USING FUNCTIONAL ANALYSIS AND SOBOLEV SPACES TO SOLVE POISSON'S EQUATION

YI WANG

Abstract. We study Banach and Hilbert spaces with an eye towards defining weak solutions to elliptic PDE. Using Lax-Milgram we prove that weak solutions to Poisson's equation exist under certain conditions.

Contents

1. Introduction

We will discuss the following problem in this paper: let Ω be an open and connected subset in $\mathbb{R}$ and f be an L^2 function on Ω, is there a solution to Poisson's equation

$$-\Delta u = f? \tag{1}$$

From elementary partial differential equations class, we know if $\Omega = \mathbb{R}$, we can solve Poisson's equation using the fundamental solution to Laplace's equation. However, if we just take Ω to be an open and connected set, the above method is no longer useful. In addition, for arbitrary Ω and f, a C^2 solution does not always exist. Therefore, instead of finding a strong solution, i.e., a C^2 function which satisfies (1), we integrate (1) against a test function ϕ (a test function is a

Date: September 28, 2016.

smooth function compactly supported in Ω), integrate by parts, and arrive at the equation

$$\int_\Omega \nabla u \nabla \phi = \int_\Omega f\phi, \quad \forall \phi \in C_c^\infty(\Omega). \tag{2}$$

So intuitively we want to find a function which satisfies (2) for all test functions and this is the place where Hilbert spaces come into play. In the first 5 sections of the paper we will set the stage for the Hilbert spaces and in the last section we will utilize Hilbert spaces to solve the main problem. A solid background in real analysis is required for the full understanding of this paper.

2. Banach spaces

In this section we shall present definition and examples of Banach spaces as well as prove the famous Hahn-Banach theorems which enable us to extend linear functionals and separate sets in Banach spaces.

Definition 2.1. *A Banach space is a complete normed vector space.*

The motivation behind Banach spaces is that we want to generalize $\mathbb{R}^n$ to spaces of infinite dimensions. There are several characteristics of $\mathbb{R}^n$ which make us love them so much: they are linear spaces, they are metric spaces, and they are complete. All these 3 properties of $\mathbb{R}^n$ are included in the definition of Banach spaces.

Example 2.2. $\mathbb{R}^n$ *is a Banach space for any positive integer n, with the norm of the vector* $(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ *defined to be* $\sqrt{a_1{}^2 + a_2{}^2 + \ldots + a_n{}^2}$.

Definition 2.3. *Let* (X, σ, μ) *be a* σ*-finite measure space. For* $1 \leq p < +\infty$*, define the* L^p *norm of a function* f *by*

$$\|f\|_p \equiv \left(\int_X |f(x)|^p d\mu\right)^{1/p}.$$

For p *with* $p = +\infty$*, define the* L^p *norm of* f *by*

$$\|f\|_\infty \equiv \inf\{M \mid \mu(\{x \mid |f(x)| > M\}) = 0\}.$$

The space L^p *is defined to be*

$$L^p \equiv \{f : X \to \mathbb{R} | f \text{ measurable and } \|f\|_p < +\infty\}.$$

Remark 2.4. *For any* p *such that* $1 \leq p \leq +\infty$*,* L^p *is a Banach space.*

In the following, E denotes a Banach space.

Definition 2.5. *Given a linear function, $f : E \to \mathbb{R}$, the norm of f, denoted by $\|f\|$, is defined to be*

$$\|f\| \equiv \sup_{x\in E, \|x\|\leq 1} |f(x)|.$$

If $\|f\| < +\infty$, we say f is a bounded linear functional on E.

Remark 2.6. *It is not hard to check that a bounded linear functional f is also a continuous linear functional, and vice versa.*

Definition 2.7. *The dual space of E, denoted by E^*, is defined to be the collection of all bounded linear functionals on E with the norm given above.*

Example 2.8. *Let $E = \mathbb{R}$. It is easy to see for every $f \in E^*$, there is an $r_f \in E$ such that $f(x) = r_f x$ for all $x \in E$, and the converse is also true. Thus, we can identify E^* with E.*

Theorem 2.9 (Hahn-Banach analytic form)**.** *Let $p : E \to \mathbb{R}$ be a function satisfying*

$$p(\lambda x) = \lambda p(x) \quad \forall x \in E \quad \forall \lambda > 0 \tag{3}$$

$$p(x + y) \leq p(x) + p(y) \quad \forall x, y \in E \tag{4}$$

Let $G \subset E$ be a linear subspace and let $g : G \to \mathbb{R}$ be a linear functional such that $g(x) \leq p(x) \quad \forall x \in G$. Then there exists a linear functional $f : E \to \mathbb{R}$ which extends g and

$$f(x) \leq p(x) \quad \forall x \in E.$$

Proof. Suppose not. Let X be defined as the collection of linear extensions of g such that $f(x) \leq p(x)$ on the domain of f. For any two elements, f_1 and f_2 in X, we say $f_1 < f_2$ if and only if $domain(f_1) \subset domain(f_2)$, and f_2 is an extension of f_1. Then for any chain $f_1 < f_2 < f_3 < ......$, let $D_1 \subset D_2 \subset ......$ be their domains and let D be defined as $\cup_i D_i$. It is clear that D is a linear subspace. We define a function $f : D \to \mathbb{R}$ as follows: for any $x \in D$, by our construction there exists an i such that $x \in D_i$. we define $f(x)$ to be $f_i(x)$. It is easy to verify that f is well defined and $f \in X$. Thus, the chain $f_1 < f_2 < ......$ has an upper bound in X. By applying Zorn's lemma, we know there exists a maximal element h which is an extension of g to some linear subspace G_1 of E.

Suppose $G_1 \neq E$. Choose $x_0 \in E - G_1$, and consider the subspace

$$G_2 \equiv \{x + \lambda x_0 \big| \ x \in G_1, \ \lambda \in \mathbb{R}\}.$$

It is clear that G_2 is a linear subspace of E which is strictly larger than G_1. In addition, for all x, y in G_1,

$$h(x) + h(y) = h(x + y) \leq p(x + y) \leq p(x + x_0) + p(y - x_0).$$

Hence,

$$h(y) - p(y - x_0) \leq p(x + x_0) - h(x).$$

Thus, we can choose an $\alpha \in \mathbb{R}$ such that

$$h(y) - p(y - x_0) \leq \alpha \leq p(x + x_0) - h(x)$$

for all $x, y \in G_1$. Define a linear function $f : G_2 \to \mathbb{R}$ by $f(y + \lambda x_0) \equiv h(y) + \lambda\alpha,\ \forall y \in G_1,\ \lambda \in \mathbb{R}$. It is obvious that f is an extension of h to G_2. In addition, by our construction of α, we have $f(x + x_0) = h(x)+\alpha \leq p(x+x_0)$ and $f(x-x_0) = h(x)-\alpha \leq p(x-x_0) \quad \forall x \in G_1$. By (1) and the fact that h is a linear functional defined on a linear subspace of E, it is clear that $f(x + \lambda x_0) \leq p(x + x_0)\ \forall x \in G_1$ *and* $\lambda \in \mathbb{R}$, which contradicts the maximality of h. □

Definition 2.10. *a hyperplane H is a subset of E of the form*

$$H = \{x \in E |\ f(x) = \alpha\},$$

where f is a nontrivial linear functional and α is a constant in $\mathbb{R}$.

It is not hard to show that H is a closed hyperplane if and only if its corresponding f is bounded.

Now, we can use the analytic form of Hahn-Banach theorem to separate sets. Before we start, we look at a few definitions and lemmas which help us to prove the geometric form of Hahn-Banach theorem.

Definition 2.11. *Let A and B be two subsets of E. We say that the hyperplane, $H = [f = \alpha]$, separates A and B if*

$$f(x) \leq \alpha \quad \forall\, x \in A \quad \textit{and} \quad f(x) \geq \alpha \quad \forall x \in B.$$

We say that H strictly separates A and B if there exists some $\epsilon > 0$ such that

$$f(x) \leq \alpha - \epsilon \quad \forall x \in A \quad \textit{and} \quad f(x) \geq \alpha + \epsilon \quad \forall x \in B.$$

Lemma 2.12. *Let $C \subset E$ be an open convex set with $0 \in C$. For every $x \in E$ set*

$$p(x) \equiv \inf\{a > 0 \,\big|\, a^{-1}x \in C\}.$$

We call p the gauge of C. Then p satisfies (3) (4) and the following properties:

$$\exists M \textit{ such that } 0 \leq p(x) \leq M\|x\|, \quad \forall x \in E \tag{5}$$

$$C = \{x \in E \,\big|\, p(x) < 1\} \tag{6}$$

Proof. It is clear that p is linear, so it satisfies (3). (4) follows from convexity of C. (5) is true because C is open and (6) follows from definition of p. □

Lemma 2.13. *Let $C \subset E$ be a nonempty open convex set and let x_0 be an element in E with $x_0 \notin C$. Then there exists $f \in E^*$ such that $f(x) < f(x_0) \quad \forall x \in C$. In particular, the hyperplane, $H = [f = f(x_0)]$, separates $\{x_0\}$ and C.*

Proof. After a translation we may assume $0 \in C$. We introduce the gauge of C, which we denote by p. Consider the linear subspace $G = \{\lambda x_0 \mid \lambda \in \mathbb{R}\}$ and the linear functional $g : G \to \mathbb{R}$ defined by

$$g(tx_0) = t, \ t \in \mathbb{R}.$$

It is clear that

$$g(x) \leq p(x) \quad \forall x \in G.$$

By Theorem 2.9 we know we can extend g to f defined on E such that

$$f(x) \leq p(x) \quad \forall x \in G.$$

So we must have

(1) $f(x_0) = 1$ (f is an extension of g and $g(x_0) = 1$).
(2) $f(x) \leq p(x) \leq 1 \quad \forall x \in C$ (p is the gauge of C and (4) is true).
(3) f is continuous (by (4)).

Thus, we are done. □

Corollary 2.14 (Hahn-Banach, first geometric form)**.** *let $A \subset E$ and $B \subset E$ be two nonempty convex subsets such that $A \cap B = \emptyset$. Assume that one of them is open. Then there exists a closed hyperplane that separates A and B.*

Proof. Assume A is open. Let $C = \{x - y;\ x \in A, y \in B\}$. As A and B are convex, it is clear that C is convex. As $C = \cup_{x \in B}(A - x)$, the union of open sets, we know C is open. In addition, C does not contain 0 as $A \cap B = \emptyset$. By Lemma 2.13, there exists an $f \in E^*$ such that $f(x) < 0$, $\forall x \in C$, which implies $f(x) < f(y) \quad \forall x \in A, y \in B$. Let $\alpha = sup_{x \in A} f(x)$. We know $[f = \alpha]$ is a hyperplane which separates A and B. □

Corollary 2.15 (Hahn-Banach, second geometric form)**.** *let $A \subset E$ and $B \subset E$ be two nonempty convex subsets such that $A \cap B = \emptyset$. Assume that A is closed and B is compact. Then there exists a closed hyperplane that strictly separates A and B.*

Proof. The proof is similar to that of Corollary 2.14. □

3. Weak topology, weak star topology and reflexivity

In this section I will introduce definitions of weak and weak star topology. The motivation behind those topologies is that a topology with fewer open sets has more compact sets.

In the following, E denotes a Banach space.

Definition 3.1. *The weak topology $\sigma(E, E^*)$ on E is defined to be the coarsest topology on E such that for all f in E^*, f is continuous.*

Remark 3.2. *Such a topology exists. Consider $A = \{$all topologies on E such that for all f in E^*, f is continuous $\}$. We know A is nonempty as it contains the discrete topology on A. We define $\sigma(E, E^*)$ to be the intersection of all elements in A and it is easy to verify we get the topology we want. In addition, the weak topology is coarser than the strong topology as for any $f \in E^*, f$ is continuous with respect to the strong topology by Remark 2.6.*

Proposition 3.3. *Let (x_n) be a sequence in E. Then*

(1) *$x_n \rightharpoonup x$ weakly in $\sigma(E, E^*)$ if and only if $f(x_n) \to f(x)$, $\forall f \in E^*$.*
(2) *if (x_n) converges strongly, then (x_n) converges weakly in $\sigma(E, E^*)$.*

Here is an example that shows the weak topology is strictly coarser than the strong topology.

Definition 3.4. *Let ℓ^2 denote the collection of sequences $(x_1, x_2, ...) \in \mathbb{R}^\infty$ such that $\sum_{i=1}^{\infty} {x_i}^2 < \infty$. We define a bilinear form $(-,-)$ on ℓ^2 by*

$$\forall a = (a_1, a_2, ...),\ b = (b_1, b_2, ...) \in \ell^2, (a, b) \equiv \sum_{i=1}^{\infty} a_i b_i.$$

It is easy to verify that ℓ^2 is a Hilbert space (Definition 5.2) under the scalar product $(-,-)$.

Example 3.5. *Consider a sequence $(e_n) \in \ell^2$, where $e_1 = (1, 0, 0, 0, ...)$, $e_2 = (0, 1, 0, 0, ...)$, $e_3 = (0, 0, 1, 0, ...)$ etc. It is clear that (e_n) does not converge in the strong topology. However, as we will prove later (Theorem 5.9), for any $f \in (\ell^2)^*$, there exists an element $F \in \ell^2$ such that $f(x) = (F, x) \quad \forall x \in \ell^2$ and then it is easy to show that $f(e_n) \to f(0) = 0$. Thus, (e_n) converges to 0 in the weak topology.*

Here is another example that shows the weak topology is strictly coarser than the strong topology.

Example 3.6. *The unit sphere $S = \{x \in E \mid \|x\| = 1\}$, with E infinite-dimensional, is not closed in $\sigma(E, E^*)$. More precisely, the closure of S with respect to the weak topology $\sigma(E, E^*)$ is B_E.*

Proof. If a set is closed, its closure should be the same as itself. So it suffices for us to show the second part, namely, $\overline{S} = B_E$.

We prove $B_E \subset \overline{S}$ first. For any $x_0 \in B_E$, we choose a neighborhood V of x. We may assume V is of the form

$$V = \{x \in E \big| f_i(x - x_0) < \epsilon \text{ for } i = 1, 2, 3, ...n, f_i \in E^* \text{and } \epsilon > 0\}.$$

Choose a $y_0 \in E$ such that $f_i(y_0) = 0 \; \forall \; i = 1, 2, 3....., n$. We know such a y_0 exists or the function from E to $\mathbb{R}^n$ sending $x \in E$ to $(f_1(x), f_2(x), ..., f_n(x))$ would be an injection from E, a space of infinite dimension, to a space of finite dimension, which is a contradiction. Define a function $g : \mathbb{R} \to \mathbb{R}$ by

$$g(t) = \|x_0 + ty_0\|.$$

It is clear that g is a continuous function such that $g(0) \leq 1$ and $\lim_{t\to+\infty} g(x) = +\infty$. So there exists a t such that $\|x_0 + ty_0\| = 1$. So $x_0 + ty_0$ is in both V and S. Thus, $V \cap S$ is nonempty and we are done.

To prove the other direction, it suffices to show B_E is closed. As

$$B_E = \cap_{f\in E^*,\|f\|\leq 1}\{x \in E \mid |f(x)| \leq 1\},$$

writing equation we see that B_E is an intersection of closed sets and thus closed. □

Lemma 3.7. *let Z be a topological space, E be a Banach space with weak topology and let ϕ be a function from Z to E. Then ϕ is continuous if and only if $f \circ \phi$ is continuous for all f in E^*.*

Proof. The proof follows easily from the definition of weak topology. □

Theorem 3.8. *Let C be a convex subset of E, then C is closed in the weak topology $\sigma(E, E^*)$ if and only if it is closed in the strong topology.*

Proof. As the weak topology is coarser than the strong topology, it suffices for us to prove the converse. Let C be a closed and convex set in the strong topology and let $x_0 \in E$ be a point in C^c. As C is a closed and convex set, $\{x_0\}$ is a closed convex and compact set, by Theorem 2.15 we know there exists an $\alpha \in \mathbb{R}$ and a bounded linear functional $f \in E^*$ such that

$$f(x_0) < \alpha < f(y) \quad \forall y \in C.$$

$V = \{x \in E \mid f(x) < \alpha\}$ is an open set in the weak topology which contains x_0 and disjoint from C. Thus, C^c is open, or C is closed. □

Remark 3.9. *Let J be a function from E to E^{**} defined as follows: For any $x \in E$, $J(x) \in E^{**}$ is the element which takes $g \in E^*$ to*

$g(x)$. It is easy to verify that for any $x \in E$, $J(x)$ defined above is a bounded linear functional on E^*, thus, is an element in E^{**}. Moreover, J preserves norms, i.e.,

$$\forall x \in E, \|x\|_E = \|J(x)\|_{E^{**}}.$$

We call J the canonical injection of E into E^{**}.

Remark 3.10. *J is clearly an injection, but not necessarily a surjection. For example, J defined on the space*

$$\ell^1 = \{(x_n) \in \mathbb{R}^\infty | \sum_{i=1}^{\infty} |x_i| < \infty\}$$

is not surjective (the proof involves showing $(\ell^1)^* = \ell^\infty$ and finding an element in $(\ell^\infty)^*$ which is not in ℓ^1). This property of J is the motivation of another topology coarser than the weak topology: the weak star topology.

Definition 3.11. *The weak star topology $\sigma(E^*, E)$ is defined to be the coarsest topology on E^* such that for all $x \in E$, the function $J(x)$ is continuous (See Remark 3.9 for definition).*

Proposition 3.12. *Let (f_n) be a sequence in E^*. Then*

1. *$f_n \overset{*}{\rightharpoonup} f$ in $\sigma(E^*, E)$ if and only if $f_n(x) \to f(x) \quad \forall x \in E$.*
2. *If $f_n \to f$ strongly, then $f_n \rightharpoonup f$ weakly. If $f_n \rightharpoonup f$ weakly, then $f_n \overset{*}{\rightharpoonup} f$.*

Proof. The proof is similar to that of Proposition 3.3, except that the second part uses the canonical injection J. □

Remark 3.13. *From above we can define 3 different topologies on E^*: the strong topology, the weak topology and the weak star topology. The following theorem will show that the closed unit ball on E^* is compact with the weak star topology (which is not always the case with the weak topology and never the case with the strong topology for space E with infinite dimensions). In fact, for any Banach space E such that E is not reflexive (See Definition 3.15), the unit ball is not compact in the weak topology (See Theorem 3.18).*

Theorem 3.14 (Banach-Alaoglu-Bourbaki)**.** *The closed unit ball,*

$$B_{E^*} = \{f \in E^* \mid \|f\| \leq 1\},$$

is compact in $\sigma(E^*, E)$.

Proof. $Y = \mathbb{R}^E$ (the collection of all functions from E to $\mathbb{R}$) equipped with the product topology. We denote an element ω in Y by a sequence

$(\omega_x)_{x\in E}$. Define a function $\phi : E^* \to Y$ such that, $\forall f \in E^*, \phi(f) = (f(x))_{x\in E}$. It is clear that ϕ is a surjection from E^* to $\phi(E^*)$. Moreover, for any two elements $f, g \in E^*$, if $\phi(f) = \phi(g)$, then by definition of ϕ, for any $x \in E$, we have $f(x) = g(x)$, which means $f = g$. Therefore, ϕ is also an injection. We claim ϕ is a homeomorphism. It suffices to show that ϕ and ϕ^{-1} are continuous. Referring to Lemma 3.7 and noticing that $\forall x \in E$, the function $E^* \to \mathbb{R}$ sending f to $f(x)$ is continuous, we know ϕ is continuous. The continuity of ϕ^{-1} can also be proven by using similar means. Thus, ϕ is a homeomorphism from E^* to $\phi(E^*)$. Let A be a subset of Y defined as follows:

$$A = \{\omega \in Y |\ |\omega_x| \le \|x\|, \omega_{x+y} = \omega_x + \omega_y, \omega_{\lambda x} = \lambda\omega_x, \forall x, y \in E \textit{ and } \forall \lambda \in \mathbb{R}.\}$$

It is clear that $\phi(B_{E^*}) = A$, so it suffices to show that A is compact. Define B, C as follows:

$$B = \{\omega \in Y ||\omega_x| \le \|x\|, \forall x \in E\},$$

$$C = \{\omega \in Y |\omega_{x+y} = \omega_x + \omega_y, \omega_{\lambda x} = \lambda\omega_x, \forall x, y \in E \textit{ and } \forall \lambda \in \mathbb{R}.\}$$

Then by Tychonoff's theorem (See [4, chapter 5]), B is compact. As C is intersection of closed sets, C is closed, and $A = B \cap C$ is compact. ☐

Definition 3.15. *Let E be a Banach space and let $J : E \to E^{**}$ be the canonical injection from E into E^{**}. The space E is said to be reflexive if J is surjective, i.e., $J(E) = E^{**}$. An example of a non-reflexive space is given in Remark 3.10.*

In the last part of section 3 I will present several results which will contribute to the proof of Corollary 4.6.

Lemma 3.16. *Let E be a Banach space. Let $f_1, f_2, ..., f_k$ be given in E^* and let $\gamma_1, \gamma_2, ..., \gamma_k$ be given in $\mathbb{R}$. The following properties are equivalent:*

(1) *$\forall \epsilon > 0$, $\exists x_\epsilon \in E$ such that $\|x_\epsilon\| \le 1$ and*

$$|f_i(x_\epsilon) - \gamma_i| < \epsilon, \quad \forall i = 1, 2, ..., k.$$

(2) $|\sum_{i=1}^k \beta_i\gamma_i| \le \|\sum_{i=1}^k \beta_i f_i\| \quad \forall \beta_1, \beta_2, ..., \beta_k \in \mathbb{R}$.

Proof. (1) $\to$ (2): Fix $\beta_1, \beta_2, ..., \beta_k \in \mathbb{R}$, let

$$S = \sum_{i=1}^k |\beta_i|,$$

from (1) we have

$$|\sum_{i=1}^k \beta_i f_i(x_\epsilon) - \sum_{i=1}^k \beta_i\gamma_i| \le \epsilon S,$$

which implies

$$|\sum_{i=1}^{k}\beta_i\gamma_i| \le \|\sum_{i=1}^{k}\beta_i f_i\|\|x_\epsilon\| + \epsilon S \le \|\sum_{i=1}^{k}\beta_i f_i\| + \epsilon S.$$

Let ϵ goes to 0 and we obtain (2).

(2) → (1): Let $\gamma = (\gamma_1, \gamma_2, ..., \gamma_k)$ and consider $\phi : E \to \mathbb{R}^k$ defined by

$$\phi(x) = (f_1(x), f_2(x), ..., f_k(x)).$$

(1) is equivalent to $\gamma \in$ closure of $\phi(B_E)$. Suppose by contradiction that this is false, then the closure of $\phi(B_E)$ is a closed convex set and $\{\gamma\}$ is a compact convex set which is disjoint from the closure. Thus by Theorem 2.15 we can separate them strictly by a closed hyperplane $[f = \alpha]$ for some $f \in E^*$ and $\alpha \in \mathbb{R}$, which means there exists a $\beta = (\beta_1, \beta_2, ..., \beta_k) \in \mathbb{R}^k$ such that

$$\sum_{i=1}^{k}\beta_i f_i(x) < \alpha < \sum_{i=1}^{k}\beta_i\gamma_i, \quad \forall x \in B_E,$$

and therefore

$$\|\sum_{i=1}^{k}\beta_i f_i\| \le \alpha < \sum_{i=1}^{k}\beta_i\gamma_i,$$

a contradiction. □

Lemma 3.17. *$J(B_E)$ is dense in $B_{E^{**}}$ with respect to the $\sigma(E^{**}, E^*)$ topology.*

Proof. Let $\theta \in B_{E^{**}}$, and let V be a neighborhood of θ. We might assume

$$V = \{g \in B_{E^{**}} | |g(f_i) - \theta(f_i)| < \epsilon, \quad \forall i = 1, 2, ..., k\}.$$

We want to show that V intersects $J(B_E)$ non-trivially, i.e., we want to find $x \in E$ such that

$$|f_i(x) - \theta(f_i)| < \epsilon \quad \forall i = 1, 2, ..., k.$$

Define $\gamma_i = \theta(f_i)$ for $i = 1, 2, ..., k$. By the previous lemma we know it suffices to show

$$|\sum_{i=1}^{k}\beta_i\gamma_i| \le \|\sum_{i=1}^{k}\beta_i f_i\| \quad \forall \beta_i \in \mathbb{R},$$

which is clear since $\sum_{i=1}^{k}\beta_i\gamma_i = \theta(\sum_{i=1}^{k}\beta_i f_i)$ and $\|\theta\| \le 1$. □

Theorem 3.18. *Let E be a Banach space. Then E is reflexive if and only if*

$$B_E = \{x \in E \mid \|x\| \leq 1\}$$

is compact with the weak topology $\sigma(E, E^*)$.

Proof. Suppose E is reflexive. As J preserves norm, by Theorem 3.14, it suffices to show J is a homeomorphism from E equipped with weak topology $\sigma(E, E^*)$ to E^{**} equipped with weak star topology $\sigma(E^{**}, E^*)$, which is easy if we use Lemma 3.7 and reflexivity of E.

Now suppose the unit ball is compact with the weak topology. If we can show $J(B_E) = B_{E^{**}}$, then we are done. From the forward direction proven above, we know $J(B_E)$ is compact, thus closed in the weak star topology. So it suffices to show it is also dense in $B_{E^{**}}$, which we have already proven in Lemma 3.17. □

Corollary 3.19. *Let E be a reflexive Banach space. Let $K \subset E$ be a bounded, closed, and convex subset of E. Then K is compact in $\sigma(E, E^*)$.*

Proof. By Theorem 3.8 we know K is closed for the topology $\sigma(E, E^*)$. As K is bounded, there exists an m such that $K \subset mB_E$, and mB_E is compact by Theorem 3.14. □

4. Lower semicontinuity

In this section the definition of lower semicontinuity will be introduced. We will see that on a reflexive Banach space, lower semicontinuity of a functional guarantees the existence of minimizer under certain conditions.

Definition 4.1. *A function, $f : E \to \mathbb{R}$, is lower semicontinuous (l.s.c) if for every $x \in E$ and for every $\epsilon > 0$ there is some neighborhood V of x such that*

$$f(y) \geq f(x) - \epsilon \quad \forall y \in V$$

In particular, if f is l.s.c, then for every sequence $(x_n) \in E$ such that $x_n \to x$, we have

$$\liminf_{n\to\infty} f(x_n) \geq f(x).$$

Lemma 4.2. *Given a function $\phi : E \to (-\infty, +\infty]$, the following property is equivalent to the definition we gave for a function to be lower semicontinuous:*

$$\forall \lambda \in \mathbb{R}, \text{the set } A_\lambda = \{x \in E \mid \phi(x) \leq \lambda\} \text{ is closed.} \tag{7}$$

Proof. The proof is a standard analysis argument. □

Example 4.3. *Define a function $f : \mathbb{R} \to \mathbb{R}$ by*

$$f(x) = \begin{cases} 1, & x > 0 \\ -1, & x \leq 0 \end{cases} \tag{8}$$

It is easy to verify that f is lower semicontinuous.

Corollary 4.4. *Assume that $\phi : E \to (-\infty, +\infty]$ is convex and l.s.c. with respect to the strong topology, Then ϕ is l.s.c with respect to the weak topology.*

Proof. $\forall \lambda \in \mathbb{R}$, the set

$$A = \{x \in E \mid \phi(x) \leq \lambda\}$$

is convex (by convexity of ϕ) and closed (by l.s.c in the strong topology) in the strong topology. Thus, it is also closed in the weak topology by Theorem 3.8. As the choice of λ is arbitrary, by Lemma 4.2, ϕ is l.s.c in the weak topology. □

Lemma 4.5. *If E is compact and f is l.s.c, then $\inf_E f$ is achieved.*

Proof. The proof is a standard analysis argument. □

What follows is a big theorem which guarantees the existence of minimizer under certain conditions. We will use it to prove Corollary 5.14.

Theorem 4.6. *Let E be a reflexive Banach space. Let $A \subset E$ be a nonempty, closed, and convex subset of E. Let $\phi : A \to (-\infty, +\infty]$ be a convex l.s.c. function such that $\phi \not\equiv \infty$ and*

$$\lim_{\substack{x \in A \\ \|x\| \to \infty}} \phi(x) = +\infty. \tag{9}$$

Then ϕ achieves its minimum on A, i.e. there exists some x_0 in A such that

$$\phi(x_0) = \inf_{x \in A} \phi(x).$$

Proof. Choose an $a \in A$ such that $\phi(a) < +\infty$ and define the set

$$B = \{x \in A \mid \phi(x) < \phi(a)\}.$$

We know B is closed (by semicontinuity of ϕ), convex (by convexity of ϕ), and bounded (by 9). Thus, by Corollary 3.19 it is compact in the weak topology. And by Corollary 4.4, ϕ is also l.s.c. in the weak topology. We apply Lemma 4.5 and know ϕ achieves minimum on B, which clearly implies that it achieves minimum on A. □

5. Hilbert spaces

Definition 5.1. *Let H be a vector space. A scalar product, (u, v), is a bilinear form on $H \times H$ with values in $\mathbb{R}$ such that*

$$(10) \qquad \begin{aligned} (u, v) &= (v, u) \quad \forall u, v \in H \\ (u, u) &\geq 0 \quad \forall u \in H \\ (u, u) &\neq 0 \quad \forall u \neq 0. \end{aligned}$$

It is not hard to see that the quantity,

$$\|u\| = \sqrt{(u, u)},$$

is a norm.

Definition 5.2. *A Hilbert space is a vector space H equipped with a scalar product such that H is complete under the norm $\| \; \|$.*

Example 5.3. *For any positive integer n, the space $\mathbb{R}^n$ is a Hilbert space with scalar product defined as the inner product.*

Example 5.4. *$L^2(\mathbb{R})$ is a Hilbert space with scalar product defined as*

$$(f, g) = \int_R fg, \quad \forall f, g \in L^2(\mathbb{R}).$$

In the following, H denotes a Hilbert space.

Remark 5.5. *All Hilbert spaces are Banach spaces. In addition to the properties of a Banach space, a Hilbert space also has "angles". For example, if $f, g \in H$ satisfies $(f, g) = 0$, we say f is orthogonal to g. Using this newly acquired property, we can deduce more characteristics of Hilbert spaces. The following theorem shows that the projection we see in $\mathbb{R}^n$ also exists in a Hilbert space.*

Theorem 5.6. *Let $K \subset H$ be a nonempty closed convex set. Then for every $f \in H$ there exists a unique element $u \in K$ such that*

$$(11) \qquad |f - u| = \min_{v \in K} |f - v|.$$

Moreover, u is characterized by the property

$$(12) \qquad u \in K \text{ and } (f - u, v - u) \leq 0, \quad \forall v \in K.$$

Notation. *The above element u is denoted by*

$$u = P_K f.$$

Proof. (1) Fix $f \in H$, define $\phi : H \to \mathbb{R}$ by $\phi(x) = |f - x|$. It is clear that ϕ is convex (by the triangle inequality of norm), continuous (thus l.s.c) and

$$\lim_{|x| \to +\infty} \phi(x) = +\infty.$$

By Corollary 4.6, we know ϕ achieves a minimum.

(2) We want to show (11) and (12) are equivalent.

Suppose u minimizes ϕ. Then for any v in K and for any $t \in [0, 1]$, by convexity of K, we know that

$$w = tu + (1-t)v \in K.$$

As u is the minimizer of ϕ on K, we know

$$|f - u| \le |f - w| = |(f - u) - (1 - t)(v - u)|,$$

which implies

$$|f - u|^2 \le |f - u|^2 + (1 - t)^2 |v - u|^2 - 2(1 - t)(f - u, v - u),$$

or equivalently

$$2(f - u, v - u) \le (1 - t)|v - u|^2$$

for all t in $[0, 1]$. Let t go to 0, the claim is true.

Suppose there exists an u in K satisfying (12). Then we have $|f - u|^2 - |f - v|^2 = 2(f - u, v - u) - |u - v|^2 \le 0$ for all $v \in K$.

(3) We claim that such an u is unique.

Suppose we have two minimizers u and v. Then by (2) we have

$$(f - u, w - u) \le 0 \quad \forall w \in K$$

and

$$(f - v, w - v) \le 0 \quad \forall w \in K.$$

Putting $v = w$ in the first inequality and $u = w$ in the second inequality and add them we get $|v - u|^2 \le 0$, which implies $v = u$.

□

The next lemma shows that the projection map does not increase distance.

Lemma 5.7. *Let $K \subset H$ be a nonempty closed convex set. Then*

$$|P_K f_1 - P_K f_2| \le |f_1 - f_2|, \quad \forall f_1, f_2 \in H$$

.

Proof. Define $u_1 = P_K f_1$ and $u_2 = P_K f_2$. By Theorem 5.6, we have

$$(f_1 - u_1, w - u_1) \le 0, \quad \forall w \in K, \tag{13}$$

$$(f_2 - u_2, w - u_2) \le 0, \quad \forall w \in K. \tag{14}$$

Select $w = u_2$ in (13) and $w = u_1$ in (14) and add the two inequalities together, we have

$$(f_1 - f_2 + u_2 - u_1, u_2 - u_2) \le 0,$$

or equivalently

$$|u_2 - u_1|^2 \le (f_1 - f_2, u_1 - u_2),$$

which implies

$$|u_1 - u_2| \le |f_1 - f_2|.$$

□

Corollary 5.8. *Assume that $M \subset H$ is a closed linear subspace. Let $f \in H$. Then $u = P_M f$ is characterized by*

$$u \in M \quad \textit{and} \quad (f - u, v) = 0, \quad \forall v \in M$$

Proof. A linear subspace is also a convex set, thus we can apply Theorem 5.6 and get

$$(f - u, v - u) \le 0 \quad \forall v \in M.$$

For any $t \in R$, tv is also in M. So

$$(f - u, tv - u) \le 0 \quad \forall v \in M, \quad \forall t \in R.$$

So we must have

$$(f - u, v) = 0, \quad \forall v \in M.$$

The converse is also easy to show. □

Theorem 5.9. *Given any $\phi \in H^*$, there exists a unique $f \in H$ such that*

$$\phi(u) = (f, u) \quad \forall u \in H.$$

Moreover,

$$|f| = \|\phi\|_{H^*}.$$

Proof. Choose a $\phi \in H^*$ If

$$\phi(x) = 0, \quad \forall x \in H,$$

we can choose $f = 0$ and we are done. If not, let M, a subset of H, be defined as follows:

$$M = \{x \in H; \phi(x) = 0\}.$$

As $M = \phi^{-1}(0)$, we know M is closed. It is clear that M is also a subspace. By assumption, there exists an $x \in H$ but not in M. Let $y = P_M x$ and $z = x - y$ ($z \notin M$ so $\phi(z) \neq 0$). By Corollary 5.8 we have

$$(z, v) = 0 \quad \forall v \in M.$$

Besides, it is clear that $z \neq 0$. So $(z, z) \neq 0$. Define $\alpha = \frac{\phi(z)}{(z,z)}$. Let $x_0 = \alpha z$. It is clear that

$$\phi(x_0) = (x_0, x_0) \neq 0.$$

Claim:

$$\forall x \in H, \alpha(x) = (x_0, x).$$

proof: for any $x \in H$, write it in the form

$$(x - \frac{\phi(x)}{(x_0, x_0)} x_0) + \frac{\phi(x)}{(x_0, x_0)} x_0.$$

As

$$\phi(x - \frac{\phi(x)}{(x_0, x_0)} x_0) = \phi(x) - \frac{\phi(x)}{(x_0, x_0)} \phi(x_0) = 0,$$

we have

$$x - \frac{\phi(x)}{(x_0, x_0)} x_0 \in M.$$

Thus,

$$(x_0, x) = (x_0, (x - \frac{\phi(x)}{(x_0, x_0)} x_0) + \frac{\phi(x)}{(x_0, x_0)} x_0) = (x_0, x - \frac{\phi(x)}{(x_0, x_0)} x_0) + (x_0, \frac{\phi(x)}{(x_0, x_0)} x_0)$$

$$= (x_0, \frac{\phi(x)}{(x_0, x_0)} x_0) = \phi(x)$$

and we are done. □

Remark 5.10. *We can identify H and H^* by Theorem 5.9.*

Definition 5.11. *A bilinear form $\alpha : H \times H \to \mathbb{R}$ is said to be continuous if there is a constant C such that*

$$|\alpha(u, v)| \leq C|u||v| \qquad \forall u, v \in H;$$

coercive if there is a constant $a > 0$ such that

$$\alpha(v, v) \geq a|v|^2 \qquad \forall v \in H.$$

Remark 5.12. *It is obvious that for any Hilbert space H, the scalar product $(-, -)$ on H is a continuous and coercive bilinear form. It is also true that for any positive definite $n \times n$ matrix, the inner product $\alpha(\vec{u}, \vec{v}) \equiv (A \cdot \vec{u}) \cdot \vec{v}^{\perp}$ is coercive and continuous on $\mathbb{R}^n$*

Theorem 5.13 (Contraction Mapping Theorem). *Let X be a nonempty complete metric space and let $f : X \to X$ be a strict mapping, i.e. there exists $k \in (0, 1)$ such that*

$$d(f(x), f(y)) \leq kd(x, y) \quad \forall x, y \in X.$$

Then there exists a unique $x \in X$ such that $f(x) = x$.

Proof. See [3, chapter 9]. □

We are now ready to prove the Lax-Milgram theorem. We will use it in Section 6 to prove the existence of weak solution to the Laplace equation.

Corollary 5.14 (Lax-Milgram). *Assume that $\alpha(u, v)$ is a continuous coercive bilinear form on H. Then, given any $\phi \in H^*$, there exists a unique element $u \in H$ such that*

$$\alpha(u, v) = \phi(v) \qquad \forall v \in H.$$

Moreover, if α is symmetric, then u is characterized by the property

$$u \in K \qquad \textit{and} \qquad \frac{1}{2}\alpha(u, u) - \phi(u) = \min_{v \in K}\{\frac{1}{2}\alpha(v, v) - \phi(v).\}$$

Proof. For any $\phi \in H^*$, by Theorem 5.9, we know there exists a g in H such that

$$\phi(x) = (g, x) \quad \forall x \in H.$$

In addition, if we fix u, $\alpha(u, -)$ is a continuous linear functional on H. Thus, there exists a unique element in H denoted by A_u such that

$$\alpha(u, v) = (A_u, v) \quad \forall v \in H.$$

There are some properties of A_u which we will use later. First, as α is continuous, there exists C greater than 0 such that

$$|\alpha(u, v)| \leq C\|u\|_H\|v\|_H \quad \forall u, v \in H.$$

Then $\forall u \in H$, we choose $v = A_u$ so we have

$$\alpha(u, v) = \alpha(u, A_u) = (A_u, A_u) = (\|A_u\|_H)^2 \leq C\|u\|_H\|A_u\|_H,$$

which means

$$\|A_u\|_H \leq C\|u\|_H \quad \forall u \in H. \tag{15}$$

Moreover, by coercivity of α it is easy to see there exists $a > 0$ such that

$$(A_u, u) \geq a\|u\|_H{}^2 \quad \forall u \in H. \tag{16}$$

Now, what we want to prove is equivalent to finding some $u \in K$ such that

$$(A_u, v) = (g, v) \quad \forall v \in K.$$

Select a constant $\zeta > 0$ to be determined later. It suffices to show that

$$(\zeta g - \zeta A_u + u - u, v) = 0 \quad \forall v \in K. \tag{17}$$

Let θ be a linear operator on K defined by $\theta(x) = P_K(\zeta g - \zeta A_x + x)$. We need to find an $x \in K$ such that $\theta(x) = x$. By Theorem 5.13, it suffices to find a ζ which makes θ a strict contraction, or equivalently, find a ζ such that

$$d(\theta(u_1), \theta(u_2)) \leq kd(u_1, u_2), \quad \forall u_1, u_2 \in K \quad and \quad k < 1.$$

As P_K does not increase distance (by Theorem 5.7). We have

$$|\theta(x) - \theta(y)| \leq |\theta(A_y - A_x) + (x - y)|,$$

and

$$|\theta(x)-\theta(y)|^2 \leq |x-y|^2+\zeta^2|A_x-A_y|^2-2\zeta(A_x-A_y, x-y) \leq |x-y|^2(C^2\zeta^2-2a\zeta+1).$$

As $C > 0$ and $a > 0$, we know we can select a $\zeta_0 > 0$ such that $(C^2\zeta_0^2 - 2a\zeta_0 + 1) < 1$ and the corresponding function θ is a strict contraction. Setting $\zeta = \zeta_0$, we find that (17) holds.

Now, assume α is symmetric, then it is easy to check α satisfies all the axioms of a scalar product. By coercivity and continuity of α, we know the new norm acquired using the new scalar product is equivalent to the old norm, i.e. there exists a constant $C > 0$ such that

$$\frac{1}{C}\|u\|_{a(,)} \leq \|u\|_{(,)} \leq C\|u\|_{a(,)} \quad \forall u \in H.$$

Choose any $\phi \in H^*$, using Theorem 5.9, we know there exists a $g \in H$ such that $\phi(v) = \alpha(g, v) \quad \forall v \in H$. So what we want to show is equivalent to finding a $u \in K$ such that

$$\alpha(g - u, v) = 0 \quad \forall v \in K.$$

By Corollary 5.8 we know this amounts to finding the projection of g onto K under the scalar product given by α. We know such a $u \in K$ should minimize

$$\alpha(g - u, g - u)^{1/2},$$

which is equivalent to minimizing

$$\alpha(g - u, g - u) = \alpha(g, g) + \alpha(u, u) - 2\alpha(g, u).$$

As $\alpha(g, g)$ is a constant, such a $u \in K$ should be the minimizer of

$$\frac{1}{2}\alpha(x, x) - \alpha(g, x) = \frac{1}{2}\alpha(x, x) - \phi(x) \quad \forall x \in K.$$

□

6. Sobolev spaces

Definition 6.1. *Let $I = (a, b)$ be an open interval, possibly unbounded. For a function $u \in L^p(I)$, if there exists a $g \in L^p(I)$ such that*

$$\int_I u\phi' = -\int_I g\phi, \quad \forall \phi \in C_c^1(I),$$

we say u has a weak derivative, g.

Example 6.2. *Consider $I = (-1, 1)$, and a function $f : I \to \mathbb{R}$ defined by*

$$f(x) = \begin{cases} -x, & x > 0 \\ x, & x \leq 0 \end{cases}. \tag{18}$$

f does not have a classical derivative at 0*. However, it is easy to verify that $g : I \to \mathbb{R}$ defined by*

$$g(x) = \begin{cases} -1, & x > 0 \\ 1, & x \leq 0 \end{cases} \tag{19}$$

is the weak derivative of f.

Remark 6.3. *It is easy to see that if u given above has two weak derivatives, they differ only by a set of measure 0. In this sense, the weak derivative is unique.*

Definition 6.4. *Let $I = (a, b)$ be an open interval, possibly unbounded and let p lie in the interval $[1, +\infty]$. The Sobolev space $W^{1,p}(I)$ is defined to be*

$$W^{1,p}(I) = \{u \in L^p;\ \exists g \in L^p \text{ such that } \int_I u\phi' = -\int_I g\phi \quad \forall \phi \in C_c^1(I)\}.$$

We set

$$H^1(I) = W^{1,2}(I).$$

For $u \in W^{1,p}(I)$, we denote $u' = g$.

Sometimes we are unable to find a strong solution (See introduction for more information) to some partial differential equations. Instead, we seek a weak solution (See Definition 6.7). In order to do so, Sobolev spaces are introduced.

Remark 6.5. *The space $W^{1,p}$ is a Banach space equipped with the norm*

$$\|u\|_{W^{1,p}} = \|u\|_{L^p} + \|u'\|_{L^p}.$$

Or for $1 < p < +\infty$, the equivalent norm

$$\|u\|_{W^{1,p}} = (\|u\|_{L^p} + \|u'\|_{L^p})^{\frac{1}{p}}.$$

The space H^1 is a Hilbert space equipped with the scalar product

$$(u,v)_{H^1} = (u,v)_{L^2} + (u',v')_{L^2}.$$

Definition 6.6. *For any simply connected and open set $\Omega \subset \mathbb{R}$, the space $W_0^{1,p}(\Omega)$ is defined to be the collection of functions in $W^{1,p}(\Omega)$ which have zero boundary values.*

Definition 6.7. *Given a differential equation of the form:*

$$-\Delta u = f, \quad in \quad \Omega \tag{20}$$

$$u = 0, \quad on \quad \partial\Omega \tag{21}$$

for simply connected and open set $\Omega \subset \mathbb{R}$ and $f \in L^2(\Omega)$, we define the weak solution u of the system to be a function $u \in H_0^1(\Omega)$ such that

$$\int_\Omega \nabla u \nabla \phi = \int_\Omega f\phi, \quad \forall \phi \in C_c^\infty(\Omega).$$

Theorem 6.8 (Poincaré inequality)**.** *Suppose that $1 \leq p < +\infty$, and Ω is a bounded open set. There exists a constant C such that*

$$\|u\|_{L^p(\Omega)} \leq C\|\nabla u\|_{L^p(\Omega)} \quad \forall u \in W_0^{1,p}(\Omega).$$

Proof. The Poincaré inequality is really important. See [1, chapter 9.4] for more information of the proof. □

The assumption that $u \in W_0^{1,p}$ is crucial. The inequality is not true for general u in $W^{1,p}$.

The last theorem in this paper shows the existence of weak solution of Poisson's equation using all the machinery we have built up so far.

Theorem 6.9. *For any simply connected open set Ω in $\mathbb{R}$ and for any f in $L^2(\Omega)$, there exists a weak solution to the system (20)(21).*

Proof. Define a bilinear form $a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ by

$$a(u,v) = \int_\Omega \nabla u \nabla v.$$

It is clear that a is a symmetric bilinear form. Moreover,

$$|a(u,v)| = |\int_\Omega \nabla u \nabla v| \leq \left(\int_\Omega (\nabla u)^2\right)^{1/2} \left(\int_\Omega (\nabla u)^2\right)^{1/2} = \|\nabla u\|_{L^2}\|\nabla v\|_{L^2} \leq \|u\|_{H_0^1}\|v\|_{H_0^1},$$

so a is continuous.

In addition, by Theorem 6.8, we know there exists a constant C such that

$$\|u\|_{L^2(\Omega)} \leq C\|\nabla u\|_{L^2(\Omega)} \quad \forall u \in H_0^1(\Omega).$$

We choose a proper constant $C_0 \in \mathbb{R}, C_0 > 0$ such that

$$C_0 + C_0 C^2 \leq 1.$$

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

$$C_0(\|v\|_{H_0^1(\Omega)})^2 = C_0\left((\|v\|_{L^2(\Omega)})^2 + (\|\nabla v\|_{L^2(\Omega)})^2\right) \leq C_0C^2(\|\nabla v\|_{L^2(\Omega)})^2 + C_0(\|\nabla v\|_{L^2(\Omega)})^2$$

$$= (C_0 + C_0C^2)(\|\nabla v\|_{L^2(\Omega)})^2 \leq (\|\nabla v\|_{L^2(\Omega)})^2 = \int_\Omega (\nabla v)^2 = |a(v,v)|.$$

So a is coercive. Given the function $f \in L^2(\Omega)$, we define a linear map $\phi : H_0^1 \to \mathbb{R}$ by

$$\phi : u \to \int_\Omega fu.$$

ϕ is also continuous because

$$\left|\int_\Omega fu\right| \leq \left(\int_\Omega f^2\right)^{1/2}\left(\int_\Omega u^2\right)^{1/2} = \|f\|_{L^2}\|u\|_{L^2} \leq \|f\|_{L^2}\|u\|_{H_0^1},$$

where $\|f\|_{L^2}$ is less than infinity. By Theorem 5.14, there exists a unique element $u \in H_0^1(\Omega)$ such that

$$a(u,v) = \phi(v) \quad \forall v \in H_0^1(\Omega),$$

or equivalently

$$\int_\Omega \nabla u \nabla v = \int_\Omega fv \quad \forall v \in H_0^1(\Omega).$$

So we are done. □

Acknowledgements. I would like to thank my mentor, Max Engelstein, for his continuous guidance and support throughout this summer. Without his suggestions this paper would not have been possible. I would also like to thank Prof. May and the math department for organizing such an amazing program.

References

[1] Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Piscataway, NJ, 2010.
[2] Richard Bass. *Real analysis for graduate students.* Storrs, CT, 2016.
[3] Walter Rudin. *Principles of Mathematical Analysis.* Madison, WI, 1976.
[4] James Munkres. *Topology.* Cambridge, MA, 1975.