Chapter 5

Variational formulation of elliptic partial differential equations (PDEs)

5.1. Introduction

In this chapter we introduce important concepts from *functional analysis* that will provide a rigorous variational framework for partial differential equations posed on a d-dimensional domain with general boundary conditions. We will also discuss well-posedness of the variational formulation.

5.2. Lebesgue integration

In this section we provide a brief, very incomplete overview of Lebesgue integration theory that will be used to construct the notation of a "weak" derivative, i.e., a notion of differentiation that makes sense even for functions that are non-smooth or not even defined pointwise.

Lebesgue integration provides a natural and rigorous setting for functional analysis because it can be used to define *complete* spaces, i.e., spaces where limits of integrable functions are themselves integrable. The familiar Riemann integral does not possess these properties. While the Lesbegue theory provides a more general definition of integrability, conveniently, for functions that are both Lesbegue and Riemann integrable, the definitions agree. A formal definition of the Lebesgue integral relies on a measure theoretic construction, which is beyond the scope of this course. Instead we summarize the key results below and provide a few examples.

Let $\Omega \subset \mathbb{R}^n$ (open) equipped with Lebesgue measure dx. We restrict out attention to real-valued functions on Ω that are Lebesgue measurable. Measurable functions in Ω are defined *almost everywhere* (a.e.) in Ω, i.e., if we change the values of a measurable function f on a subset of Ω of measure zero, the measurable function does not change. A measure can be thought of as a generalization of distance, e.g., intuitivitely, think of length if $\Omega \subset \mathbb{R}$ or area if $\Omega \subset \mathbb{R}^2$ so a finite collection of points in $\mathbb{R}$ and a finite collection of curves in $\mathbb{R}^2$ have zero measure. Two measurable functions $f, g \in \mathcal{F}_{\Omega\to\mathbb{R}}$ are called equal almost everywhere in Ω if there exists $E \subset \Omega$ such that the Lebesgue measure of E is zero and $f(x) = g(x)$ for all $x \in \Omega\backslash E$. In this sense, $f \in \mathcal{F}_{\Omega\to\mathbb{R}}$ is the *zero function* (zero a.e.) if there exists $E \subset \Omega$ of measure zero such that $f(x) = 0$ for all $x \in \Omega\backslash E$. We denote the Lebesgue integral of f as

$$\int_\Omega f(x)\,dx. \tag{5.1}$$

For $1 \leqslant p < \infty$, let

$$\|f\|_{L^p(\Omega)} := \left(\int_\Omega |f(x)|^p\,dx\right)^{1/p} \tag{5.2}$$

and for $p = \infty$ define

$$\|f\|_{L^\infty(\Omega)} := \inf\{C \in \mathbb{R}_{>0} \text{ such that } |f(x)| \leqslant C \text{ a.e. in } \Omega\}. \tag{5.3}$$

Then we define the *Lebesgue spaces* as

$$L^p(\Omega) := \left\{ f \in \mathcal{F}_{\Omega\to\mathbb{R}} \;\middle|\; \|f\|_{L^p(\Omega)} < \infty \right\}. \tag{5.4}$$

Through a minor abuse of notation, we also use $L^p(\Omega)$ to denote the set of *equivalence classes* of functions indistinguishable with respect to $\|\cdot\|_{L^p(\Omega)}$, i.e., let $f \in L^p(\Omega)$ then the equivalence class associated with f is

$$[f] := \left\{ g \in L^p(\Omega) \;\middle|\; \|f - g\|_{L^p(\Omega)} = 0 \right\}. \tag{5.5}$$

The Lebesgue spaces are Banach (complete normed) spaces, which we establish in the next three propositions.

Proposition 5.1 (Minkowski's Inequality)**.** *For* $1 \leqslant p \leqslant \infty$ *and* $f, g \in L^p(\Omega)$*, we have*

$$\|f + g\|_{L^p(\Omega)} \leqslant \|f\|_{L^p(\Omega)} + \|g\|_{L^p(\Omega)}. \tag{5.6}$$

Proposition 5.2. *For* $1 \leqslant p \leqslant \infty$, $(L^p(\Omega), \|\cdot\|_{L^p(\Omega)})$ *is a normed space.*

Proof. We first show that $L^p(\Omega)$ is a linear subspace of the linear space $\mathcal{F}_{\Omega\to\mathbb{R}}$. Let $f, g \in L^p(\Omega)$ and $\alpha, \beta \in \mathbb{R}$, then

$$\|\alpha f + \beta g\|_{L^p(\Omega)} \leqslant \|\alpha f\|_{L^p(\Omega)} + \|\beta g\|_{L^p(\Omega)} \leqslant \alpha \|f\|_{L^p(\Omega)} + \beta \|g\|_{L^p(\Omega)} < \infty, \tag{5.7}$$

which follows from the Minkowski inequality and definition of $\|\cdot\|_{L^p(\Omega)}$. This shows $L^p(\Omega)$ is closed and therefore a linear subspace of $\mathcal{F}_{\Omega\to\mathbb{R}}$. Next we show $\|\cdot\|_{L^p(\Omega)}$ is a valid norm. The triangle inequality follows directly from the Minkowski inequality. Linearity w.r.t. scalar multiplication follows from the definition

$$\|\lambda f\|_{L^p(\Omega)} = \left(\int_\Omega |\lambda f(x)|^p \, dx \right)^{1/p} = |\lambda| \left(\int_\Omega |f(x)|^p \, dx \right)^{1/p} = |\lambda| \, \|f\|_{L^p(\Omega)} \tag{5.8}$$

for $\lambda \in \mathbb{R}$, $f \in L^p(\Omega)$ and $1 \leqslant p < \infty$ (the case for $p = \infty$ follows similarly). Non-negativity follows trivially from the definition (the integrand is non-negative so the integral must be non-negative). Positive definiteness follows from $\|f\|_{L^p(\Omega)} = \int_\Omega |f(x)|^p \, dx = 0$ and the integrand $|f(x)|^p \geqslant 0$; the only way for both of these conditions to be true is if $f = 0$ a.e. ☐

Proposition 5.3. *For* $1 \leqslant p \leqslant \infty$, $(L^p(\Omega), \|\cdot\|_{L^p(\Omega)})$ *is a Banach space.*

This is an extremely important result in the study of the variational formulation of PDEs; however, its proof is beyond the scope of the course.

5.3. Weak differentiation

Now we turn to a generalized definition of differentiation that is well-defined even for Lebesgue functions that do not necessarily satisfy traditional smoothness conditions. For the remainder of the course, we will restrict ourselves to considering Lipschitz domains that possess Lipschitz continuous boundaries.

Definition 5.3.1 (Lipschitz domain)**.** A domain $\Omega \subset \mathbb{R}^d$ is called a Lipschitz domain if its boundary is Lipschitz continuous (corners allowed, cusps are not).

Definition 5.3.2 (Weak derivatives in $\mathbb{R}$)**.** Let $\Omega \subset \mathbb{R}$ be a bounded domain. The *weak derivative* of a function $f : \Omega \to \mathbb{R}$, denoted $D^1 f$ or Df or $\frac{df}{dx}$, exists if there exists $g \in L^2(\Omega)$ such that

$$\int_\Omega v g \, dx = -\int_\Omega \frac{dv}{dx} f \, dx \qquad \forall v \in \mathcal{C}_c^\infty(\Omega), \tag{5.9}$$

in which case $D^1 f := g$. The kth weak derivative, denoted $D^k f$ or $\frac{d^k f}{dx^k}$, exists if there exists $g^{(k)} \in \text{Ł}^2(\Omega)$ such that

$$\int_\Omega v g^{(k)} \, dx = (-1)^k \int_\Omega \frac{d^k v}{dx^k} f \, dx \qquad \forall v \in \mathcal{C}_c^\infty(\Omega), \tag{5.10}$$

in which case $D^k f := g^{(k)}$, for $k = 1, \dots, d$.

Example 5.1: Weak derivative of continuous piecewise polynomial functions

In this example, we will show that piecewise polynomial continuous functions have weak derivatives that agree with our informal intuition (piecewise polynomial of lower degree). Consider $\Omega := (0, L) \subset \mathbb{R}$ partitioned into two intervals $\Omega_1 := (0, \hat{x})$ and $\Omega_2 := (\hat{x}, L)$. Let $f \in \mathcal{C}^0(\Omega)$ such that $f|_{\Omega_1} \in \mathcal{P}^k(\Omega_1)$, $f|_{\Omega_2} \in \mathcal{P}^k(\Omega_2)$ ($k \in \mathbb{N}$) and take $v \in \mathcal{C}_c^\infty(\Omega)$ (notice this implies $v(0) = v(L) = 0$). Then, following the definition of the weak derivative, we consider

$$-\int_\Omega \frac{dv}{dx} f\,dx = -\int_{\Omega_1} \frac{dv}{dx} f\,dx - \int_{\Omega_2} \frac{dv}{dx} f\,dx = -\lim_{\epsilon\to 0}\left(\int_0^{\hat{x}-\epsilon} \frac{dv}{dx} f\,dx + \int_{\hat{x}+\epsilon}^L \frac{dv}{dx} f\,dx\right). \tag{5.11}$$

Since each of the integrals in the last expression involve $\mathcal{C}^1(\Omega)$ functions (v is smooth everywhere and the point of non-smoothness in f has been removed), we can apply standard integration-by-parts to yield

$$-\int_0^{\hat{x}-\epsilon} \frac{dv}{dx} f\,dx = \int_0^{\hat{x}-\epsilon} v\frac{df}{dx}\,dx - [vf]_0^{\hat{x}-\epsilon}, \qquad -\int_{\hat{x}+\epsilon}^L \frac{dv}{dx} f\,dx = \int_{\hat{x}+\epsilon}^L v\frac{df}{dx}\,dx - [vf]_{\hat{x}+\epsilon}^L. \tag{5.12}$$

Since $v = 0$ at the boundaries, this reduces (5.11) to

$$-\int_\Omega \frac{dv}{dx} f\,dx = \lim_{\epsilon\to 0}\left(\int_0^{\hat{x}-\epsilon} vg\,dx + \int_{\hat{x}+\epsilon}^L vg\,dx + [vf]_{\hat{x}-\epsilon}^{\hat{x}+\epsilon}\right), \tag{5.13}$$

where we have defined $g \in \mathcal{F}_{\Omega\to\mathbb{R}}$ to agree with the derivative of f on each interval

$$g|_{\Omega_i} = \frac{d}{dx}\, f|_{\Omega_i} \in \mathcal{P}^{k-1}(\Omega_i). \tag{5.14}$$

Clearly g is a piecewise polynomial function; however, we cannot make any claims regarding continuity at $\hat{x}$. Next, we apply the limit in (5.13) to obtain

$$-\int_\Omega \frac{dv}{dx} f\,dx = \int_\Omega vg\,dx. \tag{5.15}$$

The term $[vf]_{\hat{x}-\epsilon}^{\hat{x}+\epsilon}$ vanishes because both v and f are continuous. The remaining term resulted from the additive property of integration and the fact that the integral is indepedent of the integrand over sets of measure zero (individual points in $\mathbb{R}$), therefore the fact that g is not defined at $\hat{x}$ is not important. Therefore, the weak derivative of f does exist and is equal to the function g in (5.14). This result is trivial to extend to functions that are piecewise polynomial over any finite number of intervals.

While simple, this result has important implications for the FEM. First, notice that if $f \notin \mathcal{C}^0(\Omega)$ the boundary term would not drop out and the weak derivative would not exist. Therefore, this would not be a valid FE subspace; the bilinear form, even for second-order PDEs involves the first weak derivative of functions in the trial and test space. This is precisely the reason why we required the FE subspace in Chapter 4 to only consist of continuous functions. Second, notice that the second weak derivative of f does not exist, i.e., the first weak derivative is g a piecewise polynomial function that is not continuous, therefore its weak derivative does not exist. This means the space of piecewise polynomial $\mathcal{C}^0(\Omega)$ functions is not an acceptable FE space for fourth-order (or higher) PDEs since the bilinear form requires two weak derivatives.

Example 5.2: Weak derivatives of continuously differentiable piecewise polynomial functions

In this example, we will show that piecewise polynomial continuously differentiable functions have *two* weak derivatives that agree with our informal intuition (piecewise polynomials of lower degree). Consider $\Omega := (0, L) \subset \mathbb{R}$ partitioned into two intervals $\Omega_1 := (0, \hat{x})$ and $\Omega_2 := (\hat{x}, L)$. Let $f \in \mathcal{C}^1(\Omega)$ such that $f|_{\Omega_1} \in \mathcal{P}^k(\Omega_1)$ and $f|_{\Omega_2} \in \mathcal{P}^k(\Omega_2)$ ($k \geqslant 1$). Since f is differentiable in the conventional sense, the weak derivative agrees with the classical derivative $g := \frac{df}{dx}$ and, because f is a continuously differentiable piecewise polynomial function, we know that g is a continuous piecewise polynomial function

$$g|_{\Omega_i} = \frac{d}{dx} f|_{\Omega_i} \in \mathcal{P}^{k-1}(\Omega_i), \qquad g \in \mathcal{C}^0(\Omega). \tag{5.16}$$

In the previous example, we found the weak derivative of a continuous, piecewise polynomial function to be a piecewise polynomial function of one lower degree (not necessarily defined at $\hat{x}$)

$$h|_{\Omega_i} := \frac{d}{dx} g|_{\Omega_i} = \frac{d^2}{dx^2} f|_{\Omega_i}. \tag{5.17}$$

From this simple argument, we see that f has two weak derivatives whose definitions align with our (informal) intuition. Also note that unlike spaces of continuous piecewise polynomial functions, spaces of continuously differentiable piecewise polynomial functions provide a valid FE subspace for fourth-order PDEs.

Definition 5.3.3 (Weak derivatives in $\mathbb{R}^d$)**.** Let $\Omega \subset \mathbb{R}^d$. The kth *weak partial derivative* of a function $f : \Omega \to \mathbb{R}$, denoted $\frac{\partial f}{\partial x_k}$, exists if there exists $g_k \in L^2(\Omega)$ such that

$$\int_\Omega v g_k \, dx = -\int_\Omega \frac{\partial v}{\partial x_k} f \, dx \qquad \forall v \in \mathcal{C}_c^\infty(\Omega). \tag{5.18}$$

The associated weak gradient is

$$\nabla f := \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_d} \right) \in \left(L^2(\Omega) \right)^d. \tag{5.19}$$

Higher weak derivatives are defined by recursively applying definition (5.18), e.g., the (i, j)th second weak partial derivative of a function $f : \Omega \to \mathbb{R}$, denoted $\frac{\partial^2 f}{\partial x_i \partial x_k}$, exists if there exists $g_{ij} \in L^2(\Omega)$ such that

$$\int_\Omega v g_{ij} \, dx = -\int_\Omega \frac{\partial v}{\partial x_j} \frac{\partial f}{\partial x_i} \, dx \qquad \forall v \in \mathcal{C}_c^\infty(\Omega). \tag{5.20}$$

Remark 5.1. Notice (5.9) and (5.18) closely resemble the integration-by-parts formula derived in Chapter 2 for $\mathcal{C}^1$ functions without the boundary terms (because $v \in \mathcal{C}_c^\infty(\Omega)$ vanishes at the boundary). This is not by accident; after observing how foundational the integration-by-parts property is to the variational formulation of PDEs, we simply *define* a new type of derivative using this key property. Since the definition is based on integrability (rather than pointwise limits), it applies to a much broader class of functions and *extends* or *generalizes* the familiar concept of a derivative, e.g., to functions that are not smooth, continuous, or even defined pointwise (recall Lebesgue integration is invariant with respect to the behavior of a function over sets of measure zero).

Remark 5.2. For functions that possess standard and weak derivatives, the definitions agree. The concept of a weak derivative is useful in that it formalizes some intuitive, yet non-rigorous results, e.g., the derivative of a piecewise linear function is a piecewise constant functions (see Example 5.1).

With the concept of a weak derivative at our disposal, we turn to the construction of Sobolev spaces, which are function spaces characterized by the integrability of its functions and their weak derivatives.

5.4. Sobolev norms and spaces

Sobolev spaces are characterized by the existance and integrability of the functions and their weak derivatives. For compactness, we represent partial derivatives of functions using *multi-index* notation. An n-dimensional multi-index is an element $\alpha \in \mathbb{N}_0^n$ with entries $\alpha = (\alpha_1, \dots, \alpha_n)$. The *order* or magnitude of the multi-index is given by its sum: $|\alpha| = \sum_{i=1}^n \alpha_i$. In this chapter, multi-index notation will be used to define a partial derivative of a multi-dimensional function:

$$(D^\alpha f)(x) := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}(x). \tag{5.21}$$

The order of the derivative is given by $|\alpha|$.

To construct the Sobolev spaces, we first define the *Sobolev norm*. Then the Sobolev spaces are defined as the set of functions from Ω to $\mathbb{R}$ with finite Sobolev norm.

Definition 5.4.1 (Sobolev norm)**.** Let $k \in \mathbb{N}_0$ and $f \in \mathcal{F}_{\Omega \to \mathbb{R}}$ and suppose the weak derivatives $D^\alpha f$ exists for all $|\alpha| \leqslant k$. The Sobolev norm is defined as

$$\|f\|_{W_p^k(\Omega)} := \left(\sum_{|\alpha| \leqslant k} \|D^\alpha f\|_{L^p(\Omega)}^p \right)^{1/p} \tag{5.22}$$

for $1 \leqslant p < \infty$ and for the case $p = \infty$

$$\|f\|_{W_\infty^k(\Omega)} := \max_{|\alpha| \leqslant k} \|D^\alpha f\|_{L^\infty(\Omega)}. \tag{5.23}$$

Definition 5.4.2 (Sobolev spaces)**.** The *Sobolev spaces* are

$$W_p^k(\Omega) := \left\{ f \in \mathcal{F}_{\Omega \to \mathbb{R}} \;\middle|\; \|f\|_{W_p^k(\Omega)} < \infty \right\} \tag{5.24}$$

for $1 \leqslant p < \infty$ and $k \in \mathbb{N}_0$.

The Sobolev space could be equivalently defined as the collection of functions where all weak derivatives up to order k belong to $L^p(\Omega)$

$$W_p^k(\Omega) := \{ f \in \mathcal{F}_{\Omega \to \mathbb{R}} \mid D^\alpha f \in L^p(\Omega) \; \forall |\alpha| \leqslant k \}. \tag{5.25}$$

Theorem 5.1. *The Sobolev space* $(W_p^k(\Omega), \|\cdot\|_{W_p^k(\Omega)})$ *is a Banach space.*

Definition 5.4.3 ($H^k(\Omega)$ spaces)**.** The special case of the Sobolev spaces $W_p^k(\Omega)$ with $p = 2$, i.e., function spaces with square-integrable derivatives, are often denoted $H^k(\Omega)$

$$H^k(\Omega) := W_2^k(\Omega). \tag{5.26}$$

Theorem 5.2. *The Sobolev space* $H^k(\Omega)$ *is a Hilbert (complete inner-product) space with the inner product*

$$(w, v)_{H^k(\Omega)} := \sum_{|\alpha| \leqslant k} (D^\alpha w, D^\alpha v)_{L^2(\Omega)}. \tag{5.27}$$

Remark 5.3. The Sobolev norm in (5.22) for the case $p = 2$, $k \in \mathbb{N}_0$ agrees with the norm induced by the inner product $\|v\|_{H^k(\Omega)} := \sqrt{(v, v)_{H^k(\Omega)}}$.

Example 5.3: Functions of a single variable

Let $\Omega \subset \mathbb{R}$ and consider $f \in \mathcal{F}_{\Omega \to \mathbb{R}}$. The multi-index derivative operator reduces to standard differentiation

$$D^\alpha f = \frac{d^\alpha f}{dx^\alpha} \tag{5.28}$$

where $\alpha \in \mathbb{N}$. In this case, the Sobolev norm is

$$\|f\|_{W_p^k(\Omega)}^p = \sum_{\alpha=0}^{k} \left\| \frac{d^\alpha f}{dx^\alpha} \right\|_{L^p(\Omega)} = \|f\|_{L^p(\Omega)} + \left\| \frac{df}{dx} \right\|_{L^p(\Omega)} + \left\| \frac{d^2 f}{dx^2} \right\|_{L^p(\Omega)} + \cdots + \left\| \frac{d^k f}{dx^k} \right\|_{L^p(\Omega)}. \tag{5.29}$$

The special case where $k = 1$, $p = 2$ is of particular interest in our study of the finite element method

$$\|f\|_{H^1(\Omega)}^2 = \|f\|_{L^2(\Omega)} + \left\| \frac{df}{dx} \right\|_{L^2(\Omega)} = \int_\Omega \left(f^2 + \frac{df}{dx}^2 \right) dx. \tag{5.30}$$

Similarly, as we will see $H^1(\Omega)$ and $H^2(\Omega)$ are the most widely used Sobolev spaces

$$\begin{aligned} H^1(\Omega) &= \left\{ f \in \mathcal{F}_{\Omega\to\mathbb{R}} \;\middle|\; f, \frac{df}{dx} \in L^2(\Omega) \right\} \\ H^2(\Omega) &= \left\{ f \in \mathcal{F}_{\Omega\to\mathbb{R}} \;\middle|\; f, \frac{df}{dx}, \frac{d^2f}{dx^2} \in L^2(\Omega) \right\}. \end{aligned} \tag{5.31}$$

Example 5.4: Functions of two variables

Let $\Omega \subset \mathbb{R}^2$ and consider $f \in \mathcal{F}_{\Omega\to\mathbb{R}}$. Several examples of the multi-index derivative operator in two variables are

$$\begin{aligned} \alpha = (0,0)^T &\qquad\Longrightarrow\qquad D^\alpha f = f \\ \alpha = (1,0)^T &\qquad\Longrightarrow\qquad D^\alpha f = \frac{\partial f}{\partial x_1} \\ \alpha = (0,1)^T &\qquad\Longrightarrow\qquad D^\alpha f = \frac{\partial f}{\partial x_2} \\ \alpha = (2,0)^T &\qquad\Longrightarrow\qquad D^\alpha f = \frac{\partial^2 f}{\partial x_1^2} \\ \alpha = (1,1)^T &\qquad\Longrightarrow\qquad D^\alpha f = \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \alpha = (0,2)^T &\qquad\Longrightarrow\qquad D^\alpha f = \frac{\partial^2 f}{\partial x_2^2} \\ \alpha = (m,n)^T &\qquad\Longrightarrow\qquad D^\alpha f = \frac{\partial^{m+n} f}{\partial x_1^m \partial x_2^n}. \end{aligned} \tag{5.32}$$

The $H^1(\Omega)$ and $H^2(\Omega)$ function spaces are

$$\begin{aligned} H^1(\Omega) &= \left\{ f \in \mathcal{F}_{\Omega\to\mathbb{R}} \;\middle|\; f, \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \in L^2(\Omega) \right\} \\ H^2(\Omega) &= \left\{ f \in \mathcal{F}_{\Omega\to\mathbb{R}} \;\middle|\; f, \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial^2 f}{\partial x_1^2}, \frac{\partial^2 f}{\partial x_1 \partial x_2}, \frac{\partial^2 f}{\partial x_2^2} \in L^2(\Omega) \right\}. \end{aligned} \tag{5.33}$$

Example 5.5: Spaces of piecewise polynomial functions

In Examples 5.1-5.2 in the previous section, we showed, in the case where $\Omega \subset \mathbb{R}$, that piecewise polynomial $\mathcal{C}^0(\Omega)$ functions have one and only one weak derivative, which is a piecewise polynomial function of one lower degree. Since piecewise polynomial functions are integrable (regardless of continuity), both the function and its weak derivative are integrable, which implies $\mathcal{C}^0(\Omega)$ piecewise polynomial functions belong to $W_p^1(\Omega)$ for any $1 \leqslant p \leqslant \infty$ ($H^1(\Omega)$ if $p = 2$). Similarly piecewise polynomial $\mathcal{C}^1(\Omega)$ functions belong to $W_p^2(\Omega)$ ($H^2(\Omega)$ if $p = 2$).

In this course, we will focus primarily on the Lebesgue space $L^2(\Omega)$ and Sobolev spaces $H^1(\Omega)$ and $H^2(\Omega)$.

5.5. Poisson equation

Let $\Omega \subset \mathbb{R}^d$ (Lipschitz) and partition the boundary into $\partial\Omega = \overline{\partial\Omega_D \cup \partial\Omega_N}$. The strong formulation of the Poisson equation with essential and natural boundary conditions is: find $u \in H^2(\Omega)$ such that

$$-\Delta u = f \quad \text{in } \Omega, \qquad u = g \quad \text{on } \partial\Omega_D, \qquad \nabla u \cdot n = h \quad \text{on } \partial\Omega_N, \tag{5.34}$$

where $f \in L^2(\Omega)$ is a source term, $g \in H^1(\partial\Omega)$ is the essential boundary condition, $h \in L^2(\partial\Omega)$ is the natural boundary conditions, and $n : \partial\Omega \to \mathbb{R}^d$ is the unit outward normal. Notice that we had to choose $u \in H^2(\Omega)$ to ensure the weak second derivative exists.

Following the procedure detailed in Chapter 3, we derive the weak formulation from the weighted residual formulation: find u such that

$$\int_{\Omega} w(-u_{,ii} - f)\, dx \tag{5.35}$$

for all w. Next we apply integration-by-parts to yield

$$\int_{\Omega} w_{,i} u_{,i} = \int_{\Omega} wf\, dx + \int_{\partial\Omega} w u_i n_i\, ds. \tag{5.36}$$

We require the test function to be zero on $\partial\Omega_D$ and apply the natural boundary condition to reduce the boundary term to

$$\int_{\partial\Omega} w u_i n_i\, ds = \int_{\partial\Omega_D} \underbrace{w}_{=0 \text{ on } \partial\Omega_D} u_i n_i\, ds + \int_{\partial\Omega_N} w \underbrace{u_i n_i}_{=h \text{ on } \partial\Omega_N} ds = \int_{\partial\Omega_N} wh\, ds. \tag{5.37}$$

Finally, we define the weak formulation to be: find $u \in \mathcal{V}$ such that

$$\int_{\Omega} w_{,i} u_{,i} = \int_{\Omega} wf\, dx + \int_{\partial\Omega_N} wh\, ds \tag{5.38}$$

for all $w \in \mathcal{V}^0$, where the spaces $\mathcal{V}, \mathcal{V}^0 \subset H^1(\Omega)$ are the subsets of $H^1(\Omega)$ that satisfy the non-homogeneous and homogeneous essential boundary conditions, respectively,

$$\mathcal{V} := \left\{ v \in H^1(\Omega) \mid v|_{\partial\Omega_D} = g \right\}, \qquad \mathcal{V}^0 := \left\{ v \in H^1(\Omega) \mid v|_{\partial\Omega_D} = 0 \right\}. \tag{5.39}$$

However, notice that the derivation is a formal argument, not a rigorous one since integration-by-parts, as presented in Chapter 2, requires the functions to be continuously differentiable. Fortunately, even in this more general Sobolev setting, the strong and weak formulations are equivalent (Theorem 5.3) because integration-by-parts holds for functions in H^1 (Proposition 5.4).

Proposition 5.4 (Integration-by-parts in $H^1(\Omega)$)**.** *Let $v, w \in H^1(\Omega)$. Then, for $i = 1, \ldots, d$*

$$\int_{\Omega} \left(\frac{\partial v}{\partial x_i}\right) w\, dx = -\int_{\Omega} v \left(\frac{\partial w}{\partial x_i}\right) dx + \int_{\partial\Omega} v w n_i. \tag{5.40}$$

Proposition 5.5 (Divergence theorem in $H^2(\Omega)$)**.** *Let $u \in H^2(\Omega)$ and $v \in H^1(\Omega)$. Then*

$$\int_{\Omega} -u_{,ii} v\, dx = \int_{\Omega} u_{,i} v_{,i}\, dx - \int_{\partial\Omega} v u_{,i} n_i\, ds. \tag{5.41}$$

Proof. Apply Proposition 5.4 with $v := -u_{,i}$ and $w := v$ and sum over $i = 1, \ldots, d$. □

Theorem 5.3 (Poisson equation: Strong-weak formulation equivalence)**.** *If $u \in H^2(\Omega)$ is a solution to the strong formulation of the Poisson equation (5.34), then it is a solution of the weak formulation. If u is a solution of the weak formulation of the Poisson equation (5.38), then it is a solution of the strong formulation provided $u \in H^2(\Omega)$.*

Proof. Suppose $u \in H^2(\Omega)$ satisfies the strong formulation (5.34). Then clearly $u \in \mathcal{V}$ because $H^2(\Omega) \subset H^1(\Omega)$ and $u|_{\partial\Omega_D} = g$. Then for any $w \in \mathcal{V}^0$ we have

$$\int_{\Omega} w(-u_{,ii} - f)\, dx = 0. \tag{5.42}$$

From Proposition 5.5, this is equivalent to

$$\int_{\Omega} u_{,i} w_{,i}\, dx = \int_{\Omega} wf\, dx + \int_{\partial\Omega} w u_{,i} n_i. \tag{5.43}$$

Finally, we use the fact that $w|_{\partial\Omega_D} = 0$ because $w \in \mathcal{V}^0$ and $(u_i n_i)|_{\partial\Omega_N} = h$ to obtain the weak formulation. Therefore any solution of the strong formulation satisfies the weak formulation. To prove the converse, suppose u satisfies the weak formulation (5.38) and $u \in H^2(\Omega)$. From Proposition 5.5, this implies

$$\int_\Omega w(-u_{,ii} - f)\, dx = \int_\Omega (w_{,i}u_{,i} - wf)\, dx - \int_\Omega wu_{,i}n_i\, ds = \int_{\partial\Omega_N} wh\, ds - \int_{\partial\Omega} wu_{,i}n_i\, ds, \tag{5.44}$$

for all $w \in \mathcal{V}^0$, which reduces to

$$\int_\Omega w(-u_{,ii} - f)\, dx = \int_{\partial\Omega_N} w(h - u_{,i}n_i)\, ds \tag{5.45}$$

by observing $w|_{\partial\Omega_D} = 0$. This only holds for an all $w \in \mathcal{V}^0$ if $-u_{,ii} = f$ in Ω and $u_{,i}n_i = h$ on $\partial\Omega_N$, which is precisely the strong formulation. □

To close this section, we write the weak formulation as a bilinear form:

$$\text{find } u \in \mathcal{V} \text{ such that } B(w,u) = \ell(w) \quad \forall w \in \mathcal{V}^0, \tag{5.46}$$

where $B : \mathcal{V}^0 \times \mathcal{V} \to \mathbb{R}$ is a bilinear functional and $\ell : \mathcal{V}^0 \to \mathbb{R}$ is a linear functional defined as

$$B(w,u) := \int_\Omega w_{,i}u_{,i}\, dx, \quad \ell(w) := \int_\Omega wf\, dx + \int_{\partial\Omega_N} wh\, ds. \tag{5.47}$$

As we will see in the next section, it is advantageous to write the weak formulation in such a way that the trial and test space are the same Hilbert space; as it stands currently the test space is a subspace of $H^1(\Omega)$ and therefore a Hilbert space under the inner product $(\cdot,\cdot)_{H^1(\Omega)}$ defined in (5.27), but the trial space is an affine subspace of $H^1(\Omega)$. To recast the weak formulation consider any $\varphi \in \mathcal{V}$, then any $u \in \mathcal{V}$ can be written as $u = \varphi + \bar{u}$ for some $\bar{u} \in \mathcal{V}^0$. Substituting this into the bilinear form we have

$$B(w,u) - \ell(w) = B(w, \varphi + \bar{u}) - \ell(w) = B(w, \bar{u}) - \ell(w) + B(w, \varphi) = B(w, \bar{u}) - \bar{\ell}(w) \tag{5.48}$$

where the functional $\bar{\ell} : \mathcal{V}^0 \to \mathbb{R}$ is defined as $\bar{\ell}(w) := \ell(w) - B(w, \varphi)$. Then the weak formulation reduces to

$$\text{find } \bar{u} \in \mathcal{V}^0 \text{ such that } B(w, \bar{u}) = \bar{\ell}(w) \quad \forall w \in \mathcal{V}^0, \tag{5.49}$$

where B is now interpreted as a bilinear form $B : \mathcal{V}^0 \times \mathcal{V}^0 \to \mathbb{R}$.

5.6. Well-posedness of variational problems

To close this chapter, we state conditions a bilinear form $a(\cdot,\cdot)$ and linear functional $b(\cdot)$ must satisfy such that the *variational problem*

$$\text{find } u \in V \text{ such that } a(w,u) = b(w) \quad \forall w \in V, \tag{5.50}$$

has a unique solution. From this variational problem, we introduce a finite-dimensional Galerkin approximation problem

$$\text{find } u_h \in V_h \text{ such that } a(w_h, u_h) = b(w_h) \quad \forall w_h \in V_h, \tag{5.51}$$

where $V_h \subset V$ is a finite-dimensional subspace. The results in this section are profound because, by employing this abstract variational setting, any boundary value problem that can be transformed into a variational problem that satisfies the required conditions will be guaranteed to possess a unique solution. This result can be used to conclude that both the infinite-dimensional weak formulation of a PDE and its finite-dimensional approximation possess a unique solutions. We begin with a defintion of the dual space to a Banach space.

Definition 5.6.1 (Dual space)**.** Consider a Banach space $(H, \|\cdot\|_H)$. The dual space to H, denoted H', is the set of linear functionals on H, i.e.,

$$H' := \{L \in \mathcal{F}_{H\to\mathbb{R}} \mid L(u + av) = L(u) + aL(v) \ \ \forall u, v \in H, a \in \mathbb{R}\}. \tag{5.52}$$

The concept of continuity was discussed briefly in Chapter 2. For the case of linear functionals, continuity is equivalent to *boundedness*. We generalize this notion of continuity/boundedness to bilinear forms and introduce a new concept called coercivity.

Proposition 5.6 (Continuity of linear functional)**.** *A linear functional L on a Banach space $(H, \|\cdot\|_H)$ is continuous if and only if it is bounded, i.e., if there $\exists C < \infty$ such that $|L(v)| \leqslant C \|v\|_H$ for all $v \in H$.*

Definition 5.6.2 (Continuity, coercivity of bilinear form)**.** Consider a normed linear space $(H, \|\cdot\|_H)$. A bilinear form $a : H \times H \to \mathbb{R}$ is *continuous* (or bounded) if $\exists C < \infty$ such that

$$|a(v,w)| \leqslant C \|v\|_H \|w\|_H \qquad \forall v, w \in H \tag{5.53}$$

and *coercive* on $V \subset H$ if $\exists \alpha > 0$ such that

$$a(v,v) \geqslant \alpha \|v\|_H^2 \qquad \forall v \in V. \tag{5.54}$$

With these definitions, we introduce the Lax-Milgram theorem (Theorem 5.4) that states variational problems of the form (5.50) possess a unique solution provided (1) bilinear form is continuous and coercive, (2) the linear functional is continuous, and (3) the trial and test space are the same Hilbert space.

Theorem 5.4 (Lax-Milgram)**.** *Given a Hilbert space $(V, (\cdot,\cdot))$, a continuous, coercive bilinear form $a(\cdot,\cdot)$ and a continuous linear functional $b \in V'$, there exists a unique $u \in V$ such that*

$$a(w,u) = b(w) \qquad \forall v \in V. \tag{5.55}$$

It can be shown that the bilinear form $B(\cdot,\cdot)$ defined in (5.49) is continuous and coercive, the linear form $\bar{\ell}(\cdot)$ in (5.49) is continuous, and $\mathcal{V}^0$ is a Hilbert space. Thus, the variational problem in (5.49), i.e., the weak formulation of the Poisson equation, has a unique solution (Lax-Milgram theorem). Furthermore, for any linear subspace $\mathcal{V}_h^0 \subset \mathcal{V}^0$, the finite-dimensional approximation: find $\bar{u}_h \in \mathcal{V}_h^0$ such that $B(w_h, \bar{u}_h) = \bar{\ell}(w_h)$ for all $w_h \in \mathcal{V}_h^0$, also has a unique solution (Lax-Milgram theorem). In the next chapter, we use the theortical foundations established in this chapter to construct a finite element space $\mathcal{V}_h^0$ such that it is a linear subspace of $\mathcal{V}^0$ (5.39), which is enough to guarantee well-posedness of the FE formulation.

5.7. Summary

In this chapter, we introduced fundamental concepts from functional analysis that lead to a rigorous variational formulation of boundary value problems and conditions under which it (and corresponding finite-dimensional approximations) is well-posed.

(1) Lebesgue integration provides a natural and rigorous setting to study variational formulations of boundary value problems because it can be used to define *complete* spaces, which is crucial in the proof of the Lax-Milgram theorem.

(2) The Lebesgue spaces $L^p(\Omega)$ are collections of Lebesgue integrable functions and form Banach spaces when combined with the corresponding norm $\|\cdot\|_{L^p(\Omega)}$.

(3) Since functions in Lebesgue spaces are not defined pointwise, a generalized concept of differentiation is required, called *weak derivatives*. Weak derivatives coincide with standard derivatives for classically differentiable functions and provide a rigorous definition of a derivative of non-smooth functions.

(4) Sobolev spaces are characterized by the integrability of the functions and their weak derivatives that comprise it. The most general Sobolev spaces are $W_p^k(\Omega)$; however, the most widely used ones are $H^k(\Omega)$ since they are Hilbert spaces when combined with a suitable inner product.
    - Piecewise polynomial $\mathcal{C}^0(\Omega)$ functions belong to $H^1(\Omega)$.
    - Piecewise polynomial $\mathcal{C}^1(\Omega)$ functions belong to $H^2(\Omega)$.

(5) Using these constructs, a rigorous definition of the strong and weak formulation of the Poisson equation were provided and their equivalence was proven.

(6) The Lax-Milgram theorem provides conditions that a general variational problem must satisfy to possess a unique solution. The weak formulation of the Poisson equation satisfies these conditions and therefore it possesses a unique solution. Furthermore, any finite-dimensional approximation of the weak formulation of the Poisson equation will possess a unique solution provided the approximation space is a subspace of the original Hilbert space.