3 Banach, Hilbert and Sobolev Spaces

We intend to approach the question of existence of solutions to Dirichlet problems through theory associated with Sobolev Spaces. Therefore, the first part of this section looks at Banach and Hilbert spaces from a functional analytic point of view, The Riesz Representation Theorem and further, the Lax-Milgram Theorem. The second part of this section focuses on Sobolev spaces and introduces the concepts of *weak* derivative and *weak* solution to equations akin to (3).

3.1 Banach and Hilbert Spaces

Let $\mathcal{V}$ be a linear space over $\mathbb{R}$. A *norm* $||\cdot||$ on $\mathcal{V}$ is a mapping $||\cdot|| : \mathcal{V} \to \mathbb{R}$ satisfying

$$\begin{aligned} &(i) && ||x|| \geq 0 \text{ for all } x \in \mathcal{V}, ||x|| = 0 \text{ if and only if } x = 0; \\ &(ii) && ||\alpha x|| = |\alpha| \; ||x|| \text{ for all } \alpha \in \mathbb{R}, x \in \mathcal{V}; \\ &(iii) && ||x+y|| \leq ||x|| + ||y|| \text{ for all } x, y \in \mathcal{V}. \end{aligned} \tag{9}$$

Definiton 3.8. *A Banach space $\mathscr{B}$ is a normed linear space complete with respect to a norm $||\cdot||_{\mathscr{B}}$. That is, every sequence $\{x_n\}_{n=1}^{\infty} \subset \mathscr{B}$ Cauchy with respect to the norm $||\cdot||_{\mathscr{B}}$ converges to an element $y \in \mathscr{B}$.*

An inner product on a linear space $\mathcal{V}$ is a mapping $\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ that has positivity, linearity, and symmetry. That is, it satisfies

$$\begin{aligned} &(i) && \langle x, x \rangle \geq 0 \text{ for all } x \in \mathcal{V}, \langle x, x \rangle = 0 \text{ if and only if } x = 0; \\ &(ii) && \langle \lambda_1 x_1 + \lambda_2 x_2, y \rangle = \lambda_1 \langle x_1, y \rangle + \lambda_2 \langle x_2, y \rangle \\ & && \text{for all } \lambda_1, \lambda_2 \in \mathbb{R}, x_1, x_2, y \in \mathcal{V}; \\ &(iii) && \langle x, y \rangle = \langle y, x \rangle \text{ for all } x, y \in \mathcal{V}. \end{aligned} \tag{10}$$

Writing $||x|| = \langle x, x \rangle^{1/2}$ for $x \in \mathcal{V}$, each of the following holds for all $x, y \in \mathcal{V}$.

Proposition 3.9.

Cauchy-Schwarz inequality	$|\langle x, y \rangle| \leq ||x|| ||y||;$	(11)
Triangle inequality	$||x+y|| \leq ||x|| + ||y||;$	(12)
Parallelogram law	$||x+y||^2 + ||x-y||^2 = 2(||x||^2 + ||y||^2).$	(13)

Refer to appendix for proofs.

Definiton 3.10. *A Hilbert space $\mathscr{H}$ is a Banach space with norm $||\cdot||_{\mathscr{H}}$ endowed with an inner product $\langle\cdot,\cdot\rangle : \mathscr{H} \times \mathscr{H} \to \mathbb{R}$ compatible with its norm. That is, for any element $x \in \mathscr{H}$, $||x||_{\mathscr{H}} = \langle x, x\rangle^{1/2}$.*

It is easy to see that $||x||_{\mathscr{H}}$ is a norm. Notice (i) and (ii) of (9) are satisfied immediately by properties (i) and (ii) of (10) respectively and (9 iii) is the triangle inequality (12).

The above combines the beginning of chapter 5 with section 5.5 from [GT]. In order to study the partial differential equations we are interested in, we need a theory about linear maps on Banach and Hilbert spaces. The following three Lemmas are built from the beginning of chapter 3 of [S] and also incorporates details from [R]. Let $\mathscr{B}$ and $\mathscr{D}$ be Banach spaces. A mapping $\varphi : \mathscr{B} \to \mathscr{D}$ is said to be linear if and only if φ satisfies

$$\varphi(x+y) = \varphi(x) + \varphi(y) \text{ for all } x, y \in \mathscr{B} \text{ and} \tag{14}$$

$$\varphi(ax) = a\varphi(x) \text{ for all } a \in \mathbb{R} \text{ and } x \in \mathscr{B}. \tag{15}$$

Note that by (14), $\varphi(0)$ must equal $0 \in \mathscr{D}$. A description of continuity of linear maps is necessary to apply the theorems to follow. We now give the definition of continuity in a general Banach space setting, where the continuity of a generic function is given in terms of the Banach space norm, which is then focused to the case of linear maps.

Definiton 3.11. *Let $\mathscr{B}, \mathscr{D}$ be Banach spaces with norms $||\cdot||_{\mathscr{B}}$ and $||\cdot||_{\mathscr{D}}$. A function $\psi : \mathscr{B} \to \mathscr{D}$ is said to be continuous at $x_0 \in \mathscr{B}$ if and only if for each $\epsilon > 0$ there is a $\delta > 0$ so that if $||x - x_0||_{\mathscr{B}} < \delta$ then $||\psi(x) - \psi(x_0)||_{\mathscr{D}} < \epsilon$. If ψ is continuous at each $x_0 \in \mathscr{B}$, we say that ψ is continuous.*

Notice that if $\mathscr{B} = \mathscr{D} = \mathbb{R}$, then this definition is identical to the usual concept of continuity of real functions of one variable. That is, $f : \mathbb{R} \to \mathbb{R}$ is continuous at $x_0 \in \mathbb{R}$ if and only if for every $\epsilon > 0$ there is a $\delta > 0$ such that if $|x - x_0| < \delta$ then $|f(x) - f(x_0)| < \epsilon$.

That a linear mapping from $\mathbb{R}$ into $\mathbb{R}$ is continuous if it is continuous at $x_0 = 0$ is a well known result. The same property holds for linear maps defined on Banach spaces as shown in the following lemma.

Lemma 3.12. *Let $\mathscr{B}$ and $\mathscr{D}$ be Banach spaces and $\varphi : \mathscr{B} \to \mathscr{D}$ be a linear mapping. If φ is continuous at $x = 0$ then φ is continuous at each $x \in \mathscr{B}$.*

Proof:
Let $\epsilon > 0$ be given. Assume that $\varphi : \mathscr{B} \to \mathscr{D}$ is continuous at $0 \in \mathscr{B}$, then there is a $\delta > 0$ so that if $||x - 0||_{\mathscr{B}} < \delta$ then $||\varphi(x) - \varphi(0)||_{\mathscr{D}} < \epsilon$. Now, let $x_0, x \in \mathscr{B}$ be such that $||x - x_0||_{\mathscr{B}} < \delta$. Then $||\varphi(x - x_0) - \varphi(0)||_{\mathscr{D}} < \epsilon$, but by linearity this means $||\varphi(x) - \varphi(x_0)||_{\mathscr{D}} < \epsilon$. That is, φ is continuous at x_0. Since x_0 was arbitrary, φ is continuous.

★

Using this lemma, the continuity of a linear map can be determined by checking continuity at $x = 0$. The next lemma gives a necessary and sufficient condition for continuity at the origin for linear maps defined on Banach spaces.

Lemma 3.13. *Let $\mathscr{B}, \mathscr{D}$ be Banach spaces and $\varphi : \mathscr{B} \to \mathscr{D}$ a linear map. Then, φ is continuous if and only if φ is bounded, that is, if and only if there is a $\kappa > 0$ so that*

$$||\varphi(x)||_{\mathscr{D}} \leq \kappa ||x||_{\mathscr{B}} \text{ for all } x \in \mathscr{B}. \tag{16}$$

Proof: Assume that φ is bounded. Let $\epsilon > 0$ be given and set $\delta = \dfrac{\epsilon}{\kappa}$. If $x \in \mathscr{B}$ satisfies $||x||_{\mathscr{B}} < \delta$, then

$$||\varphi(x)||_{\mathscr{D}} \leq \kappa ||x||_{\mathscr{B}} < \kappa \frac{\epsilon}{\kappa} = \epsilon. \tag{17}$$

Since $\epsilon > 0$ was arbitrary, φ is continuous at $x = 0$ and therefore φ is a continuous mapping by Lemma 3.12.

Now assume that φ is continuous. Then, in particular, φ is continuous at $x = 0$. Thus, there is a $\delta > 0$ so that $||\varphi(x)||_{\mathscr{D}} \leq 1$ for $||x||_{\mathscr{B}} < \delta$. Next, let $y \neq 0 \in \mathscr{B}$. Then, $z = \delta \dfrac{y}{||y||_{\mathscr{B}}}$ satisfies $||z||_{\mathscr{B}} < \delta$ and from the linearity of φ we can see that

$$\begin{aligned} ||\varphi(y)||_{\mathscr{D}} &= \left\| \frac{||y||_{\mathscr{B}}}{\delta} \frac{\delta}{||y||_{\mathscr{B}}} \varphi(y) \right\|_{\mathscr{D}} \\ &= \frac{||y||_{\mathscr{B}}}{\delta} \left\| \varphi\Big(\frac{\delta y}{||y||_{\mathscr{B}}} \Big) \right\|_{\mathscr{D}} \\ &= \frac{||y||_{\mathscr{B}}}{\delta} ||\varphi(z)||_{\mathscr{D}} \leq \frac{||y||_{\mathscr{B}}}{\delta}. \end{aligned} \tag{18}$$

Therefore, with $\kappa = \dfrac{1}{\delta}$ we have

$$||\varphi(y)||_{\mathscr{D}} \leq \kappa ||y||_{\mathscr{B}}. \tag{19}$$

Since κ is independent of y and $y \neq 0 \in \mathscr{B}$ was arbitrary, we have that (19) holds for every $y \neq 0 \in \mathscr{B}$. Finally, to see that (19) holds for $y = 0$, recall that $\varphi(0) = 0 \in \mathscr{B}$. That is, since φ is linear, for any $x \in \mathscr{B}$ we have $\varphi(x) = \varphi(x) + \varphi(0)$. Thus, $\varphi(0) = 0$ and (19) holds trivially. Therefore, (19) holds for all $y \in \mathscr{B}$ and condition (17) is satisfied.

★

Next we give a technical result about the range of a linear map φ, $Rng(\varphi)$, that we will need in the proof of the Lax-Milgram theorem at the end of this section.

Lemma 3.14. *Let $\mathscr{B}$ be a Banach space and $\varphi : \mathscr{B} \to \mathscr{B}$ be a continuous linear mapping. Suppose that there is a $\nu > 0$ so that φ satisfies*

$$\nu ||x||_{\mathscr{B}} \leq ||\varphi(x)||_{\mathscr{D}} \tag{20}$$

for all $x \in \mathscr{B}$. Then, $Rng(\varphi)$ is a closed set.

Proof Suppose $Rng(\varphi)$ is not closed. Then, there is a $z \in \mathscr{B}$ that is a limit point of $Rng(\varphi)$ not lying in $Rng(\varphi)$. That is, $z \neq \varphi(x)$ for any $x \in \mathscr{B}$. As z is a limit point, choose a sequence $\{y_i\} \subset Rng(\varphi)$ converging to z in $\mathscr{D}$. Now, since $\{y_i\} \subset Rng(\varphi)$ we have that $y_i = \varphi(x_i)$ for some $x_i \in \mathscr{B}$. Since $\{y_i\}$ converges to z in $\mathscr{D}$, it is Cauchy in $\mathscr{D}$ and (20) gives us that the sequence $\{x_i\}$ is also Cauchy in $\mathscr{B}$. As $\mathscr{B}$ is complete, $\{x_i\}$ converges to an element $x \in \mathscr{B}$. Now, let $\epsilon > 0$ be given. Then,

$$\begin{aligned} ||\varphi(x) - z||_{\mathscr{D}} &= ||\varphi(x) - \varphi(x_i) + \varphi(x_i) - z||_{\mathscr{D}} \\ &\leq ||\varphi(x - x_i)||_{\mathscr{D}} + ||y_i - z||_{\mathscr{D}} \\ &\leq \kappa ||x - x_i||_{\mathscr{D}} + ||y_i - z||_{\mathscr{D}} \text{ by the continuity of } \varphi \\ &< \epsilon \end{aligned}$$

for i sufficiently large. Thus, $||\varphi(x) - z||_{\mathscr{D}} = 0$, that is, $\varphi(x) = z$, must be true. Therefore we conclude that $z \in Rng(\varphi)$, a contradiction.

★

Linear maps $\varphi : \mathscr{B} \to \mathbb{R}$ are important in the context of Hilbert spaces as they are compatible with the inner product. For example, consider the Banach space

$\mathbb{R}$ with inner product defined by $\langle x, y\rangle = xy$. It is clear that with this inner product, $\mathbb{R}$ is a Hilbert space since

$$||x||_{\mathbb{R}} = |x| = (x^2)^{1/2} = \langle x, x\rangle^{1/2}.$$

Linear maps on $\mathbb{R}$ are of the form $y = mx$ where m is a real constant. To see that these mappings are given by the inner product, consider the following. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be linear. Then, $\varphi(x) = mx$ for a constant m. Setting $y = m$ we have that $\varphi(x) = \langle x, y\rangle$ for all $x \in \mathbb{R}$. That is, there is a unique $y \in \mathbb{R}$ so that $\varphi(x) = \langle x, y\rangle$ for all $x \in \mathbb{R}$. This example is a special case of the Riesz Representation Theorem below. Before stating the result, we introduce the notion of linear functional and dual space. The following definition and three theorems are presented in sections 5.4, 5.7, 5.6, and 5.8 of [GT] respectively. The proof of Theorem 3.17 includes detail from section 2.1 of [S].

Definiton 3.15. *Let $\mathscr{B}$ be a Banach space. A linear mapping $\varphi : \mathscr{B} \to \mathbb{R}$ is called a* linear functional. *The collection of all continuous linear functionals defined on $\mathscr{B}$ is called the* dual space *of $\mathscr{B}$ and is denoted by $\mathscr{B}^*$.*

Theorem 3.16. *(The Riesz Representation Theorem) Let $\mathscr{H}$ be a Hilbert space with norm $||\cdot||_{\mathscr{H}}$ and inner product $\langle\cdot,\cdot\rangle$. Then, for every bounded linear functional φ on $\mathscr{H}$, there is a unique $y \in \mathscr{H}$ so that $\varphi(x) = \langle x, y\rangle$ for all $x \in \mathscr{H}$, that is, each linear functional $\varphi \in \mathscr{H}^*$ is given by the inner product.*

Before proving this result we require a technical theorem known as the projection theorem. This theorem specifies that given a closed subspace $\mathscr{M} \subset \mathscr{H}$ we may write any $x \in \mathscr{H}$ as $x = z + w$ where $z \in \mathscr{M}$ and $w \in \mathscr{M}^\perp$, where $\mathscr{M}^\perp$ denotes the orthogonal compliment of $\mathscr{M}$ in $\mathscr{H}$. In the context of $\mathbb{R}^2$, we may think of $\mathscr{M} = span\{(1,0)\}$ and $\mathscr{M}^\perp = span\{(0,1)\}$. Then, any $(x, y) \in \mathbb{R}^2$ may be written as $(x, y) = (x, 0) + (0, y) = z + w$ where $z \in \mathscr{M}$ and $w \in \mathscr{M}^\perp$.

Theorem 3.17. *(The Projection Theorem) Let $\mathscr{M}$ be a closed subspace of a Hilbert Space $\mathscr{H}$. Then for $x \in \mathscr{H}$, $x = y + z$ where $y \in \mathscr{M}$ and $z \in \mathscr{M}^\perp$.*

Proof:
Let $\mathscr{M}$ be a closed subspace of $\mathscr{H}$. If $x \in \mathscr{M}$, set $y = x$ and $z = 0$. Thus, we may assume $\mathscr{M} \neq \mathscr{H}$ and that $x \notin \mathscr{M}$. Set

$$d = dist(x, \mathscr{M}) = \inf_{y\in\mathscr{M}} ||x - y||_{\mathscr{H}} > 0.$$

By the definition of d, there is a sequence $\{y_n\}$ of elements of $\mathscr{M}$ such that $||x - y_n||_{\mathscr{H}} \to d$. Apply the parallelogram law to $x - y_n$ and $x - y_m$ to obtain

$$||(x - y_n) + (x - y_m)||^2_{\mathscr{H}} + ||(x - y_n) - (x - y_m)||^2_{\mathscr{H}} = 2||x - y_n||^2_{\mathscr{H}} + 2||x - y_m||^2_{\mathscr{H}}$$

or equivalently,

$$4\left\|x - \left[\frac{(y_n + y_m)}{2}\right]\right\|^2_{\mathscr{H}} + ||y_m - y_n||^2_{\mathscr{H}} = 2||x - y_n||^2_{\mathscr{H}} + 2||x - y_m||^2_{\mathscr{H}}.$$

Since $\mathscr{M}$ is a subspace, $\dfrac{(y_n + y_m)}{2}$ is in $\mathscr{M}$. Hence the left hand side is not smaller than

$$4d^2 + ||y_m - y_n||^2_{\mathscr{H}}.$$

This implies

$$||y_m - y_n||^2_{\mathscr{H}} \leq 2||x - y_n||^2_{\mathscr{H}} + 2||x - y_m||^2_{\mathscr{H}} - 4d^2.$$

As $m, n \to \infty$, the right-hand side approaches $2d^2 + 2d^2 - 4d^2 = 0$. Therefore

$$||y_m - y_n||^2_{\mathscr{H}} \to 0 \text{ as } m, n \to \infty$$

Thus, $\{y_n\}$ is Cauchy in $\mathscr{H}$. Using that $\mathscr{H}$ is complete, there is a $y \in \mathscr{H}$ such that the sequence $\{y_n\}$ converges to y. As $\mathscr{M}$ is closed in $\mathscr{H}$, $y \in \mathscr{M}$, and $d = \lim_{n\to\infty} ||x - y_n||_{\mathscr{H}} = ||x - y||_{\mathscr{H}}$.

Now write $x = y + z$ where $z = x - y$. All that remains is to show that $z \in \mathscr{M}^{\perp}$, in other words, that $\langle t, z\rangle = 0$ for all $t \in \mathscr{M}$. For any $t \in \mathscr{M}$ and $\alpha \in \mathbb{R}$, $y + \alpha t \in \mathscr{M}$ and so

$$\begin{aligned} d^2 &\leq ||x - y - \alpha t||^2_{\mathscr{H}} \\ &= \langle z - \alpha t, z - \alpha t\rangle \\ &= ||z||^2_{\mathscr{H}} - 2\alpha\langle t, z\rangle + \alpha^2||t||^2_{\mathscr{H}} \end{aligned} \tag{21}$$

Since $\|z\|^2 = d^2$, (21) simplifies to give

$$\begin{aligned} \langle t, z\rangle &\leq \frac{\alpha}{2}||t||^2_{\mathscr{H}} \text{ for } \alpha > 0, \\ \langle t, z\rangle &\geq \frac{\alpha}{2}||t||^2_{\mathscr{H}} \text{ for } \alpha < 0, \end{aligned}$$

that is,

$$-\frac{|\alpha|}{2}||t||^2_{\mathscr{H}} \leq \langle t, z\rangle \leq \frac{|\alpha|}{2}||t||^2_{\mathscr{H}}. \tag{22}$$

Since (22) holds for all $\alpha \in \mathbb{R}$, it must be that $\langle t, z\rangle = 0$ for all $t \in \mathscr{M}$, which can only happen if z is orthogonal to t, that is $z \in \mathscr{M}^\perp$.

★

Proof of Theorem 3.16: Let $\mathscr{N}$ denote the null space of φ, that is $\mathscr{N} = \{x \mid \varphi(x) = 0\}$. If $\mathscr{N} = \mathscr{H}$, the result is proved by taking $y = 0$. Otherwise, since $\mathscr{N}$ is a closed subspace of $\mathscr{H}$, there exists by Theorem 3.17 an element $z \neq 0 \in \mathscr{N}^\perp \subset \mathscr{H}$ such that $\langle x, z\rangle = 0$ for all $x \in \mathscr{N}$. Hence $\varphi(z) \neq 0$ and, moreover, for any $x \in \mathscr{H}$,

$$\varphi\Big(x - \frac{\varphi(x)}{\varphi(z)}z\Big) = \varphi(x) - \frac{\varphi(x)}{\varphi(z)}\varphi(z) = 0$$

So that the element $x - \frac{\varphi(x)}{\varphi(z)}z \in \mathscr{N}$. This means

$$\Big\langle x - \frac{\varphi(x)}{\varphi(z)}z, z\Big\rangle = 0$$

that is,

$$\langle x, z\rangle = \frac{\varphi(x)}{\varphi(z)}||z||^2_{\mathscr{H}}$$

and hence $\varphi(x) = \langle x, y\rangle$ where $y = \frac{\varphi(z)}{||z||^2_{\mathscr{H}}}z$.
To see that y is unique, assume the contrary. Let z_1 be an element satisfying $\varphi(x) = \langle x, z_1\rangle$ for all $x \in \mathscr{H}$. Then

$$\varphi(x) - \varphi(x) = \langle x, z_1 - y\rangle = 0 \qquad \text{for all } x \in \mathscr{H}.$$

Choosing $x = z_1 - y$ shows that $||z_1 - y||_{\mathscr{H}} = 0$, and so $z_1 = y$.

★

Now that we have the Riesz Representation Theorem, it can be extended to bilinear forms. This generalization is known as the Lax-Milgram Theorem, the main tool used to prove the final theorem of this paper. The following proof elaborates on the proof of the Lax-Milgram Theorem given in [GT] (page 83) and I thank Dr. Rodney for related notes. Recall that a real Bilinear form $L(x, y)$ on a Banach space $\mathscr{B}$ is a function $L : \mathscr{B} \times \mathscr{B} \to \mathbb{R}$ that is linear in both variables separately. That is,

$$\begin{aligned} (i) \qquad & L(ax + by, z) = aL(x, z) + bL(y, z) \text{ and} \\ (ii) \qquad & L(x, aw + bz) = aL(x, w) + bL(x, z). \end{aligned}$$

Theorem 3.18. *(The Lax-Milgram Theorem) Let $\mathscr{H}$ be a Hilbert space and $L(x, y)$ a real bilinear form defined on $\mathscr{H}$ that satisfies the two following properties.*

(i) L *is* bounded *on $\mathscr{H}$. That is, there is a $\kappa > 0$ so that $|L(x, y)| \le \kappa ||x||_{\mathscr{H}} ||y||_{\mathscr{H}}$ for all $x, y \in \mathscr{H}$.*

(ii) L *is* coercive *on $\mathscr{H}$. That is, there is a $\nu > 0$ so that $L(x, x) \ge \nu ||x||^2_{\mathscr{H}}$ for all $x \in \mathscr{H}$.*

Then, given $\varphi \in \mathscr{H}^*$ there is a unique $z \in \mathscr{H}$ so that $\varphi(x) = L(x, z)$ for all $x \in \mathscr{H}$.

Proof: Let $\mathscr{H}$ be a Hilbert space and $L(x, y)$ a real bilinear form defined on $\mathscr{H}$ that is both bounded and coercive. Fix $y \in \mathscr{H}$. Then, since (i) holds, the mapping $L(x, y) = \Lambda(x)$ is a bounded linear functional on $\mathscr{H}$, that is, $\Lambda(x)$ is a continuous linear functional on $\mathscr{H}$. By the Riesz Representation Theorem (Theorem 3.16) there is a unique element $z_y \in \mathscr{H}$ so that $L(x, y) = \Lambda(x) = \langle x, z_y \rangle$ for all $x \in \mathscr{H}$. Since z_y is uniquely determined by y, we can define a mapping $T : \mathscr{H} \to \mathscr{H}$ by setting $T(y) = z_y$, which gives $L(x, y) = \langle x, T(y) \rangle$. This mapping T is linear. To see this, let $v, w \in \mathscr{H}$ and $a, b \in \mathbb{R}$. Then $s = av + bw \in \mathscr{H}$ and $T(s) = z_s = z_{av+bw}$. But, since L is linear in the second variable we have $L(x, av + bw) = aL(x, v) + bL(x, w) = a\langle x, z_v \rangle + b\langle x, z_w \rangle = \langle x, az_v + bz_w \rangle$. This means that $T(s) = aT(v) + bT(w)$ giving that T is linear. As a linear mapping, T is bounded. To see this notice that as $T(y) \in \mathscr{H}$ for $y \in \mathscr{H}$ we have

$$||T(y)||^2_{\mathscr{H}} = |\langle T(y), T(y) \rangle| = |L(T(y), y)| \le \kappa ||T(y)||_{\mathscr{H}} ||y||_{\mathscr{H}} \tag{23}$$

by (i), and so dividing by $||T(y)||_{\mathscr{H}}$ gives $||T(y)||_{\mathscr{H}} \leq \kappa||y||_{\mathscr{H}}$. Also, by (ii) we have

$$\nu||x||^2_{\mathscr{H}} \leq L(x,x) = \langle x, T(x)\rangle \leq ||x||_{\mathscr{H}}||T(x)||_{\mathscr{H}} \tag{24}$$

giving that $\nu||x||_{\mathscr{H}} \leq ||T(x)||_{\mathscr{H}}$. These estimates combined with Lemma 3.14 give that T is a continuous linear map with closed range. Notice, as well, T is both one to one and onto and therefore invertible. T is one to one by definition, since $T(y) = z_y$ and z_y was uniquely determined by y. To see that T is onto, notice that, since $Rng(T)$ is closed, if T is not onto then we can apply the Projection Theorem (Theorem 3.17) to find a non-zero $z \in \mathscr{H}$ so that $\langle z, T(x)\rangle = 0$ for all $x \in \mathscr{H}$. Choosing $x = z$ gives $\langle z, T(z)\rangle = L(z,z) = 0$. This means that $z = 0$ by (ii), the coercive property of L, a contradiction.

And so we have that T is a continuous linear map that is both one-to-one and onto. Therefore, T^{-1} exists and is also continuous, one-to-one and onto. Finally, let $\varphi \in \mathscr{H}^*$ be a continuous linear functional. Then, by the Riesz Representation Theorem, there is a unique $g \in \mathscr{H}$ so that $\varphi(x) = \langle x, g\rangle$ for all $x \in \mathscr{H}$. So, applying T^{-1} we obtain

$$\varphi(x) = \langle x, g\rangle = L(x, T^{-1}(g))$$

for all $x \in \mathscr{H}$. Thus for $\varphi \in \mathscr{H}^*$ we have a unique $z = T^{-1}(g) \in \mathscr{H}$ such that $\varphi(x) = L(x,z)$ for all $x \in \mathscr{H}$.

★

3.2 Sobolev Spaces and Weak Solutions

In this section we apply the Banach and Hilbert space theory to the Sobolev space we are interested in working with. To define this Sobolev space we require the definition of the classical Banach spaces of analysis.

For $1 \leq p < \infty$ we let $L^p(\Omega)$ denote the Banach space consisting of functions on Ω whose p^{th} power has finite integral. That is,

$$L^p(\Omega) = \{f : \Omega \to \mathbb{R} \mid \Big(\int_\Omega |f|^p dx\Big)^{1/p} < \infty\}. \tag{25}$$

The norm on $L^p(\Omega)$ is defined by

$$||u||_p = ||u||_{L^p(\Omega)} = \Big(\int_\Omega |u|^p dx\Big)^{1/p}. \tag{26}$$

The above definition and norm are given in section 7.1 of [GT], however, that $L^p(\Omega)$ is a Banach space with norm (26) for $1 \leq p < \infty$ is a deep mathematical result known as the Riesz-Fischer Theorem that will be taken as fact here. One may find the statement and proof of the Riesz-Fischer theorem in [P] p. 277.

Before defining the Sobolev space to be considered, we begin by examining the space $L^2(\Omega)$ of square integrable functions.

Let $u, v \in L^2(\Omega)$, then we define their inner product and norm by setting

$$< u, v >_2 = \int_\Omega u(x)v(x)dx. \tag{27}$$

That (27) is an inner product is easily verified as the three properties of (10) are satisfied.

$$(i) \qquad < u, u >_2 = \int_\Omega u(x)u(x)dx = \int_\Omega u(x)^2 dx > 0, \qquad \text{since } u^2(x) > 0.$$

$$\begin{aligned}(ii) \qquad < \lambda_1 u_1 + \lambda_2 u_2, v >_2 &= \int_\Omega [\lambda_1 u_1(x) + \lambda_2 u_2(x)]v(x)dx \\ &= \int_\Omega [\lambda_1 u_1(x)v(x) + \lambda_2 u_2(x)v(x)]dx \\ &= \lambda_1 \int_\Omega u_1(x)v(x)dx + \lambda_2 \int_\Omega u_2(x)v(x)dx \\ &= \lambda_1 < u_1, v >_2 + \lambda_2 < u_2, v >_2 .\end{aligned}$$

$$(iii) \qquad < u, v >_2 = \int_\Omega u(x)v(x)dx = \int_\Omega v(x)u(x)dx = < v, u >_2 .$$

Using this, we see that the $L^2(\Omega)$ norm given by (26) is compatible with (27) since

$$\|u\|_{L^2(\Omega)} = \Big(\int_\Omega |u(x)|^2 dx\Big)^{1/2} = \Big(\int_\Omega u(x)u(x)dx\Big)^{1/2} = < u, u >^{1/2} . \tag{28}$$

Therefore, the space $L^2(\Omega)$ is a complete normed linear space with a compatible inner product, in other words, a Hilbert Space. With (27) and (28), we have a special form of the Cauchy-Schwarz inequality, (11),

$$
\begin{aligned}
| < u, v >_2 | = | \int_\Omega u(x)v(x)dx| &\leq \Big(\int_\Omega |u(x)|^2 dx \Big)^{1/2} \Big(\int_\Omega |v(x)|^2 dx \Big)^{1/2} \\
&= ||u||_2 ||v||_2
\end{aligned}
$$

for all $u, v \in L^2(\Omega)$.

Now look at the Sobolev space defined by

$$W_0^{1,2}(\Omega) = \{u \in L^2(\Omega), Du \in L^2(\Omega)\}.$$

Consider a related inner product on $C_0^1(\Omega)$ that accounts for the size of the derivative of a function.

$$< u, v >_W = \int_\Omega u(x)v(x)dx + \int_\Omega Du(x)Dv(x)dx. \tag{29}$$

Using a familiar argument, see (27), it is simple to see that (29) is indeed an inner product. This inner product defines a norm on $C_0^1(\Omega)$ by setting

$$||u||_{W^{1,2}(\Omega)} = < u, u >_W^{1/2} = \Big(\int_\Omega |u(x)|^2 dx + \int_\Omega |Du(x)|^2 dx \Big)^{1/2}. \tag{30}$$

Definiton 3.19. *Let Ω be a domain of $\mathbb{R}^n$. The Sobolev space $W_0^{1,2}(\Omega)$ is defined as the closure of $C_0^1(\Omega)$ in the norm (30).*

A more general definition for $W_0^{k,p}(\Omega)$, $k \geq 1$, is given by [GT] in section 7.5. Using our definition, a generic element of $W_0^{1,2}(\Omega)$ is an equivalence class of sequences of continuously differentiable functions Cauchy with respect to (30), defined in Ω and with compact support. Also, since both inequalities

$$
\begin{aligned}
||u||_{L^2(\Omega)} &\leq ||u||_{W^{1,2}(\Omega)} \text{ and} \qquad (31) \\
||Du||_{L^2(\Omega)} &\leq ||u||_{W^{1,2}(\Omega)} \qquad (32)
\end{aligned}
$$

are true, any sequence $\{u_j\}$ Cauchy in the $W_0^{1,2}(\Omega)$ norm is also Cauchy in $L^2(\Omega)$ and likewise the sequence $\{Du_j\}$ of vectors in $\mathbb{R}^n$ is Cauchy in $L^2(\Omega)$. Since $L^2(\Omega)$ is complete by the Riesz-Fischer theorem, there are elements $u, \vec{w} \in L^2(\Omega)$ so that $u_j \to u$ and $Du_j \to \vec{w}$ in $L^2(\Omega)$. Therefore, to each equivalence class in $W_0^{1,2}(\Omega)$ there is a unique pair $(u, \vec{w})$ with norm

$$||(u, \vec{w})||_{W^{1,2}(\Omega)} = (||u||_2^2 + ||\vec{w}||_2^2)^{1/2} = \lim_{j\to\infty} (||u_j||_2^2 + ||Du_j||_2^2)^{1/2}.$$

This definition of norm is identical to (30), and, with the inner product given by (29), $W_0^{1,2}(\Omega)$ is a Hilbert space.

We now discuss the notion of weak derivative. Let $(u, \vec{w}) \in W_0^{1,2}(\Omega)$, and select a sequence $\{u_j\} \subset C_0^1(\Omega)$ such that $||u_j - u||_2 \to 0$ and $||Du_j - \vec{w}||_2 \to 0$ as $j \to \infty$. Then, for each $j \in \mathbb{N}$, we can integrate by parts against any fixed $v \in C_0^1(\Omega)$ to obtain

$$\int_\Omega u_j(x) D_i v(x) = -\int_\Omega v(x) D_i u_j(x) dx. \tag{33}$$

Furthermore, by the Cauchy-Schwarz inequality we have

$$\begin{aligned}
\Big|\int_\Omega u_j D_i v dx - \int_\Omega u D_i v dx\Big| &= \Big|\int_\Omega (u_j - u) D_i v dx\Big| \\
&\leq \int_\Omega |u - u_j||D_i v| dx \\
&\leq \Big(\int_\Omega |u - u_j|^2 dx\Big)^{1/2}\Big(\int_\Omega |D_i v|^2 dx\Big)^{1/2} \\
&\leq ||u_j - u||_{L^2(\Omega)} ||D_i v||_{L^2(\Omega)},
\end{aligned}$$

and

$$\begin{aligned}
\Big|\int_\Omega v D_i u_j dx - \int_\Omega v\vec{w}\Big| &= \Big|\int_\Omega v(D_i u_j - \vec{w}) dx\Big| \\
&\leq \int_\Omega |v||D_i u_j - \vec{w}| dx \\
&\leq \Big(\int_\Omega |v|^2 dx\Big)^{1/2}\Big(\int_\Omega |D_i u_j - \vec{w}|^2 dx\Big)^{1/2} \\
&\leq ||D_i u_j - \vec{w}||_{L^2(\Omega)} ||v||_{L^2(\Omega)}.
\end{aligned}$$

Therefore, as $j \to \infty$ we have that the functions $u, \vec{w}$ satisfy

$$\int_\Omega u D_i v dx = -\int_\Omega v w_i dx \tag{34}$$

where $v \in C_0^1(\Omega)$ is arbitrary and w_i is the i^{th} component of $\vec{w}$. Equality (34) looks very similar to an integration by parts formula, and so the function $w_i \in L^2(\Omega)$ is thought of as the i^{th}*weak derivative* of u in Ω and the vector function $\vec{w}$ is the weak gradient of u in Ω. We will write $w_i = D_i u$ and $\vec{w} = Du$ for convenience. With this notation, we will refer to a generic element (u, Du) of $W_0^{1,2}(\Omega)$ by writing $u \in W_0^{1,2}(\Omega)$. Furthermore, we may now realize the Hilbert space $W_0^{1,2}(\Omega)$ as the collection of those weakly differentiable functions $u \in L^2(\Omega)$ whose weak derivative satisfies $Du \in L^2(\Omega)$. That is,

$$W_0^{1,2}(\Omega) = \{u \in L^2(\Omega) \mid Du \in L^2(\Omega)\}.$$

For clarity, we now give the formal definition of weak derivative.

Definiton 3.20. *Let $u \in L^2(\Omega)$. A vector valued function $\vec{w} = (w_1, ..., w_n)$ is called the weak gradient of u in Ω if and only if*

$$\int_\Omega u D_i v dx = -\int_\Omega v w_i dx$$

for all $v \in C_0^1(\Omega)$ and $i = 1, ..., n$. If such a $\vec{w}$ exists we say that u is weakly differentiable.

A much more general form of (34) can be found on page 149 of [GT]

Definition 4.21 below allows the interpretation of $W_0^{1,2}(\Omega)$ as the collection of all those $u \in L^2(\Omega)$ with weak gradient $Du \in L^2(\Omega)$ and $u = 0$ on $\partial\Omega$. This interpretation is not obvious. The interested reader may refer to [GT] for the development of this idea.