infty(\mathbb{R}^n)$ and define the sequence of dilations

$$f_k(x) = k^{-n/2} f(x/k).$$

and note that clearly the sequence $\{f_k\}$ is bounded in $L^2(\mathbb{R}^n)$. Furthermore, for each $p \geq 2$ we have

$$\|f_k\|_{L^p(\mathbb{R}^n)} = k^{-n/2} \left(\int_{\mathbb{R}^n} |f(x/k)|^p dx \right)^{1/p} = k^{n/p - n/2} \|f\|_{L^p(\mathbb{R}^n)}$$

In particular, $\|f_k\|_{L^2(\mathbb{R}^n)} = \|f\|_{L^2(\mathbb{R}^n)}$ for all $k \in \mathbb{N}$, while $\|f_k\|_{L^p(\mathbb{R}^n)} \to 0$ as $j \to \infty$ for all $p > 2$. In this example, the sequence of functions spreads out everywhere in space which, since the mass is conserved, locally the mass must tend to zero. This illustrates the "vanishing" part of the trichotomy in Theorem 9.

(c) *Sequences may "oscillate to death": see Figure 2(c) As an example, for each $k \in \mathbb{N}$ set $\phi_k(x) = \cos(kx)$ and note that although $\{\phi_k\}$ is bounded in $L^2(0, 2\pi)$ the fact that*

$$\|\phi_j - \phi_k\|_{L^2(0,2\pi)} = \sqrt{2\pi} \quad \text{for all } j \neq k$$

implies that it has no convergent subsequence. Nevertheless, I claim that

$$\phi_k \rightharpoonup 0$$

in $L^2(0, 2\pi)$. Indeed, recall that given a $f \in L^2(0, 2\pi)$ we can express f as a Fourier series. In particular, we have

$$f(x) = \frac{1}{2\pi} \int_0^{2\pi} f(z)dz + \frac{1}{\pi} \sum_{k=1}^{\infty} a_k \cos(kx) + b_k \sin(kx)$$

in $L^2(0, 2\pi)$, where here $a_k = \langle f, \cos(k\cdot)\rangle_{L^2(0,2\pi)}$ and $b_k = \langle f, \sin(k\cdot)\rangle_{L^2(0,2\pi)}$. In particular, one has (using the bi-orthogonality of the set $\{\cos(kx), \sin(jx)\}_{k=0,j=1}^{\infty}$)

$$\int_0^{2\pi} f(x)^2 dx = \left(\frac{1}{2\pi} \int_0^{2\pi} f(z)dz \right)^2 + \sum_{k=1}^{\infty} a_k^2 + b_k^2$$

so that, in particular,

$$\sum_{k=1}^{\infty} \left| \langle f, \phi_k \rangle_{L^2(0,2\pi)} \right|^2 < \infty.$$

It follows that $\lim_{k\to\infty} \langle f, \phi_k\rangle_{L^2(0,2\pi)} = 0$, which since $f \in L^2(0, 2\pi)$ was arbitrary, implies $\phi_k \rightharpoonup 0$ in $L^2(0, 2\pi)$, as claimed.

(d) *Sequences may "concentrate": see Figure 2(d). As an example, fix $\phi \in C_c^\infty(\mathbb{R})$ and set*

$$\phi_k(x) = k^{1/2}\phi(kx)$$

for each $k \in \mathbb{N}$. This example is essentially "dual" to the vanishing example: now, eventually we find all the mass of function in an arbitrarialy small neighborhood of the origin. This duality is readily seen through the Fourier transform, as functions which "spread out" in the spatial side will "concentrate" on the Fourier side. In this case, you can similarly show that $\{\phi_k\}$ is bounded in $L^2(\mathbb{R}^n)$ with $\|\phi_k\|_{L^2(\mathbb{R}^n)} = \|\phi\|_{L^2(\mathbb{R}^n)}$, but that $\phi_k \rightharpoonup 0$ in $L^2(\mathbb{R}^n)$ as $k \to \infty$.

A few more remarks are in order regarding the above example. First, note that "translating" and "vanishing" can not occur on bounded domains. Furthermore, a crucial difference between the four mechanisms is their behavior under different L^p norms. In particular, observe that in the vanishing example we have $\|f_j\|_{L^p(\mathbb{R}^n)} \to 0$ when $p > 2$ and $\|f_j\|_{L^p(\mathbb{R}^n)} \to \infty$ for $1 \leq p < 2$, while in the "concentration" case the scenario is exactly reversed. In the "oscillating" and "translating" cases, however, all L^p norms are preserved. This observation suggests that one may be able to use different L^p norms to discern between “vanishing” and “concentrating”, while such an approach can not distinguish between “oscillating” and ”translating”. Finally, we note that while all four mechanisms correspond to bounded sequences in $L^2(\mathbb{R}^n)$, only “translating” and “vanishing” are also bounded sequences in $H^1(\mathbb{R}^n)$. In particular, if one were to consider rather bounded sequences in $H^1(\mathbb{R}^n)$ then two of the four mechanisms listed above no longer apply. This motivates the famous Concentration Compactness Theorem, which classifies all possible behaviors of bounded sequences in $H^1(\mathbb{R}^n)$: see Theorem 9 in Section 6.1 below.

The above example illustrates that in $L^2(\Omega)$, bounded sequences may have weak limits in $L^2(\Omega)$ even when they don't have strongly convergent subsequences. This “weak compactness” result holds much more generally, as the following result shows.

Theorem 4 (Banach-Alaoglu)**.** *Let X be a reflexive Baanch space*[7]*, and suppose that $\{u_k\}$ is a bounded sequence in X. Then there exists a subsequence $\{u_{k_j}\}$ and a $u \in X$ such that*

$$u_{k_j} \rightharpoonup u \ \text{ in } X.$$

For a proof of the above result, take Math 960. It is worth noting that if X is not reflexive, then the above result must be stated in terms of the weak-$\star$ topology on X. In reflexive spaces, however, weak and weak-$\star$ convergence is equivalent.

Before continuing, we must make an important warning. Mainly, nonlinearities are generally not continuous with respect to weak convergence. This is illustrated by the following example.

Example 2. *Recall from Example 1(c) that the sequence $\phi_k(x) = \cos(kx)$ is bounded in $L^2(0, 2\pi)$ and that, further, while it has no convergent subsequence in $L^2(0, 2\pi)$ is does satisfy*

$$\phi_k \rightharpoonup 0 \ \text{ in } L^2(0, 2\pi).$$

Now, note that while the nonlinear function

$$F : L^2(0, 2\pi) \to L^2(0, 2\pi), \quad F(g) = g^2$$

is well-defined by Cauchy-Schwartz, we clearly have

$$F(\phi_k(x)) = \frac{1}{2}\left(1 + \cos(2kx)\right) \rightharpoonup \frac{1}{2} \ \text{ in } L^2(0, 2\pi).$$

This shows that the nonlinear map F is not continuous (in the weak topology) with respect to weak convergence in $L^2(0, 2\pi)$. To further illustrate the issue, note that while $\phi_k \rightharpoonup 0$ in $L^2(0, 2\pi)$ we have

$$\int_0^{2\pi} F(\phi_k)dx = \frac{1}{2}$$

[7] That is, $(X^*)^* = X$. Note, specifically, that $W^{k,p}(\Omega)$ is reflexive for all $k \in \mathbb{N} \cup \{0\}$ and $1 < p < \infty$. Also, note since Hilbert spaces are self dual, every Hilbert space is clearly reflexive.

for all $k \in \mathbb{N}$, and hence the real-valued, nonlinear functional

$$L^2(0,2\pi) \ni v \mapsto \int_0^{2\pi} F(v)dx \in \mathbb{R}$$

is not continuous with respect to weak-convergence in $L^2(0,2\pi)$.

The above warns us that although Banach-Alaoglu gives us a compactness result, the fact that it is only with respect to the weak topology means additional work will have to be done in order to be able to use it successfully in the study of *nonlinaer* PDE. As we will see below, there are two additional tricks that are needed: one is to observe that in our applications Banach-Alaoglu can be coupled to the Rellich-Kondrachov compactness result, while the other is to note that while nonlinear functionals are not continuous with respect to weak convergence, we can still control their behavior in some sense. These additional tricks will be illustrated in the next section.

3.3 Back to our Linear Example

Now, returning to our example, recall we let $\Omega \subset \mathbb{R}^n$ be open and bounded and set

$$F(u) = \int_\Omega |Du|^2 dx, \quad u \in H_0^1(\Omega)$$

and

$$\mathcal{A} = \left\{ u \in H_0^1(\Omega) : \int_\Omega u^2 \ dx = 1 \right\}.$$

We have seen that if $u \in H_0^1(\Omega)$ is a local min of $F|_{\mathcal{A}}$ then three exists a $\lambda \in \mathbb{R}$ such that $-\Delta u = \lambda u$ weakly in $H_0^1(\Omega)$. Further, to prove such a local min exists, we set

$$\lambda = \inf_{u \in \mathcal{A}} F(u)$$

and noted that F is bounded below on $\mathcal{A}$ and, in particular, $\lambda \geq 0$. By definition, there exists a minimizing sequence $\{u_k\}$ in $\mathcal{A}$ such that

$$F(u_k) \to \lambda \ \text{ as } k \to \infty$$

and that, since F and the constraint set $\mathcal{A}$ together control the H^1 norm, we know the sequence $\{u_k\}$ is bounded in $H_0^1(\Omega)$. Since $H_0^1(\Omega)$ is reflexive, it now follows by Banach-Alaoglu that there exists a subsequence $\{u_{k_j}\}$ and a function $\phi \in H_0^1(\Omega)$ such that $u_{k_j} \rightharpoonup \phi$ in $H_0^1(\Omega)$, i.e.

$$u_{k_j} \rightharpoonup \phi \ \text{ in } L^2(\Omega)$$

and

$$Du_{k_j} \rightharpoonup D\phi \ \text{ in } L^2(\Omega; \mathbb{R}^n).$$

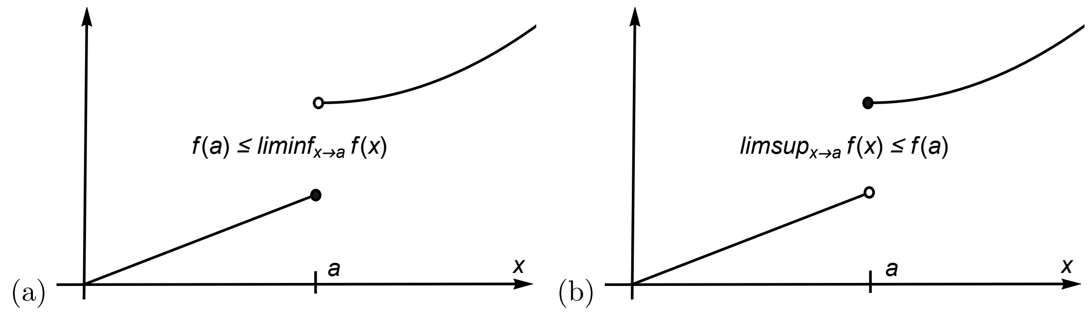

Figure 3: Grahpical depictions of (a) lower-semicontinuity and (b) upper semicontinuity of a function $f : \mathbb{R} \to \mathbb{R}$ at $x = a$.

Moreover, since weak limits are unique, and since norm convergence implies weak convergence, Rellich-Kondrachov implies that, after possibly passing to another subsequence, we can assume that

$$u_{k_j} \to \phi \ \text{ in } L^2(\Omega).$$

In particular, it follows that $\|\phi\|_{L^2(\Omega)} = 1$ so that $\phi \in \mathcal{A}$. This illustrates the power of coupling the weak convergence result in Banach-Alaoglu to the normed convergence result in Rellich-Kondrachov.

It is natural to expect that ϕ is a global minima of $F|_{\mathcal{A}}$, which would be true if we can show that

$$F(\phi) = \lambda.$$

Unfortunatley, as we saw in the previous example, nonlinear functionals are not in general continuous with respect to weak convergence in $H^1(\Omega)$. Nevertheless, we do have the following result.

Lemma 2. *Let $\{u_j\}$ be a sequence in $H^1_0(\Omega)$, and suppose there exists a $\phi \in H^1_0(\Omega)$ such that*

$$u_j \rightharpoonup \phi \ \text{ in } H^1_0(\Omega).$$

Then if $F(u) := \int_\Omega |Du|^2 dx$ we have

$$F(\phi) \leq \liminf_{j\to\infty} F(u_j),$$

i.e. the functional F is weakly lower semicontinuous with respect to weak convergence in $H^1_0(\Omega)$.

Remark 1. *Recall that given a function $f : \mathbb{R} \to \mathbb{R}$, we say that f is lower (respectively. upper) semicontinuous if the function value can not jump up (respectively, down) at discontinuities of f: see Figure 3. This same notion applies to general maps $F : X \to \mathbb{R}$ defined on metric spaces X, as well as to general topological spaces X (obviously, with some modifications).*

Proof. Note that with $\phi \in H_0^1(\Omega)$ as above, the mapping

$$\eta \mapsto \int_\Omega D\phi \cdot D\eta \; dx$$

is a bounded linear functional on $H_0^1(\Omega)$, so that, by the definition of weak convergence,

$$\lim_{j\to\infty} \int_\Omega D\phi \cdot Du_j \; dx = \int_\Omega |D\phi|^2 = F(\phi).$$

Since Cauchy-Schwartz also gives the bound

$$\int_\Omega D\phi \cdot Du_j \; dx \leq \left(\int_\Omega |D\phi|^2 dx\right)^{1/2} \left(\int_\Omega |Du_j|^2 dx\right)^{1/2} = F(\phi)^{1/2} F(\phi_j)^{1/2},$$

it follows that

$$F(\phi) \leq F(\phi)^{1/2} \liminf_{j\to\infty} F(\phi_j)^{1/2},$$

which is equivalent to the desired result. $\square$

Using above lemma, it follows that

$$\lambda \leq F(\phi) \leq \liminf_{j\to\infty} F(u_{k_j}) = \lambda \tag{3}$$

so that λ is indeed a minimum of $F|_{\mathcal{A}}$ with minimizer ϕ. Note the first inequality in (3) follows by the definition of λ and the fact that ϕ is *some* element of $\mathcal{A}$, while the last equality follows by the virtue of $\{u_{k_j}\}$ being a minimizing sequence for $F|_{\mathcal{A}}$.

In fact, much more can be said concerning the above problem. Recall by the Lagrange multiplier theorem that there exists a $\mu \in \mathbb{R}$ such that

$$\int_\Omega D\phi \cdot Dv \; dx = \mu \int_\Omega \phi v \; dx$$

for all $v \in H_0^1(\Omega)$, i.e. $\phi \in H_0^1(\Omega)$ is a weak solution to the BVP

$$\begin{cases} -\Delta\phi = \mu\phi, & \text{in } \Omega \\ u = 0, & \text{on } \partial\Omega. \end{cases}$$

In fact, taking $v = \phi$ above we find that

$$F(\phi) = \mu \int_\Omega \phi^2 \; dx = \mu$$

so that $\mu = \lambda$, i.e. the global minimum of $F|_{\mathcal{A}}$ is an eigenvalue of $-\Delta$ on $H_0^1(\Omega)$ with eigenfunction ϕ.

Furthermore, we can show that λ is the ground state eigenvalue since if $\tilde{\lambda} \in \mathbb{C}$ is any other $H_0^1(\Omega)$-eigenvalue with eigenfunction $\tilde{\phi}$ satisfying $\|\tilde{\phi}\|_{L^2(\Omega)} = 1$, then

$$\int_\Omega D\tilde{\phi} \cdot Dv \, dx = \tilde{\lambda} \int_\Omega \tilde{\phi} v \, dx$$

for all $v \in H_0^1(\Omega)$. Taking $v = \tilde{\phi}$ then gives

$$F(\tilde{\phi}) = \tilde{\lambda}$$

which, by above, implies that $\lambda \leq \tilde{\lambda}$, as claimed.

4 Genearl Theory

While the above analysis was pertaining to a linear PDE, much of the same ideas carry through when applying the variational method to more general nonlinear PDE. The goal of this section is to provide an abstract framework, using the example in Section 3 as a guide, that will apply to more general constrained variational problems: in particular, to ones coming from nonlinear PDE.

To this end, suppose X is a reflexive Banach space and let

$$F, G : X \to \mathbb{R}$$

be (possibly nonlinear) functionals on X, with G being C^1. We are interested in developing general hypotheses under which we can answer the following.

Main Question: If $\mathcal{A} := \{u \in X : G(u) = 0\}$, when are we guaranteed that the problem

$$\lambda = \inf_{u \in \mathcal{A}} F(u)$$

has a minimizer in $\mathcal{A}$?

First, we clearly must require that F be bounded below on $\mathcal{A}$, else $\lambda = -\infty$. If F is bounded below on $\mathcal{A}$, then λ is finite and hence there exists a minimizing sequence $\{u_k\} \subset \mathcal{A}$ satisfying

$$F(u_k) \to \lambda.$$

To ensure that $\{u_k\}$ is bounded in X, it is enough to assume that F is "coercive" on $\mathcal{A}$, i.e. that

$$F(u) \to \infty \quad \text{as} \quad \|u\|_X \to \infty \quad \text{and} \quad u \in \mathcal{A}.$$

Indeed, since the sequence $\{F(u_k)\}$ converges in $\mathbb{R}$ is must be bounded, and hence coercivity of F would immediately imply the sequence $\{u_k\}$ must be bounded in X. Note in the example in Section 3, the functional $F(u) = \int_\Omega |Du|^2 dx$ was clearly coercive on all of $H_0^1(\Omega)$ by Poinceré, and hence was necessarily coercive on the constraint set $\mathcal{A}$.

Continuing, if F is coercive, then $\{u_k\}$ is bounded in the reflexive Banach space X. Consequently, Banach-Alaoglu implies there exists a $\phi \in X$ and subsequence $\{u_{k_j}\}$ such that

$$u_{k_j} \rightharpoonup \phi \text{ in } X.$$

Our goal is now to show that ϕ is our desired minimizer. To this end, note that if F is additionally weakly lower-semicontinuous on X, then we would have

$$F(\phi) \leq \liminf_{j\to\infty} F(u_{k_j}) = \lambda.$$

To conclude that ϕ is actually a minimizer, it remains to show that $\phi \in \mathcal{A}$ since then, in that case, we would additionally have

$$\lambda \leq F(\phi)$$

which implies by above that $F(\phi) = \lambda$ as desired.

Now, to establish that $\phi \in \mathcal{A}$, observe that since $G(u_{k_j}) = 0$ for all j, it would be sufficient to show that

$$\lim_{j\to\infty} G(u_{k_j}) = G(\phi),$$

i.e. it would be sufficient to have G be continuous with respect to weak convergence in X. Taken together, the above considerations lead us to the following general result.

Theorem 5. *Let X be a reflexive Banach space, and let $F, G : X \to \mathbb{R}$ be (possibly nonlinear) functionals with G being C^1. Set*

$$\mathcal{A} := \{u \in X : G(u) = 0\}.$$

If F is bounded below, coercive and weakly lower semicontinuous on $\mathcal{A}$, and if G is continuous with respect to weak convergence in X, then the problem

$$\lambda = \inf_{u\in\mathcal{A}} F(u)$$

has a minimizer in $\mathcal{A}$.

It is important to note that, unfortunately, weak convergence generally does not interact well with nonlinear maps[8]. Consequently, one should not generally expect such a weak continuity (even up to subsequences) to hold for nonlinear maps. In applications of the variational method, however, G is typically of the form

$$G(u) = \|u\|_Y - 1$$

[8]For example, let $X = L^2(0, 2\pi)$ and $G(u) = \|u\|_{L^2(0,2\pi)}$. If $f_n(x) = \sin(nx)$ then $f_n \rightharpoonup 0$ in $L^2(0, 2\pi)$ but $G(f_n) = \sqrt{\pi}$ for all n.

for some Banach space Y with X continuously embedded in Y. For example, in Section 3 we had $X = H_0^1(\Omega)$ and $Y = L^2(\Omega)$. The following result shows that if the embedding $X \subset Y$ is compact, so that the identity map $i : X \to Y$ is a compact linear operator between Banach spaces, then the nonlinear functional G is continuous with respect to weak convergence.

Lemma 3. *Suppose X and Y be Banach spaces and suppose that $T \in \mathcal{L}(X,Y)$ is a compact operator, i.e. it maps bounded sets in X to pre-compact sets[9] in Y. If $\{u_k\}$ is a sequence in X that converges weakly to some point $u \in X$, then we have*

$$\lim_{k\to\infty} T(u_k) = T(u).$$

That is, compact operators map weakly convergent sequences to strongly convergent sequences.

Consequently, in many applications the weak continuity of the constraint G is equivalent to a topological compatness problem.

5 Application to a Semilinear Elliptic BVP

In this section, we apply the general theory discussed in Section 4 to prove the existence of solutions to a nonlinear elliptic BVP. Specifically, let $\Omega \subset \mathbb{R}^n$ with $n \geq 3$ be open and bounded with smooth boundary, and consider the nonlinear elliptic BVP

$$\begin{cases} -\Delta u = |u|^{p-1}u, & \text{in } \Omega,\ p > 1 \\ \quad u = 0, & \text{on } \partial\Omega. \end{cases} \tag{4}$$

In the exercises, you will see that when $p > \frac{n+2}{n-2}$ the BVP (4) does not admit non-trivial classical solutions (at least in the case when $\Omega = B(0,r)$ for some $r > 0$). The goal of this section is to explore the converse of the above non-existence result. Specifically, we will prove the following result.

Theorem 6. *If $n \geq 3$ and $1 < p < \frac{n+2}{n-2}$, then there exists a non-trivial weak solution $u \in H_0^1(\Omega)$ of (4).*

To see this, define $F, G : H_0^1(\Omega) \to \mathbb{R}$ by

$$F(u) = \int_\Omega |Du|^2 dx, \quad G(u) = \int_\Omega |u|^{p+1} dx - 1$$

and define the admissible set

$$\mathcal{A} := \left\{u \in H_0^1(\Omega) : G(u) = 0\right\}.$$

As a first step, we need to make sure G is well-defined on $H_0^1(\Omega)$. This is handled by the following embedding theorem.

[9] In terms of sequences, T is compact if for every bounded sequence $\{x_j\}$ in X there exists a subsequence $\{x_{j_k}\}$ such that $T(x_{j_k})$ converges in Y.

Theorem 7 (Sobolev Embedding Theorem)**.** *If $\Omega \subset \mathbb{R}^n$ is an open (not necessarily bounded) domain in $\mathbb{R}^n$ and if $1 \leq p < n$, then*

$$W_0^{1,p}(\Omega) \subset L^q(\Omega)$$

is a continuous embedding for all $p \leq q \leq \frac{np}{n-p}$. In particular, there exists a constant $C = C(n,p,q) > 0$ such that

$$\|u\|_{L^q(\Omega)} \leq C\|u\|_{W^{1,p}(\Omega)}$$

for all $u \in W_0^{1,p}(\Omega)$.

Remark 2. *Note that when we discussed elliptic regularity theory we previously encountered a Sobolev Embedding Theorem for the case when $p > n$. In that case, we found that when $p > n$ functions in $W^{1,p}(\Omega)$ were continuous and bounded on Ω, and hence have better than expected regularity. The embedding result in Theorem 7 applies to the case when $1 \leq p < n$ and says that functions in $W_0^{1,p}(\Omega)$ have better than expected integrability properties. Further, a discussion of the proof of Theorem 7 is included in Section 8.1 in the Appendix.*

From Theorem 7 it follows that

$$H_0^1(\Omega) \subset L^{p+1}(\Omega) \quad \text{forall} \;\; 1 \leq p \leq \frac{2n}{n-2} - 1 = \frac{n+2}{n-2}.$$

and hence G is well defined on $H_0^1(\Omega)$ for all p under consideration in Theorem 6.

Now that we know the problem is well-defined, note that we have already seen in Section 3 that F is Gâteaux differentiable on $H_0^1(\Omega)$ with

$$dF[u]v = 2\int_\Omega Du \cdot Dv \; dx$$

for all $u, v \in H_0^1(\Omega)$. Further, one can check (see the exercises!) that G is C^1 on $H_0^1(\Omega)$ with[10]

$$DG(u)v = (p+1)\int_\Omega |u|^{p-1}uv \; dx$$

for all $u, v \in H_0^1(\Omega)$. Thus, the Lagrange Multipleir Theorem implies that if $\phi \in \mathcal{A}$ is a minimizer of $F|_{\mathcal{A}}$, then there exists a $\lambda \in \mathbb{R}$ such that

$$2\int_\Omega D\phi \cdot Dv \; dx = \lambda(p+1)\int_\Omega |\phi|^{p-1}\phi v \; dx \tag{5}$$

[10] Note by Taylor's Theorem that

$$|u + \epsilon v|^{p+1} = |u|^{p+1} + (p+1)\epsilon|u|^{p-1}uv + \frac{p(p+1)}{2}\epsilon^2|u + \theta v|^{p-1}v^2$$

for some $0 < \theta(x) < \epsilon$.

for all $v \in H_0^1(\Omega)$, that is, ϕ is a non-trivial (why?) weak solution of the BVP

$$\begin{cases} -\Delta u = \mu|u|^{p-1}u, & \text{in } \Omega \\ u = 0, & \text{on } \partial\Omega, \end{cases}$$

where $\mu = \frac{\lambda(p+1)}{2}$. Note that taking $v = \phi$ in (5) gives

$$2\int_\Omega |D\phi|^2 dx = \lambda(p+1)\int_\Omega |\phi|^{p+1}dx$$

so that $\lambda \geq 0$. However, note that if $\lambda = 0$ then $\int_\Omega |D\phi|^2 dx = 0$ so that ϕ is a constant in $H_0^1(\Omega)$, i.e. $\phi = 0$ a.e. in Ω, which is a contradiction (why?). Now rescaling ϕ as

$$w = \mu^{1/(p-1)}\phi$$

it follows that w is a non-trivial weak solution of (4), as desired.

In summary, it follows that the existence of non-trivial weak solutions of (4) can be established by showing that the problem

$$\lambda = \inf_{u \in \mathcal{A}} F(u)$$

has a minimizer in $\mathcal{A}$. To apply Theorem 5 to show that $F|_{\mathcal{A}}$ has a minimizer, we observe the following:

(a) The functional F is clearly bounded below on $H_0^1(\Omega)$.

(b) F is coercive on $H_0^1(\Omega)$. Indeed, by Poincaré we have

$$\int_\Omega u^2 dx \leq C\int_\Omega |Du|^2 dx = CF(u)$$

for some constant $C > 0$, and hence $F(u) \to \infty$ as $\|u\|_{H^1(\Omega)} \to \infty$.

(c) By our work in Section 3, F is weakly lower semicontinuous on $H_0^1(\Omega)$.

It remains to verify that $G : H_0^1(\Omega) \to \mathbb{R}$ is continuous with respect to weak convergence. This follows from the following "general" version of Rellich-Kondrachov.

Theorem 8 (Rellich-Kondrachov)**.** *Assume $\Omega \subset \mathbb{R}^n$ is open and bounded with C^1-boundary, and let $p \geq 1$. Then*

$$W_0^{1,p}(\Omega) \Subset L^q(\Omega)$$

for each

$$\begin{cases} 1 \leq q < \dfrac{np}{n-p}, & \text{if } 1 \leq p < n \\ 1 \leq q < \infty, & \text{if } p = n \\ 1 \leq q \leq \infty, & \text{if } p > n. \end{cases}$$

Proof. We previously established this fact in the case $p = q = 2$. For the general result, see Theorem 1 in Section 5.7 of Evans, or Theorem 5 in Section 6.5 of McOwen. $\square$

In particular, note that taking $p = 2$ and recalling here that $n \geq 3$ implies that

$$H_0^1(\Omega) \Subset L^{p+1}(\Omega) \quad \text{for all } 0 \leq p < \frac{2n}{n-2} - 1 = \frac{n+2}{n-2}.$$

Since $G(u) = \int_\Omega |u|^{p+1} dx - 1$, it follows from Lemma 3 that G is continuous with respect to weak convergence in $H_0^1(\Omega)$. In summary, by Theorem 5, if $1 < p < \frac{n+2}{n-2}$ then there exists a $\phi \in \mathcal{A}$ such that

$$F(\phi) = \min_{u \in \mathcal{A}} F(u)$$

and hence ϕ induces a non-trivial weak solution of the nonlinear elliptic BVP (4), as desired.

6 Concentration Compactness and Applications

In our previous examples, we used some sort of compactness to show that the proposed minimizer lies in the "admissible" set $\mathcal{A}$. Both the linear example in Section 3 and the nonlinear example in Section 5 were posed on bounded domains, which and the weak continuity of the constraint followed by an application of Rellich-Kondrachov as stated in Theorem 8. On unbounded domains, however, Theorem 8 generally fails.

Example 3. *Define $f : \mathbb{R} \to \mathbb{R}$ by*

$$f(x) = \begin{cases} 1 - |x|, & \text{if } 0 \leq |x| < 1 \\ 0, & \text{if } |x| \geq 1. \end{cases}$$

and for each $n \in \mathbb{N}$ set $f_n(x) = f(x - n)$. Then for each $p \in [1, \infty]$ we clearly have that $\{f_n\}$ is bounded in $W^{1,p}(\mathbb{R})$ and, in fact, $f_n \rightharpoonup 0$ in $W^{1,p}(\mathbb{R})$ for $p \in [1, \infty)$. However, since $\|f_n\|_{L^p(\mathbb{R})} = \|f\|_{L^p(\mathbb{R})}$ for all $n \in \mathbb{N}$ it is clear that no subsequence of $\{f_n\}$ can converge to zero in $L^p(\mathbb{R})$ for any $p \in [1, \infty]$.

Remark 3. *If one were to have some control (e.g. uniform decay) of the the sequence near $\pm\infty$, it may be possible to recover some type of norm compactness result. See the exercises.*

In order to compensate for this lack of compactness when working on unbounded domains, it is helpful to have a complete characterization of the difference between weak and strong convergence in $H^1(\mathbb{R}^n)$. This is accomplished by the following seminal result due to Pierre-Louis Lions in 1984.

6.1 Discussion of Concentration Compactness

Recall from the discussion at the end of Section 3.2 that there are in general four main mechanisms by which a bounded sequence in $L^2(\mathbb{R}^n)$ can converge weakly but not strongly. Two of those mechanisms, "oscillating" and "concentrating", do not correspond to bounded sequences in $H^1(\mathbb{R}^n)$. The next result makes this rigorous by giving a complete characterization of the possible behaviors of bounded sequences in $H^1(\mathbb{R}^n)$. Essentially, it says the sequence converges (up to translations), vanishes (in the same sense discussed in Example 1), or else the sequence "splits" into two parts with at least one of the parts "wondering off to infinity".

Theorem 9 (Concentration Compactness). *Let $\{u_k\}$ be a bounded sequence in $H^1(\mathbb{R}^n)$. Then there exists a subsequence $\{u_{k_j}\}$ satisfying one of the following:*

(i) (Convergence of Translates) There exists a sequence $\{y_j\}$ in $\mathbb{R}^n$ such that $\{u_{k_j}(\cdot - y_j)\}$ is convergent in $L^p(\mathbb{R}^n)$ for all[11] *$2 \leq p < \frac{2n}{n-2}$ if $n \geq 2$, and for all $2 \leq p \leq \infty$ if $n = 1$.*

(ii) (Vanishing) $u_{k_j} \to 0$ in $L^p(\mathbb{R}^n)$ for all $2 < p < \frac{2n}{n-2}$ if $n \geq 2$, and for all $2 < p \leq \infty$ if $n = 1$.

(iii) (Splitting) There exists sequences $\{v_j\}$ and $\{w_j\}$ in $H^1(\mathbb{R}^n)$, both bounded in $H^1(\mathbb{R}^n)$, such that

$$|v_j(x)| + |w_j(x)| \leq |u_{k_j}(x)| \quad \text{for } a.e.\ x \in \mathbb{R}^n,$$
$$\operatorname{dist}(\operatorname{spt}(v_j), \operatorname{spt}(w_j)) \to \infty,$$

and, additionally, we have

$$\left\| u_{k_j} - (v_j + w_j) \right\|_{L^p(\mathbb{R}^n)} \to 0$$

for all $2 \leq p < \frac{2n}{n-2}$ if $n \geq 2$, and for all $2 < p \leq \infty$ if $n = 1$ and

$$\liminf_{j \to \infty} \int_{\mathbb{R}^n} \left(|Du_{k_j}|^2 - |Dv_j|^2 - |Dw_j|^2 \right) dx \geq 0.$$

The proof of Theorem 9 is beyond the scope of this course. If you are interested, please let me know and I can provide a reference. We have previously seen both the "Convergence of Translates" and "Vanishing" cases in examples previously. Note that the strict inequality $p > 2$ in the vanishing case is due the discussion[12] immediately following Example 1. Also, note that the new "splitting" possibility is pretty easy to visualize: fix $f, g \in C_c^\infty(\mathbb{R}^n)$ and for each $k \in \mathbb{N}$ set

$$h_k(x) = f(x) + g(x-k) \quad \text{and} \quad w_k(x) = f(x+k) + g(x-k).$$

[11] Here, we interpret $\frac{2n}{n-2} = \infty$ when $n = 2$.

[12] Specifically, recall there we saw that studying the behavior of the sequence under different L^p norms allowed us to differentiate between "concentrating" and "vanishing".

In this case, both $\{h_k\}$ and $\{w_k\}$ are bounded sequences in $H^1(\mathbb{R}^n)$ which both satisfy the "splitting" alternative in the above trichotomy. In particular, note that since h_k and w_k have fixed L^2 norm for k sufficiently large, and since $h_k \to f$ and $g_k \to 0$ pointwise, it is clear that neither sequence can have subsequences which converge in norm.

We now consider two applications of Concentration Compactness. Both examples are in the context of constrained minimization problems. In these contexts, an important observation is the following. Suppose in Theorem 9 that the original sequence $\{u_k\}$ has a fixed L^p norm for some $2 \leq p < \frac{2n}{n-2}$, say, $\|u_k\|_{L^p(\mathbb{R}^n)} = \lambda > 0$ for all k. Then in the "splitting case", since the supports of the v_j and the w_j are eventually far apart (and hence disjoint), we may take

$$\|v_j\|_{L^p(\mathbb{R}^n)} \to \nu, \quad \text{and} \quad \|w_j\|_{L^p(\mathbb{R}^n)} \to \lambda - \nu$$

for some $\nu \in (0, \lambda)$. This observation will be useful below.

6.2 Application to Extremizers of the Sobolev Embedding Theorem

In this first application, we begin by recalling from the Sobolev Embedding Theorem (Theorem 7) that, in $\mathbb{R}^3$, the Sobolev space $H^1(\mathbb{R}^3)$ is continuously embedded in $L^p(\mathbb{R}^3)$ for all $2 \leq p \leq 6$, i.e. for all such p there exists a constant $C = C(p) > 0$ such that

$$\|f\|_{L^p(\mathbb{R}^3)} \leq C\|f\|_{H^1(\mathbb{R}^3)} \quad \text{for all } f \in H^1(\mathbb{R}^3). \tag{6}$$

There are (at least) two natural and important questions one can ask about (6):

(1) First, for a given $p \in [2, 6]$, what is the best constant one can use in (6)? That is, is it possible to identify the smallest constant $C_p > 0$ such that (6) holds for all $f \in H^1(\mathbb{R}^3)$?

(2) Secondly, if the optimal constant $C_p > 0$ can be identified, do there exist extremizers, i.e. does there exist a non-trivial $f = f_p \in H^1(\mathbb{R}^3)$ such that

$$\|f\|_{L^p(\mathbb{R}^3)} = C_p\|f\|_{H^1(\mathbb{R}^3)}$$

for each fixed $p \in [2, 6]$?

To try to answer these questions, fix $p \in [2, 6]$ and note that if $C > 0$ is such that (6) holds for all $f \in H^1(\mathbb{R}^3)$, then

$$\frac{1}{C} \leq \frac{\|f\|_{H^1(\mathbb{R}^3)}}{\|f\|_{L^p(\mathbb{R}^3)}} \quad \text{for all } f \in H^1(\mathbb{R}^3) \setminus \{0\},$$

which, by rescaling $f \mapsto f/\|f\|_{L^p(\mathbb{R}^3)}$, is equivalent to saying that

$$\frac{1}{C} \leq \|f\|_{H^1(\mathbb{R}^3)} \quad \text{for all } f \in H^1(\mathbb{R}^3) \text{ with } \|f\|_{L^p(\mathbb{R}^3)} = 1.$$

It follows that if we set

$$\mathcal{A}_p := \left\{ f \in H^1(\mathbb{R}^3) : \|f\|_{L^p(\mathbb{R}^3)} = 1 \right\}$$

and define

$$\lambda_p = \inf_{f \in \mathcal{A}_p} \|f\|_{H^1(\mathbb{R}^3)}, \tag{7}$$

then $\lambda_p > 0$ and we have $\frac{1}{\lambda_p} \leq C$ for every constant $C > 0$ satisfying (6). In particular,

$$\|f\|_{L^p(\mathbb{R}^3)} \leq \frac{1}{\lambda_p} \|f\|_{H^1(\mathbb{R}^3)} \quad \text{forall } f \in H^1(\mathbb{R}^3)$$

i.e., λ_p^{-1} is the optimal constant in (6), and extremizers of (6) correspond to minimizers of the constrained variational problem (7).

To study the minimization problem (7), we first establish the following result which demonstrates that even if the optimal constant λ_p can be identified, there may not exist extremizers.

Theorem 10. *In the case $p = 2$, we have $\lambda_2 = 1$. Furthermore, in this case there does not exist an extremizer, i.e. the only function $f \in H^1(\mathbb{R}^3)$ with*

$$\|f\|_{L^2(\mathbb{R}^3)} = \|f\|_{H^1(\mathbb{R}^3)}$$

is $f \equiv 0$.

Proof. Note that since we clearly have $\|f\|_{L^2(\mathbb{R}^3)} \leq \|f\|_{H^1(\mathbb{R}^3)}$ for all $f \in H^1(\mathbb{R}^3)$, it follows that $\lambda_2 \geq 1$. To get equality, consider the sequence $\{f_k\}_{k=1}^{\infty}$ studied in Example 1(b), i.e. fix $f \in C_c^{\infty}(\mathbb{R}^n)$ and define

$$f_k(x) = k^{-n/2} f(x/k).$$

Then clearly we have $\|f_k\|_{L^2(\mathbb{R}^n)} = \|f\|_{L^2(\mathbb{R}^n)}$ for all $k \in \mathbb{N}$ and, further,

$$\|f_k\|^2_{H^1(\mathbb{R}^3)} = \|f\|^2_{L^2(\mathbb{R}^3)} + k^{-1/2} \|Df\|^2_{L^2(\mathbb{R}^3)}.$$

Thus, for each $k \in \mathbb{N}$ we have

$$1 \leq \lambda_2^2 \leq \frac{\|f_k\|^2_{H^1(\mathbb{R}^3)}}{\|f_k\|^2_{L^2(\mathbb{R}^3)}} = \frac{\|f\|^2_{L^2(\mathbb{R}^3)} + k^{-1/2} \|Df\|^2_{L^2(\mathbb{R}^3)}}{\|f\|^2_{L^2(\mathbb{R}^3)}}$$

which, taking $k \to \infty$, yields $\lambda_2 = 1$, as claimed. To see there is not an extremizer in this case, note that if $f \in H^1(\mathbb{R}^3)$ is such that $\|f\|_{L^2(\mathbb{R}^3)} = \|f\|_{H^1(\mathbb{R}^3)}$ then we would have $Df = 0$ a.e., that is, f would have to be a constant function. Since the only constant function in $H^1(\mathbb{R}^3)$ is $f = 0$, the claim follows. $\square$

Next, we consider the minimization problem (7) in the case $p \in (2, 6)$. In this case, we must rely on Concentration Compactness.

Theorem 11. *For each $p \in (2,6)$, the variational problem (7) has a minimizer $\phi \in H^1(\mathbb{R}^3)$. In particular, we have*

$$\lambda_p = \|\phi_p\|_{H^1(\mathbb{R}^3)}.$$

Proof. Clearly[13] we have $\lambda_p > 0$, so that there exists a minimizing sequence $\{f_n\}$ such that

$$\|f_n\|_{L^p(\mathbb{R}^3)} = 1 \quad \text{and} \quad \|f_n\|_{H^1(\mathbb{R}^3)} \to \lambda_p.$$

By Concentration Compactness, see Theorem 9 above, there exists a subsequence of $\{f_n\}$ that either converges (up to translates), "vanishes", or "splits". Clearly vanishing can not occur here since, recalling $n = 3$, we have $2 < p < \frac{2n}{n-2} = 6$.

To rule out "splitting", suppose it were possible to "split" the minimizing sequence as in Theorem 9(iii). Specifically, suppose we could decompose $f_n = g_n + h_n$ with

$$\|g_n\|^p_{L^p(\mathbb{R}^3)} \to \mu > 0, \quad \|h_n\|^p_{L^p(\mathbb{R}^3)} \to 1 - \mu > 0$$

such that the supports of g_n and h_n are disjoint (at least for sufficiently large n). Given $\epsilon > 0$ small, it would then follow by the definition of λ_p in (7) that[14] for all n sufficiently large we would have

$$\|g_n\|^2_{H^1(\mathbb{R}^3)} \geq \lambda_p^2 \|g_n\|^2_{L^p(\mathbb{R}^3)} \geq \lambda_p^2 \left(\mu^{2/p} - \frac{\epsilon}{2}\right) > 0 \tag{8}$$

and, similarly,

$$\|h_n\|^2_{H^1(\mathbb{R}^3)} \geq \lambda_p^2 \|h_n\|^2_{L^p(\mathbb{R}^3)} \geq \lambda_p^2 \left((1-\mu)^{2/p} - \frac{\epsilon}{2}\right) > 0. \tag{9}$$

In particular, given $\epsilon > 0$ sufficiently small we have by disjointness of supports that

$$\begin{aligned} \|f_n\|^2_{H^1(\mathbb{R}^3)} + \lambda_p^2 \epsilon &= \|g_n\|^2_{H^1(\mathbb{R}^3)} + \|h_n\|^2_{H^1(\mathbb{R}^3)} + \lambda_p^2 \epsilon \\ &\geq \lambda_p^2 \left[\mu^{2/p} + (1-\mu)^{2/p}\right]. \end{aligned}$$

Since this would hold for all $\epsilon > 0$, it would then follow by taking $n \to \infty$ and $\epsilon > 0$ above that

$$\lambda_p^2 \geq \lambda_p^2 \left[\mu^{2/p} + (1-\mu)^{2/p}\right]. \tag{10}$$

To see the contradiction, observe that since $p > 2$ we know the function $x^{2/p}$ is strictly sub-additive on $(0, \infty)$, i.e.

$$(a+b)^{2/p} < a^{2/p} + b^{2/p} \quad \text{for all } a, b > 0$$

[13] This just follows since by Sovolev embedding we know λ_p has *some* positive lower bound.

[14] Here, we are heavily using that the quotient in (7) is homogeneous, i.e. is invariant under $f \mapsto \gamma f$ for $\gamma \in \mathbb{R}$. Indeed, note that although $g_n, h_n \notin \mathcal{A}$, the stated inequality still holds by applying (7) to $g_n / \|g_n\|_{L^p(\mathbb{R}^3)}$ and $h_n / \|h_n\|_{L^p(\mathbb{R}^3)}$.

and hence, since $\mu \in (0,1)$, it must be true that

$$\mu^{2/p} + (1-\mu)^{2/p} > 1,$$

which clearly contradicts (10). Thus, vanishing can not occur in this case.

By Concentration Compactness, it follows that there exists a sequence $\{y_j\} \in \mathbb{R}^3$ and a subsequence $\{f_{n_j}\}$ such that

$$f_{n_j}(\cdot - y_j) \to \phi \ \text{ in } L^p(\mathbb{R}^3)$$

for all $2 \leq p < 6$ and for some function $\phi \in H^1(\mathbb{R}^3)$. I claim that ϕ is a minimizer of (7). To see this, note we clearly have $\phi \in \mathcal{A}$ (since translations leave the L^p-norm invariant) so that, by definition,

$$\lambda_p \leq \|\phi\|_{H^1(\mathbb{R}^3)}.$$

Further, since the sequence $\{f_{n_j}(\cdot - y_j)\}$ is clearly bounded in $H^1(\mathbb{R}^3)$, Banach-Alaoglu along with uniqueness of weak limits implies that

$$f_{n_j}(\cdot - y_j) \rightharpoonup \phi \ \text{ in } H^1(\mathbb{R}^3).$$

Recalling that $f_{n_j}(\cdot - y_j) \to \phi$ in $L^2(\mathbb{R}^n)$ and that the map

$$H^1(\mathbb{R}^3) \ni v \mapsto \int_{\mathbb{R}^3} |Dv|^2 dx$$

is lower weakly-semicontinuous, it follows that

$$\lambda_p \leq \|\phi\|_{H^1(\mathbb{R}^3)} \leq \liminf_{j\to\infty} \|f_{n_j}(\cdot - y_j)\|_{H^1(\mathbb{R}^3)} = \lambda_p,$$

so that ϕ is indeed a minimizer of (7), as desired. □

One significant advantage in the previous argument is that the quotient in (7) is homogeneous, i.e. it is invariant under the rescaling $f \mapsto \gamma f$ for every $\gamma \in \mathbb{R}$. If that were not the case, things become more complicated as one can not use the variational problem (7) to get (8)-(9). Nevertheless, the ruling out of splitting in nonhomogeneous case still follows by a sub-additivity argument, as demonstrated in the next section.

6.3 Application to the Subcritical, Focusing NLS

We now aim to apply the method of Concentration Compactness to prove the existence and stability of nonlinear bound state solutions to one of the most famous nonlinear dispersive equations.

Let $\alpha > 0$ be fixed and consider the focusing nonlinaer Schrödinger equation

$$iu_t = -\Delta u - |u|^\alpha u, \quad x \in \mathbb{R}^n, \ t \in \mathbb{R} \tag{11}$$

where here u is in general complex valued. This equation is prevalent throughout applied mathematics, arising naturally in several areas including the following:

(1) Statistical (i.e. many particle) quantum mechanics as the temperature nears 0^oK (so-called Bose-Einstein Condesates).

(2) Wave propagation in nonlinear fiber optics. In this context, t is proportional to the longitudinal distance down the wire.

(3) The study of nonlinear "modulates" of gravity water waves in deep water. In this context, one decomposes
$$u(x,t) = A(x,t)e^{i\theta(x,t)}$$
where A models the wave amplitude and θ models the wave phase.

Of particular interest in the study of (11) is the existence and local dynamics of "bound state" solutions of the form
$$u(x,t) = e^{-i\lambda t}\phi(x) \tag{12}$$
where here $\lambda \in \mathbb{C}$ and $\phi \in H^1(\mathbb{R}^n;\mathbb{R})$. Note that u in (12) satisfies (11) if and only if ϕ satisfies the "eigenvalue" problem
$$\begin{cases} -\Delta\phi - |\phi|^\alpha\phi = \lambda\phi \\ \phi \in H^1(\mathbb{R}^n), \quad \lambda \in \mathbb{R} \end{cases} \tag{13}$$
Clearly $\phi = 0$ satisfies (13) for every $\lambda \in \mathbb{R}$. The goal of this section is to use variational methods to prove there exists non-trivial solutions of (13) for some $\lambda \in \mathbb{R}$.

To this end, we define the "energy" and "mass" functionals
$$\mathcal{H}, W : H^1(\mathbb{R}^n) \to \mathbb{R},$$
respectively, by
$$\mathcal{H}(u) = \frac{1}{2}\int_{\mathbb{R}^n} |Du|^2 dx - \frac{1}{\alpha+2}\int_{\mathbb{R}^n} |u|^{\alpha+2} dx$$
and
$$W(u) = \int_{\mathbb{R}^n} |u|^2 dx.$$
First, note that $\mathcal{H}$ is well-defined on $H^1(\mathbb{R}^n)$ provided that
$$H^1(\mathbb{R}^n) \subset L^{\alpha+2}(\mathbb{R}^n)$$
Since $H^1(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ for all $2 \leq p \leq \frac{2n}{n-2}$ if $n = 1$ or $n \geq 3$, and for all $2 \leq p < \infty$ if $n = 2$, a sufficient condition for the above embedding is
$$\alpha + 2 \leq \frac{2n}{n-2}, \quad i.e. \;\; 0 < \alpha \leq \frac{4}{n-2}$$
Consequently, we will assume throughout that $0 < \alpha \leq \frac{4}{n-2}$ if $n = 1$ or $n \geq 3$, or that $0 < \alpha < \infty$ if $n = 2$. We start with the following observation.

Lemma 4. *Let $\mu > 0$ and define the admissible set*

$$\mathcal{A}_\mu := \left\{ u \in H^1(\mathbb{R}^n) : W(u) = \mu \right\}.$$

If the problem

$$I_\mu := \inf_{u \in \mathcal{A}_\mu} \mathcal{H}(u) \tag{14}$$

has a minimizer ϕ, then ϕ is a weak solution of (13) *for some $\lambda = \lambda(\mu)$.*

Proof. This is a direct application of the Lagrange Multiplier Theorem. The calculation is left as an exercise. □

Remark 4. *An important note is that the parameter dependence of the constraint set is used here precisely because the functional $\mathcal{H}$ being minimized is not homogeneous. Specifically, the functional $\mathcal{H}$ being minimized and the constraint functional W do not scale the same way under the mapping $f \mapsto \gamma f$. As we will see, adding this parameter dependence on the constraint set is useful to rule out splitting in our application of concentration compactness.*

It remains to be seen when (14) has a minimizer. Following the general procedure introduced in Section 4, as a first step, we find sufficient conditions for $\mathcal{H}$ to be bounded below on[15] the constraint set $\mathcal{A}_\mu$.

Proposition 1. *Fix*[16] $0 < \alpha \le \frac{4}{n-2}$ *and let I_μ be defined as in* (14) *for some $\mu > 0$.*

(i) (Subcritical Case) For $0 < \alpha < \frac{4}{n}$, we have

$$-\infty < I_\mu < 0 \quad \text{for all } \mu > 0.$$

(ii) (Critical Case) For $\alpha = \frac{4}{n}$, there exists a $\mu_0 > 0$ such that $I_\mu = 0$ for $0 < \mu \le \mu_0$ and $I_\mu = -\infty$ for $\mu > \mu_0$.

(iii) (Supercritical Case) For $\frac{4}{n} < \alpha \le \frac{4}{n-2}$,

$$I_\mu = -\infty \quad \text{for all } \mu > 0.$$

Proof. This follows essentially by a scaling argument. Fix $\mu > 0$ and fix $u \in C_c^\infty(\mathbb{R}^n)$ with $\|u\|^2_{L^2(\mathbb{R}^n)} = \mu$. Consider the family of dilations

$$u_L(x) = L^{n/2} u(Lx), \quad L > 0$$

and note for all $L > 0$ we have $\|u_L\|_{L^2(\mathbb{R}^n)} = \|u\|_{L^2(\mathbb{R}^n)}$ and

$$\mathcal{H}(u_L) = \frac{L^2}{2} \int_{\mathbb{R}^n} |Du|^2 dx - \frac{L^{n\alpha/2}}{\alpha + 2} \int_{\mathbb{R}^n} |u|^{\alpha+2} dx.$$

[15] Observe $\mathcal{H}$ clearly is not bounded below on all of $H^1(\mathbb{R}^n)$. Indeed, simply notice that since $\alpha > 0$ we clearly have $\mathcal{H}(\gamma u) \to -\infty$ as $\gamma \to \infty$. Clearly though, this scaling does not preserve the constraint.

[16] With the appropriate understanding of this when $n = 1$, and the strict upper bound when $n = 2$. I'm just not going to keep saying it...

Note that if $\frac{n\alpha}{2} > 2$, i.e. $\alpha > \frac{4}{n}$, then

$$\mathcal{H}(u_L) \to -\infty \quad \text{as} \quad L \to \infty,$$

which establishes (iii).

Similarly, if $\frac{n\alpha}{2} < 2$, i.e. $\alpha < \frac{4}{n}$, then $\mathcal{H}(u_L) < 0$ for $L > 0$ sufficiently small. Thus, $I_\mu < 0$ for $\alpha < \frac{4}{n}$. To see that I_μ is finite for $\alpha < \frac{4}{n}$, we need the Sobolev embedding estimate[17]

$$\|u\|_{L^p(\mathbb{R}^n)} \leq C_{n,p}\|Du\|_{L^2(\mathbb{R}^n)}^{n(1/2-1/p)}\|u\|_{L^2(\mathbb{R}^n)}^{1-n(1/2-1/p)} \tag{15}$$

valid for all[18] $2 \leq p \leq \frac{2n}{n-2}$. Since $\alpha < \frac{4}{n-2}$ implies that $\alpha + 2 < \frac{2n}{n-2}$, taking $p = \alpha + 2$ in (15) gives

$$\mathcal{H}(u) \geq \frac{1}{2}\|Du\|_{L^2(\mathbb{R}^n)}^2 - \frac{C_{n,p}^{\alpha+2}}{\alpha+2}\|Du\|_{L^2(\mathbb{R}^n)}^{n\alpha/2}\mu^b,$$

where $b = \frac{1}{2}\left(1 - n\left(\frac{1}{2} - \frac{1}{\alpha+2}\right)\right)(\alpha+2)$. In particular, if we define $g : [0,\infty) \to \mathbb{R}$ by

$$g(R) = \frac{1}{2}R^2 - \frac{C_{n,p}^{\alpha+2}\mu^b}{\alpha+2}R^{n\alpha/2}, \tag{16}$$

then we have

$$\mathcal{H}(u) \geq g(\|Du\|_{L^2(\mathbb{R}^n)}).$$

From the graph of g, see Figure 4, it follows that for $\alpha < \frac{4}{n}$ we have

$$I_\mu \geq g_{\min}(\mu) > -\infty,$$

establishing (i).

Finally, to verify (ii) note that $\alpha = \frac{4}{n}$ implies that $\frac{n\alpha}{2} = 2$ so that the GNS inequality (15) implies

$$\mathcal{H}(u) \geq \frac{1}{2}\|Du\|_{L^2(\mathbb{R}^n)}^2\left(1 - \frac{2C_{n,p}^{\alpha+2}}{\alpha+2}\mu^{2/n}\right).$$

The idea is that dilations send the first factor on the right hand side to ∞ as $L \to \infty$, while the second factor on the right hand side is positive for $\mu > 0$ small and negative or $\mu > 0$ sufficiently large. For details, see the exercises. $\square$

Remark 5. *It is interesting to observe that in the "critical" case above, the critical mass μ_0 is related to the sharp constant $C_{n,p}$ in the GNS inequality* (15). *It turns out you can also show this constant is related to sufficient conditions for "global" existence of solutions for the critical NLS. For those interested, this would make for an interesting final project.*

With the above setup, we are now prepared to establish our main result.

[17] This follows from equation (**??**) in Section 8.1, there taking $p = 2$ and relabeling $q \mapsto p$.

[18] Again, if $n = 2$ the upper bound should be strict.

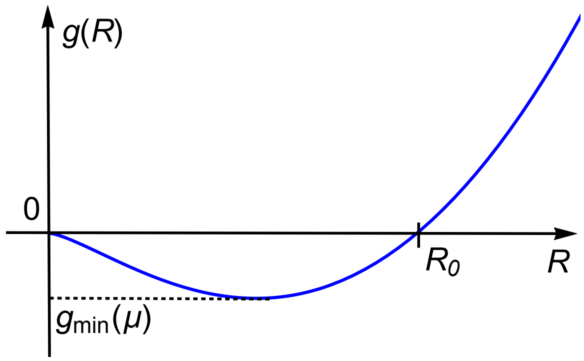

Figure 4: A representative graph of the function $g(R)$ defined in (16) in the case when $\frac{n\alpha}{2} < 2$, i.e. when $\alpha < \frac{4}{n}$.

Theorem 12 (Existence of Minimizers in Subcritical Case)**.** *For each $\mu > 0$ define the admissible set*

$$\mathcal{A}_\mu := \left\{ u \in H^1(\mathbb{R}^n) : \|u\|_{L^2(\mathbb{R}^n)}^2 = \mu \right\}.$$

If $0 < \alpha < \frac{4}{n}$ and $\mu > 0$, then the problem

$$I_\mu = \inf_{u \in \mathcal{A}_\mu} \mathcal{H}(u) \tag{17}$$

has a minimizer.

Before we begin the proof, I want to note that with more work one can prove minimizers of (17) are unique up to spatial translations and complex-rotations (i.e. multiplication by $e^{i\theta}$). You can show that if u_0 is a minimizer, then there exists a $\theta \in [0, 2\pi)$ such that $e^{i\theta} u_0$ is strictly positive on all of $\mathbb{R}^n$. For the purposes of these notes, however, we content ourselves with verifying the existence of minimizers only. If students are interested in these more subtle points, this could serve as an interesting final project in the course.

Proof. (Of Theorem 12) Since $0 < \alpha < \frac{4}{n}$, Proposition 1 implies that $I_\mu > -\infty$ and hence we can find a minimizing sequence $\{u_k\}$ in $H^1(\mathbb{R}^n)$ such that $\|u_k\|_{L^2(\mathbb{R}^n)}^2 = \mu$ for all k and

$$\mathcal{H}(u_k) \to I_\mu \quad \text{as } k \to \infty.$$

Since $I_\mu < 0$ it follows that $\mathcal{H}(u_k) < 0$ for k sufficiently large, and hence, by the GNS inequality, we have

$$g\left(\|Du_k\|_{L^2(\mathbb{R}^n)}\right) \le \mathcal{H}(u_k) < 0$$

for all k sufficiently large, where here g is defined in (16). Using the notation in Figure 4, it follows that

$$\|Du_k\|_{L^2(\mathbb{R}^n)} \le R_0$$

for all k sufficiently large. It follows that $\{u_k\}$ is a bounded sequence in $H^1(\mathbb{R}^n)$ and hence, by Concentration Compactness, we are either in the "vanishing", "splitting", or "convergence of translates" situation. Our goal below is to rule out the vanishing and splitting cases.

First, we rule out the "vanishing case". By definition, we know for all k sufficiently large we will have

$$\mathcal{H}(u_k) = \frac{1}{2}\|Du_k\|^2_{L^2(\mathbb{R}^n)} - \frac{1}{\alpha+2}\|u_k\|^{\alpha+2}_{L^{\alpha+2}(\mathbb{R}^n)} \leq \frac{I_\mu}{2} < 0,$$

and hence we have the uniform (in k) bound

$$\frac{1}{\alpha+2}\|u_k\|^{\alpha+2}_{L^{\alpha+2}(\mathbb{R}^n)} \geq -\frac{I_\mu}{2} > 0 \tag{18}$$

valid for all k sufficiently large. Note that if vanishing occurred, there would exist a subsequence $\{u_{k_j}\}$ such that $u_{k_j} \to 0$ in $L^p(\mathbb{R}^n)$ for all $2 < p < \frac{2n}{n-2}$. In particular, since $0 < \alpha < \frac{4}{n}$ implies $2 < \alpha+2 < \frac{2n}{n-2}$ we would necessarily have $\|u_{k_j}\}_{L^{\alpha+2}(\mathbb{R}^n)} \to 0$ as $j \to \infty$, which owing to the uniform bound in (18), contradicts that $I_\mu < 0$ for such μ. It follows that "vanishing" can not occur.

Next, we rule out "splitting". Note that IF splitting occured, then there would exist a subsequence $\{u_{k_j}\}$ and bounded sequences $\{v_j\}$ and $\{w_j\}$ in $H^1(\mathbb{R}^n)$ such that

$$\|v_j\|^2_{L^2(\mathbb{R}^n)} \to \nu, \quad \|w_j\|^2_{L^2(\mathbb{R}^n)} \to \mu - \nu$$

for some $\nu \in (0, \mu)$ and, furthermore, we would have

$$\int_{\mathbb{R}^n} \left|(u_{k_j}|^p - |v_j|^p - |w_j|^p\right) dx \to 0 \quad \text{as } j \to \infty$$

for all $2 \leq p < \frac{2n}{n-2}$ (so, in particular, the choice $p = \alpha + 2$ is allowed by our previous considerations) and

$$\liminf_{j\to\infty} \int_{\mathbb{R}^n} \left(|Du_{k_j}|^2 - |Dv_j|^2 - |Dw_j|^2\right) dx \geq 0.$$

Thus, for all $\epsilon > 0$ and j sufficiently large we would have

$$\begin{aligned}
I_\mu + 3\epsilon \geq \mathcal{H}(u_{k_j}) + 2\epsilon &= \frac{1}{2}\int_{\mathbb{R}^n} |Du_{k_j}|^2 dx - \frac{1}{\alpha+2}\int_{\mathbb{R}^n} |u_{k_j}|^{\alpha+2} dx + 2\epsilon \\
&\geq \frac{1}{2}\int_{\mathbb{R}^n} \left(|Dv_j|^2 + |Dw_j|^2\right) dx - \frac{1}{\alpha+2}\left(\int_{\mathbb{R}^n} \left(|v_j|^{\alpha+2} - |w_j|^{\alpha+2}\right) dx + (\alpha+2)\epsilon\right) + 2\epsilon \\
&= \mathcal{H}(v_j) + \mathcal{H}(w_j) + \epsilon.
\end{aligned}$$

Now, observe that we can the sequences $\{v_j\}$ and $\{w_j\}$ to have constant norm in $L^2(\mathbb{R}^n)$. Indeed, there exist sequences $\{a_j\}, \{b_j\} \subset \mathbb{R}^+$ such that

$$\|a_j v_j\|^2_{L^2(\mathbb{R}^n)} = \nu \quad \text{and} \quad \|b_j w_j\|^2_{L^2(\mathbb{R}^n)} = \mu - \nu$$

for all $j \in \mathbb{N}$ and, further, we clearly have $a_j, b_j \to 1$ as $j \to \infty$ by construction. For j sufficiently large then we have

$$\mathcal{H}(v_j) \geq \mathcal{H}(a_j v_j) - \frac{\epsilon}{2}, \quad \text{and} \quad \mathcal{H}(w_j) \geq \mathcal{H}(b_j w_j) - \frac{\epsilon}{2}.$$

It follows that for all j sufficiently large that

$$I_\mu + 3\epsilon \geq \mathcal{H}(a_j v_j) + \mathcal{H}(b_j w_j) \geq I_\nu + I_{\mu-\nu}.$$

Since $\epsilon > 0$ was arbitrary, it follows that IF splitting occurs, then we would have to have

$$I_\mu \geq I_\nu + I_{\mu-\nu}$$

Turns out this is impossible by the following, which rules out the possibility of "splitting".

Lemma 5. *The map $\mu \to I_\mu$ is strictly subadditive, i.e. for all $\mu > 0$ we have*

$$I_\mu < I_\nu + I_{\mu-\nu}$$

for all $\nu \in (0, \mu)$.

Proof. This follows by a scaling argument. Indeed, notice for all $\theta > 1$ and $u \in H^1(\mathbb{R}^n)$ that

$$\mathcal{H}(\theta u) = \frac{\theta^2}{2}\|Du\|^2_{L^2(\mathbb{R}^n)} - \frac{\theta^{\alpha+2}}{\alpha+2}\|u\|^{\alpha+2}_{L^{\alpha+2}(\mathbb{R}^n)} < \theta^2 \mathcal{H}(u)$$

and hence

$$I_{\theta^2\mu} < \theta^2 I_\mu \quad \text{for all } \theta > 1, \ \mu > 0.$$

If $\nu \in [\mu/2, \mu)$, i.e. $\mu - \nu \leq \nu < \mu$, it follows that

$$\begin{aligned} I_\mu &= I_{\left(\frac{\mu}{\nu}\right)\nu} < \frac{\mu}{\nu} I_\nu = I_\nu + \frac{\mu-\nu}{\nu} I_\nu \\ &= I_\nu + \frac{\mu-\nu}{\nu} I_{\frac{\nu}{\mu-\nu}(\mu-\nu)} \\ &\leq I_\nu + I_{\mu-\nu}, \end{aligned}$$

where notice the lack of strict inequality at the end occurs precisely at the end point $\nu = \frac{\mu}{2}$. Repeating the above argument with $\nu \mapsto \mu - \nu$ establishes the same estimate for $\nu \in (0, \mu/2]$, yielding the result. $\square$

By Concentration Compactness, it follows there exists a sequence $\{y_j\} \subset \mathbb{R}^n$ and a subsequence $\{u_{k_j}\}$ such that

$$u_{k_j}(\cdot - y_j) \to \phi \ \text{ in } L^p(\mathbb{R}^n)$$

for all $2 \leq p < \frac{2n}{n-2}$ for some function $\phi \in H^1(\mathbb{R}^n)$. I claim that ϕ is a minimizer of (17). To see this, note that taking $p = 2$ (and recalling translations leave the L^2-norm invariant) we see that $\|\phi\|^2_{L^2(\mathbb{R}^n)} = \mu$ so that, by definition,

$$I_\mu \leq \mathcal{H}(\phi).$$

Similarly, taking $p = \alpha + 2$ gives

$$\lim_{j\to\infty} \|u_{k_j}(\cdot - y_j)\|_{L^{\alpha+2}(\mathbb{R}^n)} = \|\phi\|_{L^{\alpha+2}(\mathbb{R}^n)}.$$

Since Banach-Alaoglu and the uniqueness of weak limits implies

$$u_{k_j}(\cdot - y_j) \rightharpoonup \phi \text{ in } H^1(\mathbb{R}^n),$$

the lower weak semicontinuity of the map

$$H^1(\mathbb{R}^n) \ni v \mapsto \int_{\mathbb{R}^n} |Dv|^2 dx$$

implies that $\int_{\mathbb{R}^n} |D\phi|^2 dx \leq \liminf_{j\to\infty} \int_{\mathbb{R}^n} |Du_{k_j}(x - y_j)|^2 dx$ and hence we have

$$\mathcal{H}(\phi) \leq \liminf_{j\to\infty} \mathcal{H}(u_{k_j}(\cdot - y_j))$$

Since $\{u_{k_j}(\cdot - y_j)\}$ is a minimizing sequence for I_μ, it follows that

$$I_\mu \leq \mathcal{H}(\phi) \leq \liminf_{j\to\infty} \mathcal{H}(u_{k_j}(\cdot - y_j)) = I_\mu$$

so that ϕ is a minimizer of (17), as desired. □

In summary, it follows that in the subcritical case $0 < \alpha < \frac{4}{n}$, given any $\mu > 0$ there exists a function $\phi_\mu \in H^1(\mathbb{R}^n)$ that (weakly) satisfies

$$\begin{cases} -\Delta u - |u|^\alpha u = \lambda u \\ \qquad \|u\|^2_{L^2(\mathbb{R}^n)} = \mu \end{cases}$$

for some constant $\lambda < 0$. Hence, the function

$$u_\mu(x,t) = e^{-i\lambda t}\phi_\mu(x)$$

is a weak solution of the subcritical NLS

$$iu_t = -\Delta u - |u|^\alpha u \text{ with } \|u_\mu\|^2_{L^2(\mathbb{R}^n)} = \mu.$$

Such solutions are known as "ground state solutions" of the NLS equation (11). Our next result uses concentration compactness to verify the "stability" of ground state solutions of the NLS.

Theorem 13. *In the subcritical case, the ground state u_μ is stable solution of* (11) *in the following sense: for all $t_0 \in \mathbb{R}$ and $\epsilon > 0$, there exists a $\delta = \delta(t_0, \epsilon) > 0$ such that all solutions $u(t)$ of* (11) *satisfying*

$$\|u(t_0) - u_\mu(t_0)\|_{H^1(\mathbb{R}^n)} < \delta$$

we have

$$\sup_{t\in\mathbb{R}} \inf_{\theta\in\mathbb{R}} \inf_{y\in\mathbb{R}^n} \left\| u(t) - e^{i\theta} u_\mu(\cdot - y, t) \right\|_{H^1(\mathbb{R}^n)} < \epsilon.$$

A consequence of Theorem 13 is that there exists real-valued "modulation functions" $\theta(t), y(t)$ such that

$$\sup_{t\in\mathbb{R}} \left\| u(t) - e^{i\theta(t)} u_\mu(\cdot - y(t), t) \right\|_{H^1(\mathbb{R}^n)} < \epsilon$$

Thus, in the subcritical case, solutions of (11) which start close to a ground state remain close to the orbit of the ground state modulo the group actions of spatial translations and complex-rotations. Such a stability is called "orbital stability", and is a common feature in systems with continuous symmetries.

To motivate the need for the modulation functions, observe that solutions of (11) are invariant under spatial translations and complex rotations, i.e. if $u(x, t)$ solves (11) then so does $e^{i\theta} u(x - y, t)$ for any constants $\theta, y \in \mathbb{R}$. In particular, it follows that ground states are never "isolated" from each other, but rather always arise in two-paramter families, and hence Theorem 13 is a statement regarding the stability of the two-parameter family rather than the stability of a particular member of that family.

There are two main ways (that I know of) to approach the stability result in Theorem 13. The main way I would approach it is to use a Lyapunov functional. However, this is slightly outside the scope and methods of this class[19]. Here, I present a proof based on Concentration Compactness. As a first step, we establish the following.

Lemma 6 (Conservation of Energy and Mass)**.** *Given any solution $u \in C(\mathbb{R}, H^1(\mathbb{R}; \mathbb{C})) \cap C^1(\mathbb{R}; H^{-1}(\mathbb{R}^n; \mathbb{C}))$ of* (11)*, we have*

$$\frac{d}{dt}\mathcal{H}(u(t)) = \frac{d}{dt} W(u(t)) = 0.$$

That is, the energy $\mathcal{H}$ and the L^2 norm (i.e. mass) of solutions are conserved through time.

"Proof". While doing this rigorously is a pain, this result can be seen quite easily by working formally. First, note that multiplying (11) by $\bar{u}$ and integrating gives

$$i \int_{\mathbb{R}^n} u_t \bar{u} \, dx = - \int_{\mathbb{R}^n} (\Delta u) \bar{u} \, dx - \int_{\mathbb{R}^n} |u|^{\alpha+2} dx = \int_{\mathbb{R}^n} |Du|^2 dx - \int_{\mathbb{R}^n} |u|^{\alpha+2} dx,$$

where the last equality follows from integration by parts. Taking the complex conjugate of the above gives

$$-i \int_{\mathbb{R}^n} \bar{u}_t u \, dx = \int_{\mathbb{R}^n} |Du|^2 dx - \int_{\mathbb{R}^n} |u|^{\alpha+2} dx$$

[19] If you are interested in this, take my Math 851 class sometime :-)

and hence, formally,

$$\frac{d}{dt}\mathcal{W}(u(t)) = \frac{d}{dt}\int_{\mathbb{R}^n} |u|^2 \; dx = \int_{\mathbb{R}^n} \bar{u}_t u \; dx + \int_{\mathbb{R}^n} u_t \bar{u} \; dx = 0,$$

as desired.

Similarly, we have (again, completely formally),

$$\begin{aligned} \frac{d}{dt}\int_{\mathbb{R}^n} |Du|^2 dx &= \int_{\mathbb{R}^n} \overline{Du_t} \cdot Du \; dx + \int_{\mathbb{R}^n} \overline{Du} \cdot Du_t \; dx \\ &= -\int_{\mathbb{R}^n} \overline{u_t} \; \Delta u \; dx - \int_{\mathbb{R}^n} \overline{\Delta u} \; u_t \; dx \end{aligned} \tag{19}$$

and

$$\begin{aligned} \frac{d}{dt}\int_{\mathbb{R}^n} |u|^{\alpha+2} dx &= \frac{d}{dt}\int_{\mathbb{R}^n} \left(|u|^2\right)^{(\alpha+2)/2} dx = \frac{\alpha+2}{2}\int_{\mathbb{R}^n} |u|^\alpha \frac{d}{dt}|u|^2 \; dx \\ &= \frac{\alpha+2}{2}\left(\int_{\mathbb{R}^n} |u|^\alpha \overline{u_t} u \; dx + \int_{\mathbb{R}^n} |u|^\alpha \overline{u} \; u_t \; dx\right). \end{aligned} \tag{20}$$

It follows that, formally,

$$\begin{aligned} \frac{d}{dt}\mathcal{H}(u(t)) &= \frac{1}{2}\frac{d}{dt}\|Du\|^2_{L^2(\mathbb{R}^n)} - \frac{1}{\alpha+2}\frac{d}{dt}\|u\|^{\alpha+2}_{L^{\alpha+2}(\mathbb{R}^n)} \\ &= -\frac{1}{2}\int_{\mathbb{R}^n} \overline{u_t}\,(\Delta u + |u|^\alpha u)\, dx - \frac{1}{2}\int_{\mathbb{R}^n} u_t \overline{(\Delta u + |u|^\alpha u)} dx, \end{aligned}$$

which, using (11), gives

$$\frac{d}{dt}\mathcal{H}(u(t)) = \frac{1}{2}\int_{\mathbb{R}^n} \overline{u_t}\,(iu_t)\, dx + \frac{1}{2}\int_{\mathbb{R}^n} u_t \overline{(iu_t)} dx = 0,$$

as desired. $\square$

Remark 6. *Of course, the above proof leaves much to be desired. Specifically, notice that since we assumed $u_t \in H^1{-}1(\mathbb{R}^n)$ and $u \in H^1(\mathbb{R}^n)$ that the pairings in* (19) *might not make sense. Similarly, we have $|u|^\alpha u \notin H^1$, unless we are in dimension $n = 1$, and hence the pairings in* (20) *similarly might not make sense. The rigorous proof, of course, deals with these issues.*

With the above conservation laws established, we now provide a proof of Theorem 13.

Proof of Theorem 13. Suppose, by way of contradiction, that the ground state u_μ is not stable. Then there exists a $t_0 \in \mathbb{R}$ and an $\epsilon > 0$ such that for every $n \in \mathbb{N}$ there exists a solution ψ_n of (11) such that

$$\|\psi_n(t_0) - u_\mu(t_0)\|_{H^1(\mathbb{R}^n)} < \frac{1}{n} \tag{21}$$

and

$$\inf_{\theta\in\mathbb{R}}\inf_{y\in\mathbb{R}^n}\left\|\psi_n(t_n)-e^{i\theta}u_\mu(\cdot-y,t_n)\right\|_{H^1(\mathbb{R}^n)}\geq\epsilon \tag{22}$$

for some sequence of times $t_n\geq t_0$. Note if we decompose the solutions ψ_n as

$$\psi_n(t_n)=e^{i\lambda t_n}g_n(t_n)$$

then by our discussion directly above Theorem 13 the conditions (21)-(22) become

$$\|g_n(t_0)-\phi_\mu\|_{H^1(\mathbb{R}^n)}<\frac{1}{n} \tag{23}$$

and

$$\inf_{\theta\in\mathbb{R}}\inf_{y\in\mathbb{R}^n}\left\|g_n(t_n)-e^{i\theta}\phi_\mu(\cdot-y)\right\|_{H^1(\mathbb{R}^n)}\geq\epsilon \tag{24}$$

Our goal is to prove that (23) and (24) together imply a contradiction by using the method of concentration compactness.

To this end, observe that by the conservation of $\mathcal{H}$ we have

$$\mathcal{H}(g_n(t_n))=\mathcal{H}(g_n(t_0))\to\mathcal{H}(\phi_\mu)=I_\mu,$$

where the above limit holds since, in the subcritical case, the GNS inequality (15) implies continuity of $\mathcal{H}$ with respect to convergence in $H^1(\mathbb{R}^n)$. Similarly, by conservation of mass we have

$$\|g_n(t_n)\|^2_{L^2(\mathbb{R}^n)}=\|g_n(t_0)\|^2_{L^2(\mathbb{R}^n)}\to\|\phi_\mu\|^2_{L^2(\mathbb{R}^n)}=\mu.$$

In particluar, it follows that the sequence

$$v_n=b_ng_n(t_n),\quad b_n=\frac{\sqrt{\mu}}{\|g_n(t_n)\|_{L^2(\mathbb{R}^n)}}$$

is a minimizing sequence in $H^1(\mathbb{R}^n)$ for $\mathcal{H}$, subject to the constraint $\|u\|^2_{L^2(\mathbb{R}^n)}=\mu$. Using the same argument as before, Concentration Compactness implies there exists a subsequence v_{n_j} and a sequence $\{y_j\}$ such that

$$v_{n_j}(\cdot-y_j)\rightharpoonup\phi\ \text{ in } H^1(\mathbb{R}^n) \tag{25}$$

for some minimizer ϕ of $\mathcal{H}$ subject to $\|u\|^2_{L^2(\mathbb{R}^n)}=\mu$ and, similarly,

$$v_{n_j}\to\phi\ \text{ in } L^p(\mathbb{R}^n),\quad 2\leq p<\frac{2n}{n-2}.$$

In particular, since $\mathcal{H}(v_{n_j}(\cdot-y_j))\to I_\mu$ and $\|v_{n_j}(\cdot-y_j)\|_{L^{\alpha+2}(\mathbb{R}^n)}\to\|\phi\|_{L^{\alpha+2}(\mathbb{R}^n)}$, the definition of $\mathcal{H}$ implies that

$$\|Dv_{n_j}(\cdot-y_j)\|_{L^2(\mathbb{R}^n)}\to\|D\phi\|_{L^2(\mathbb{R}^n)}.$$

Since convergence of norms plus weak convergence implies norm convergence[20], it follows that the weak convergence in (25) can be upgraded to strong convergence, i.e. we have

$$v_{n_j}(\cdot - y_j) \to \phi \ \text{ in } H^1(\mathbb{R}^n).$$

To show this implies a contradiction, recall from the discussion below Theorem 12 that L^2-constrained minimizers of $\mathcal{H}$ are unique up to spatial translations and complex-rotations. In particular, we have

$$\phi = e^{i\theta_0}\phi_\mu(\cdot - y_0)$$

for some $\theta_0 \in \mathbb{R}$ and $y_0 \in \mathbb{R}^n$. It now follows that

$$\begin{aligned}
\inf_{\theta\in\mathbb{R}} \inf_{y\in\mathbb{R}^n} \|g_{n_j}(t_{n_j}) - e^{i\theta}\phi_\mu(\cdot - y)\|_{H^1(\mathbb{R}^n)} &\leq \|g_{n_j}(t_{n_j}) - e^{i\theta_0}\phi_\mu(\cdot - y_0 + y_j)\|_{H^1(\mathbb{R}^n)} \\
&= \|\frac{1}{b_{n_j}} v_{n_j}(\cdot - y_j) - e^{i\theta_0}\phi_\mu(\cdot - y_0)\|_{H^1(\mathbb{R}^n)} \\
&\leq \left(1 - \frac{1}{b_{n_j}}\right)\|v_{n_j}\|_{H^1(\mathbb{R}^n)} + \|v_{n_j}(\cdot - y_j) - e^{i\theta_0}\phi_\mu(\cdot - y_0)\|_{H^1(\mathbb{R}^n)} \\
&\to 0 \ \text{ as } \ j \to \infty,
\end{aligned}$$

which contradicts (24). □

One of the keys in the above proof was the fact that constrained L^2-minimizers of $\mathcal{H}$ are unique up to spatial translations and complex rotations. In many cases, one is not able to obtain such a precise description of all minimizers, and in that case the above proof establishes the stability of the set of minimizers. That is, if you start near a minimizer then you stay near a minimizer. The ability to classify all minimizers is intimately related to the so-called "non-degeneracy" of the linearization of the problem. In the above case, this corresponds to being able to show that the kernel of the linearization of the NLS about a ground state u_μ is generated exactly by the translational and phase invariances of the governing PDE. For more on this topic, talk to me or Math 851 next spring :-)

7 Exercises

Complete the following Exercises.

1. (Based on #2, Section 13.3 from McOwen) Let $U \subset \mathbb{R}^n$ be open and bounded with smooth domain, and consider the following nonlinear Dirichlet problem

$$\begin{cases} \Delta u + f(u) = 0, & \text{in } U \\ u = 0 & \text{on } \partial U, \end{cases} \tag{26}$$

[20]Indeed, suppose $\{x_j\}$ is a sequence in a Hilbert space H such that $x_j \rightharpoonup x$ in H and $\|x_j\| \to \|x\|$. Expanding

$$\|x_j - x\|^2 = \langle x_j - x, x_j - x\rangle = \|x_j\|^2 - 2\langle x_j, x\rangle + \|x\|^2$$

it then follows that $x_j \to x$ in H, as claimed.

where $f : \mathbb{R} \to \mathbb{R}$ is a continuous function.

(a) If $F(u) = \int_0^u f(t)dt$, use integration by parts to show that

$$n \int_U F(u)dx + \int_U f(u) \sum_{i=1}^n x_i \frac{\partial u}{\partial x_i}\, dx = 0$$

for any $u \in C^1(\bar{U})$ with $u = 0$ on ∂U.

(b) If $u \in C^2(U) \cap C(\bar{U})$ satisfies (26), then prove *Pohozaev's identity*

$$\frac{n-2}{2} \int_U |Du|^2 dx - n \int_U F(u)dx + \frac{1}{2} \int_{\partial U} \left(\frac{\partial u}{\partial \nu}\right)^2 (x \cdot \nu) dS = 0,$$

where ν is the exterior unit normal vector to ∂U.

(c) When $U = B(0, r)$ is a ball in $\mathbb{R}^n$, show that (26) does not admit a non-trivial classical solution $u \in C^2(U) \cap C(\bar{U})$ when $f(u) = |u|^{p-1}u$ with $p > \frac{n+2}{n-2}$.

2. Let $U \subset \mathbb{R}^n$ be open and define the functional $F : H_0^1(U) \to \mathbb{R}$ by $F(u) = \|u\|^2_{L^2(U)}$.

(a) Prove that F is Gâteaux differentiable on $H_0^1(U)$, and identify its Gâteaux derivative at an arbitrary $u \in H_0^1(U)$ in an arbitrary direction $v \in H_0^1(U)$. Further, show that, for each $v \in H_0^1(U)$, the map $dF[u] : H_0^1(U) \to \mathbb{R}$ is a bounded liner map.

(b) Prove that F is Fréchet differentiable and that the map $DG : H_0^1(U) \to B(H_0^1(U), \mathbb{R})$ is continuous. Here, $B(X, \mathbb{R})$ denotes the set of all bounded linear maps from a Banach space X into $\mathbb{R}$.

(c) (Suggested) Discuss the differentiability properties of L^p norms on on $H_0^1(U)$.

3. For $n \geq 3$ and $0 < \alpha < \frac{4}{n-2}$, define the functional $\mathcal{H} : H^1(\mathbb{R}^n) \to \mathbb{R}$ by

$$\mathcal{H}(u) = \frac{1}{2} \int_{\mathbb{R}^n} |Du|^2 - \frac{1}{\alpha + 2} \int_{\mathbb{R}^n} |u|^{\alpha+2} dx$$

and note that, as seen in class, $\mathcal{H}$ is well-defined by Sobolev embedding. For each $\mu > 0$, define

$$\mathcal{A}_\mu := \left\{ u \in H^1(\mathbb{R}^n) : \int_{\mathbb{R}^n} u^2 dx = \mu \right\}$$

and set

$$I_\mu := \inf_{u \in \mathcal{A}_\mu} \mathcal{H}(u).$$

Show that if $\alpha = \frac{4}{n}$, then there exists a $\mu_0 > 0$ such that $I_\mu = 0$ for all $\mu \in (0, \mu_0]$, while $I_\mu = -\infty$ for all $\mu > \mu_0$.

Hint: Recall the following interpolation inequality[21]: for $n \geq 3$, given any $u \in H^1(\mathbb{R}^n)$ there exists a constant $C_{n,p}$ such that

$$\|u\|_{L^p(\mathbb{R}^n)} \leq C_{p,n} \|Du\|_{L^2(\mathbb{R}^n)}^{n(1/2-1/p)} \|u\|_{L^2(\mathbb{R}^n)}^{1-n(1/2-1/p)},$$

for all $2 \leq p \leq \frac{2n}{n-2}$. First, use this inequality to prove there exists a $\mu_0 > 0$ such that $I_\mu > -\infty$ for all $\mu \in (0, \mu_0]$ and $I_\mu = -\infty$ if $\mu > \mu_0$. Next, to show $I_\mu = 0$ for all $\mu \in (0, \mu_0]$, fix $u \in H^1(\mathbb{R}^n)$ with $\|u\|_{L^2(\mathbb{R}^n)}^2 = \mu$ and consider the dialiations $u_L(x) = L^{n/2}u(Lx)$, defined for all $L > 0$. In particular, computing $\mathcal{H}(u_L)$, conclude that $I_\mu = 0$ for all $\mu \in (0, \mu_0]$.

Extended Hint: For the case $\mu > \mu_0$, it is important to prove the existence of a $u \in \mathcal{A}$ such that $\mathcal{H}(u) < 0$. **I will allow you to assume the existence of such a function $u \in \mathcal{A}$ without proof!!** However, I will award extra credit to students who are able to use the following result to prove the existence of such a function $u \in \mathcal{A}$: The problem

$$C_n^{-1} = \inf \left\{ \frac{\|Du\|_{L^2(\mathbb{R}^n)}^{n/(n+2)} \|u\|_{L^2(\mathbb{R}^n)}^{2/(n+2)}}{\|u\|_{L^{4/n+2}(\mathbb{R}^n)}} : u \in H^1(\mathbb{R}^n) \setminus \{0\} \right\}$$

has a strictly positive minimum. In particular, observe that C_n is the *best* constant in the GNS inequality described above in the critical case $\alpha = \frac{4}{n}$, in the sense that if you choose any constant $\tilde{C}$ with $\tilde{C}^{-1} > C_n^{-1}$ then there exists a function $\tilde{u}$ such that

$$\|\tilde{u}\|_{L^{4/n+2}(\mathbb{R}^n)} > \tilde{C} \|D\tilde{u}\|_{L^2(\mathbb{R}^n)}^{n/(n+2)} \|\tilde{u}\|_{L^2(\mathbb{R}^n)}^{2/(n+2)}.$$

4. Let $V : \mathbb{R}^n \to \mathbb{R}$ be such that $V \in L^\infty(\mathbb{R}^n)$ and $\lim_{|x|\to\infty} V(x) = 0$. The goal of this problem is to analyze the "Schrödinger" eigenvalue problem

$$-\Delta u + V(x)u = \lambda u, \quad x \in \mathbb{R}^n, \ \lambda \in \mathbb{R}, \tag{27}$$

posed on $H^1(\mathbb{R}^n)$. For this purpose, define the admissible set

$$\mathcal{A} := \left\{ u \in H^1(\mathbb{R}^n) : \int_{\mathbb{R}^n} |u|^2 dx = 1 \right\}$$

and consider the minimization problem

$$\mu = \inf_{u \in \mathcal{A}} \int_{\mathbb{R}^n} \left(|Du|^2 + V(x)|u|^2 \right) dx \tag{28}$$

[21] This follows from (**??**), which was a crucial component in our proof of the Gagliardo-Nirenberg-Sobolev inequality.

(a) Show that if $u \in \mathcal{A}$ is a minimizer for (28), then u is a weak solution of (27) for $\lambda = \mu$.

(b) Suppose $\{u_k\}_{k=1}^{\infty}$ is a sequence in $H^1(\mathbb{R}^n)$ such that u_k converges weakly to a function $u \in H^1(\mathbb{R}^n)$ weakly in $H^1(\mathbb{R}^n)$. Establish the following "compactness" result: under the above hypothesis on V, there exists a subsequence[22] $\{u_{k_j}\}_{j=1}^{\infty}$ such that

$$\lim_{j\to\infty} \int_{\mathbb{R}^n} V(x)|u_{k_j}(x)|^2 dx \to \int_{\mathbb{R}^n} V(x)|u(x)|^2 dx.$$

(*Hint: A fundamental results of functional analysis, known as the Banach-Steinhaus theorem, or principle of uniform boundedness, implies that all weakly convergent sequences in a Banach space are bounded*).

(c) Let V be as above and suppose there exists a function $w \in \mathcal{A}$ such that

$$\int_{\mathbb{R}^n} \left(|Dw|^2 + V(x)|w|^2\right) dx < 0.$$

Show that the minimization problem (28) has a minimizer $u \in \mathcal{A}$.

Hints:

(a) Feel free to use the result from Problem #2 above.

(b) Note that for any $k \in \mathbb{N}$ and $R > 0$, we have

$$\left|\int_{\mathbb{R}^n} V(x)\left(u_k^2 - u^2\right) dx\right| \leq \int_{B(0,R)} |V(x)| \cdot \left|u_k^2 - u^2\right| dx + \int_{\mathbb{R}^n \setminus B(0,R)} |V(x)| \cdot \left|u_k^2 - u^2\right| dx.$$

First, show the assumption on V imlpies the above integral over $\mathbb{R}^n \setminus B(0,R)$ can be made as small as desired by choosing R sufficiently large. Finally, you must control the remaining integral over the bounded domain $B(0,R)$.

(c) Follow the general strategy from class. A key part of the problem is to prove that weak limit u_0 of the minimizing sequence you get actually belongs to the constraint set $\mathcal{A}$. Note this is not immediate since the functional $G(u) = \|u\|^2_{L^2(\mathbb{R}^n)} - 1$ is NOT compact on $H^1(\mathbb{R}^n)$. However, you can establish $u_0 \in \mathcal{A}$ by contradiction argument. In particular, show that if $u_0 \notin \mathcal{A}$, then $0 < \|u_0\|^2_{L^2(\mathbb{R}^n)} < 1$ and use this observation to construct a function $v \in \mathcal{A}$ such that $F(v) < \mu$. To this end, it will be helpful to verify the following facts:

  (i) There exists a $w \in \mathcal{A}$ such that $F(w) < 0$.

[22] Actually, one can verify this result *with out* passing to a subsequence. While this is not necessary for the problem, I suggest trying to understand why this is so.

(ii) Given any $\beta \in \mathbb{R}$ and $u \in H^1(\mathbb{R}^n)$, we have $F(\beta u) = \beta^2 F(u)$, i.e. the functional F is homogeneous on $H^1(\mathbb{R}^n)$.

5. Continuing the above problem, set $F(u) := \int_{\mathbb{R}^n} \left(|Du|^2 + V(x)|u|^2\right) dx$ and let $\mathcal{A}$ be defined as above. A major question in mathematical quantum mechanics is to identify a class of potentials V for which it is true that $F(w) < 0$ for some $w \in \mathcal{A}$. Clearly, for this to be true the potential must be "negative enough": this intentionally vague notion is typically quantified by requiring that $\int_{\mathbb{R}^n} V(x)dx < 0$, in which case we say V is an "attractive" potential. However, in high dimensions, it is not the case that $F(w) < 0$ for some $w \in H^1(\mathbb{R}^n)$ for every attractive potential V. This exercise explores this idea.

(a) Show that if $n = 1, 2$ and the potential $V \in L^\infty(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$ is attractive, then there exists a $u \in \mathcal{A}$ such that $F(u) < 0$.

(b) Prove Hardy's inequality: if $n \geq 3$, then

$$\int_{\mathbb{R}^n} \frac{u^2}{|x|^2} dx \leq \frac{4}{(n-2)^2} \int_{\mathbb{R}^n} |Du|^2 dx$$

for all $u \in H^1(\mathbb{R}^n)$.

(c) Let $n \geq 3$ and define the potential $V(x) := -\frac{(n-2)^2}{4|x|^2}$ on $\mathbb{R}^n$. Conclude that even though V is an attractive potential, we have $F(u) \geq 0$ for every $u \in H^1(\mathbb{R}^n)$. Note: This part explicitly shows that for the time-dependent Schrödinger equation

$$iu_t = -\Delta u + V(x)u, \quad x \in \mathbb{R}^n, \quad t \in \mathbb{R},$$

attractive potentials in $\mathbb{R}^n$ for $n \geq 3$ may not support "bound states", i.e. solutions of the form $u(x,t) = e^{-i\lambda t}\phi(x)$ for some $\lambda \in \mathbb{R}$ and non-trivial $\phi \in H^1(\mathbb{R})$.

Hints:

(a) Fix $u \in C_c^\infty(\mathbb{R}^n) \cap \mathcal{A}$ with $u(0) \neq 0$ and define for each $L > 0$ the dilation $u_L(x) := L^{-n/2}u(x/L)$. Show then that, if $n = 1, 2$, we have $F(u_L) < 0$ for L sufficiently large. For this, it will be crucial to prove that

$$\lim_{L\to\infty} \int_{\mathbb{R}^n} V(x)|u(x/L)|^2 dx = |u(0)|^2 \int_{\mathbb{R}^n} V(x)dx.$$

(b) Notice for all $\lambda \in \mathbb{R}$ we have $|Du + \lambda \frac{x}{|x|^2} u|^2 \geq 0$. It follows then that, for each $u \in H^1(\mathbb{R}^n)$,

$$0 \leq \int_{\mathbb{R}^n} \left(|Du|^2 + 2\lambda u Du \cdot \frac{x}{|x|^2} + \lambda^2 \frac{u^2}{|x|^2} \right) dx.$$

Integrate the second term by parts and minimize the resulting inequality over $\lambda \in \mathbb{R}$ to verify the result.

6. Let $u_0 \in H^1(\mathbb{R}^n)$ be a non-trivial function such that $|x|u_0 \in L^2(\mathbb{R}^n)$, and let $u(x,t)$ be an $H^1(\mathbb{R}^n)$ solution of

$$iu_t = -\Delta u + |u|^\alpha u, \quad x \in \mathbb{R}^n, \quad t \in \mathbb{R}, \quad \alpha > 0$$

with initial data $u(x,0) = u_0(x)$, defined for $0 \le t \le T$ for some $T > 0$. The purpose of this exercise[23] is to derive a sufficient condition to ensure that the solution $u(x,t)$ blows up in $H^1(\mathbb{R}^n)$ in finite time, i.e. that there exists a $t_* > 0$ such that

$$\lim_{t \nearrow t_*} \|u(\cdot,t)\|_{H^1(\mathbb{R}^n)} = +\infty.$$

The specific argument here is due to Robert Glassey (J. Math. Phys. 1977).

(a) To begin, define the function $f : [0,T] \to \mathbb{R}$ by $f(t) := \int_{\mathbb{R}^n} |x|^2 |u(x,t)|^2 dx$. Show that for each $t \in (0,T)$ we have[24]

$$f'(t) = 4\Im \int_{\mathbb{R}^n} \bar{u}\,(x \cdot Du)\,dx.$$

In particular, conclude that $|f'(0)| \le C\|xu_0\|_{L^2(\mathbb{R}^n)}\|Du_0\|_{L^2(\mathbb{R}^n} < \infty$.

(b) (Suggested) Continuing, show that for each $t \in (0,T)$ we have

$$f''(t) = 16\mathcal{H}(u(\cdot,t)) + 4\left[\frac{2(n+2)}{\alpha+2} - n\right]\int_{\mathbb{R}^n} |u(x,t)|^{\alpha+2} dx,$$

where

$$\mathcal{H}(u) = \frac{1}{2}\int_{\mathbb{R}^n} |Du|^2 - \frac{1}{\alpha+2}\int_{\mathbb{R}^n} |u|^{\alpha+2} dx$$

(c) Suppose that $\mathcal{H}(u_0) < 0$ and that $\frac{4}{n} \le \alpha \le \frac{4}{n-2}$. Prove there exists[25] a $t_* > 0$ such that

$$\lim_{t \nearrow t_*} \int_{\mathbb{R}^n} |x|^2 |u(x,t)|^2 dx = 0.$$

(d) Next, prove the following: there exists a constant $C > 0$ such that, given any function $f : \mathbb{R}^n \to \mathbb{C}$ with $|x|f \in L^2(\mathbb{R}^n)$ and $Df \in L^2(\mathbb{R}^n)$, the inequality

$$\|f\|^2_{L^2(\mathbb{R}^n)} \le C\||x|f\|_{L^2(\mathbb{R}^n)}\|Df\|_{L^2(\mathbb{R}^n)}$$

holds.

[23] Note: You are not required to complete part (b) to this problem. However, you MUST complete ALL remaining parts of the problem.

[24] Here, $\Im(z)$ denotes the imaginary part of the complex number z.

[25] Here, you may use (without proof) the fact that $t_* \le T$. The fact that the solution actually exists up until $t = t*$ follows from the "blow up alternative" that we will discuss when we study semigroup methods.

(e) Conclude that if the initial data satisfies $|x|u_0 \in L^2(\mathbb{R}^n)$ and $\mathcal{H}(u_0) < 0$, then there exists[26] a $t_* > 0$ such that

$$\lim_{t \nearrow t_*} \|u(\cdot, t)\|_{H^1(\mathbb{R}^n)} = +\infty.$$

Hints:

(a) First, show that

$$f'(t) = 2 \int_{\mathbb{R}^n} |x|^2 \Re\left(\bar{u} u_t\right) dx,$$

where $\Re(z)$ denotes the real part of the complex number z. Recalling that u satisfies a particular PDE, as well as the identity (you should verify this!)

$$\bar{u}\Delta u = D \cdot (\bar{u} D u) - |Du|^2$$

the stated identity should follow.

(b) You can try to verify this if you want, but I warn you that it is pretty hard (I think). In any case, I consider finishing the remaining parts of the problem as more important.... you can always look up this inequality (and all the tricks necessary to derive it) if you want later!

(c) **IMPORTANT NOTE:** If you find it helpful, you may assume starting from this point in the problem that $\alpha = \frac{4}{n}$. In this case, the result from part (b) simply reads $f''(t) = 16\mathcal{H}(u(t))$.

To verify the claim in part (c), you need the **crucial fact**, unfortunatley not proven in class (due to time constraints...) that the functional $\mathcal{H}$ is a *conserved quantity* for the (H^1) flow induced by the given PDE. That is, if $u(x,t)$ is any (H^1) solution to the given PDE, then $\mathcal{H}(u(x,t))$ is independent of time[27]. The result should then follow from essentially integrating the results from part (a) and (b).

(d) Recall that, in $\mathbb{R}^n$ $D \cdot x = n$.

(e) Another **crucial fact** that was, again, unfortunately not proven in class, is that the functional

$$H^1(\mathbb{R}^n) \ni u \mapsto \int_{\mathbb{R}^n} |u|^2 dx \in \mathbb{R}$$

is also a *conserved quantity* for the (H^1) flow induced by the given PDE. That is, if $u(x,t)$ is any (H^1) solution to the given PDE, then $\int_{\mathbb{R}^n} |u(x,t)|^2 dx$ is independent of time[28]. Using this fact, along with the previous parts, you should find the desired result.

[26] As above, feel free to assume $t_* \leq T$.

[27] In essence, since $\mathcal{H}$ quantifies the "energy" of the system, this is a mathematical manifestation of the principle of conservation of energy.

[28] Essentially, this boils down to "conservation of mass" or "conservation of charge", depending on the context.

8 Appendix

8.1 The Sobolev Embedding Theorem

In this appendix, we discuss the proof of Theorem 7, which we restate here for completeness.

Theorem 14 (Sobolev Embedding Theorem). *If $\Omega \subset \mathbb{R}^n$ is an open (not necessarily bounded) domain in $\mathbb{R}^n$ and if $1 \leq p < n$, then*

$$W_0^{1,p}(\Omega) \subset L^q(\Omega)$$

is a continuous embedding for all $p \leq q \leq \frac{np}{n-p}$. In particular, there exists a constant $C = C(n,p,q) > 0$ such that

$$\|u\|_{L^q(\Omega)} \leq C\|u\|_{W^{1,p}(\Omega)}$$

for all $u \in W_0^{1,p}(\Omega)$.

Note in the above theorem that the domain Ω need not be bounded. Further, the result is obvious if $q = p$. To develop further, we begin by considering the case $\Omega = \mathbb{R}^n$ and asking when a Poincaré like result may hold on this unbounded domain. This is the content of the following famous result.

Theorem 15 (Gagliardo-Nirenbeg-Sobolev Inequality). *Assume that $1 \leq p < n$. Then there exits a constant $C = C(p,n) > 0$ such that*

$$\|u\|_{L^{p^*}(\mathbb{R}^n)} \leq C\|Du\|_{L^p(\mathbb{R}^n)}, \quad \text{where } \frac{1}{p} - \frac{1}{p^*} = \frac{1}{n}, \quad i.e.\ p^* := \frac{np}{n-p}$$

for all $u \in W^{1,p}(\mathbb{R}^n)$.

In Theorem 15, the number p^* is called the Sobolev conjugate of p. Note, in particular, that Theorem 15 implies that the Sobolev Embedding Theorem listed above holds when $q = p^*$. The proof of this theorem is quite long, but only relies on the Fundamental Theorem of Calculus and Hölder's inequality: see, for example, Theorem 1 in Section 5.1.6 in Evans, or Theorem 1 in Section 6.4(b) in McOwen. Here, I want to at least motivate the result through a scaling argument. To this end, let $1 \leq p < n$ and fix $u \in C_c^\infty(\mathbb{R}^n)$ and lets ask for what $q \in [1,\infty]$ could an inequality of the form

$$\|u\|_{L^q(\mathbb{R}^n)} \leq C\|Du\|_{L^p(\mathbb{R}^n)} \tag{29}$$

hold. To this end, for each $\lambda > 0$ consider the rescaled function

$$u_\lambda(x) = u(\lambda x)$$

and note that

$$\|u_\lambda\|^q_{L^q(\mathbb{R}^n)} = \int_{\mathbb{R}^n} |u(\lambda x)|^q dx = \lambda^{-n}\|u\|^q_{L^q(\mathbb{R}^n)}$$

and, similarly,
$$\|Du\|^p_{L^p(\mathbb{R}^n)} = \lambda^{p-n}\|Du\|^p_{L^p(\mathbb{R}^n)}.$$
Thus, if the inequality (29) were to hold, we would have
$$\lambda^{-n/q}\|u\|_{L^q(\mathbb{R}^n)} \leq C\lambda^{1-n/p}\|Du\|_{L^p(\mathbb{R}^n)}$$
valid for all $\lambda > 0$. By taking both $\lambda \to 0^+$ and $\lambda \to \infty$, it follows that such an inequality can only hold if
$$\lambda^{-n/q} = \lambda^{1-n/p}$$
holds for all $\lambda > 0$, i.e. if $q = p^* := \frac{np}{n-p}$. Of course, this only says that an inequality of the form (29) is possible when $q = p^*$, while in the rigorous proof one actually has to prove that such an inequality indeed holds.

Now, as stated above, the GNS inequality implies that the Sobolev Embedding Theorem holds when $q = p$ and when $q = p^*$. The proof of the Sobolev Embedding Theorem now follows from a relatively simple interpolation result. First, fix $1 \leq p < n$ and observe that using Hölder's inequality we have for all $1 \leq a \leq b \leq c$ that
$$\|u\|_{L^b(\mathbb{R}^n)} \leq \|u\|^\lambda_{L^a(\mathbb{R}^n)}\|u\|^{1-\lambda}_{L^c(\mathbb{R}^n)} \quad \text{for all } u \in C^\infty_c(\mathbb{R}^n)$$
where $\lambda \in [0,1]$ is chosen such that
$$\frac{1}{b} = \frac{\lambda}{a} + \frac{1-\lambda}{c}.$$
In words, by density, this inequality says that if you are in $L^a(\mathbb{R}^n)$ and $L^c(\mathbb{R}^n)$, then you are in $L^b(\mathbb{R}^n)$ for each b between a and c. Applying this L^p-interpolation result with $a = p$, $b = q$ and $c = p^*$ and using the GNS inequality, it follows that for all $u \in C^\infty_c(\mathbb{R}^n)$ we have
$$\|u\|_{L^q(\mathbb{R}^n)} \leq \|u\|^\lambda_{L^p(\mathbb{R}^n)}\|u\|^{1-\lambda}_{L^{p^*}(\mathbb{R}^n)} \leq C\|u\|^\lambda_{L^p(\mathbb{R}^n)}\|Du\|^{1-\lambda}_{L^p(\mathbb{R}^n)}$$
where $\lambda \in [0,1]$ is chosen so that
$$\frac{1}{q} = \frac{\lambda}{p} + \frac{1-\lambda}{p^*}.$$
In particular, notice the Sobolev embedding theorem with $q = p$ corresponds to $\lambda = 0$ while $q = p^*$ corresponds to $\lambda = 1$.

To establish the result for intermediate $q \in (p, p^*)$, we use Young's inequality
$$AB \leq \frac{A^s}{s} + \frac{B^t}{t} \quad \text{for all } A, B > 0 \text{ and } \frac{1}{s} + \frac{1}{t} = 1$$
with the choice $s = \frac{1}{\lambda}$ and $t = \frac{1}{1-\lambda}$, it follows that for all $u \in C^\infty_c(\mathbb{R}^n)$ and $q \in (p, p^*)$ we have
$$\begin{aligned}
\|u\|_{L^q(\mathbb{R}^n)} &\leq C\left(\|u\|^{2\lambda}_{L^p(\mathbb{R}^n)}\|Du\|^{2(1-\lambda)}_{L^p(\mathbb{R}^n)}\right)^{1/2} \\
&\leq C\left(\lambda\|u\|^2_{L^p(\mathbb{R}^n)} + (1-\lambda)\|Du\|^2_{L^p(\mathbb{R}^n)}\right)^{1/2} \\
&\leq C\max\{\lambda, 1-\lambda\}\|u\|_{W^{1,p}(\mathbb{R}^n)}.
\end{aligned}$$

All together, we this establishes the Sobolev Embedding Theorem for all $q \in [p, p^*]$ and $u \in C_c^\infty(\mathbb{R}^n)$, and the proof is then completed by an elementary density argument.