Chapter 2

Galerkin Method for Linear Elliptic Problems

In this chapter, some of the main ingredients of variational boundary value problems, and the approximability, stability, and convergence of Galerkin approximations for elliptic problems is presented. To this end, we first briefly touch upon some of the essential concepts of linear functional analysis. In order to fix ideas, we consider the abstract variational problem as the basis for the exposition, and as an example, the weak formulation for the Poisson equation is derived. The approximate solution using a Galerkin method is presented, and some remarks on the convergence of Galerkin methods are mentioned. For essential prerequisites on functional analysis and cogent presentations on the mathematical aspects of elliptic boundary value problems, the texts of Strang and Fix (1973), Ciarlet (1978), Braess (1997), and Reddy (1998) are particularly noteworthy.

2.1 Linear Functional Analysis

Concepts from linear functional analysis are central to a qualitative understanding of boundary value problems of mathematical physics. The notion of function spaces in describing the topology of functions in $\mathbb{R}^d$ is analogous to that of vector spaces in linear algebra. In operator notation, if one considers the PDE $Au = f$, we note that the operator $A : U \to V$ maps functions from one function space U to another function space V. Hence it is apparent that the choice and suitability of function spaces is of significance and demands attention in the analysis of PDEs. Sobolev spaces (Adams, 1975) are the proper setting in order to study partial differential equations. We first briefly discuss Sobolev spaces with emphasis on a particular class of spaces known as Hilbert spaces, and then present linear functionals and bilinear forms which play a vital role in the study of linear boundary value problems. The foundations and theoretical basis are presented for scalar-valued functions; the theory, however, readily extends to vector-valued functions.

2.1.1 Sobolev Spaces

Let Ω be an open bounded domain in $\mathbb{R}^d$ with a piecewise smooth boundary Γ. In order to study linear elliptic PDEs, we consider real Sobolev spaces that are built on the function space $L^2(\Omega)$, where $L^2(\Omega)$ consists of functions that are square-integrable in the sense of Lebesgue:

$$L^2(\Omega) = \left\{ f : \int_\Omega f^2 \, d\Omega < +\infty \right\}. \tag{2.1}$$

The space $L^2(\Omega)$ with the scalar product

$$(u, v)_{L^2(\Omega)} = \int_\Omega uv \, d\Omega, \quad u,\, v \in L^2(\Omega) \tag{2.2}$$

and the associated norm $\|u\|_{L^2(\Omega)} = (u, u)^{1/2}_{L^2(\Omega)}$ becomes a Hilbert space.

As an aid to what follows, we first introduce two definition—multi-index notation and the notion of a *weak derivative.*

Multi-Index Notation. Let $\mathbb{Z}^n_+$ denote the set of all ordered n-tuples of non-negative integers: a member $\alpha \in \mathbb{Z}^n_+$ is denoted by $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, where each component α_i is a non-negative integer. The norm of α is denoted by $|\alpha|$ with $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$ and the partial derivative $D^\alpha u$ is given by

$$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}. \tag{2.3}$$

Weak Derivative. We denote $C_0^\infty(\Omega)$ as the space of infinitely differentiable functions which are non-zero only on a compact subset of Ω. If $u \in L^2(\Omega)$ is a locally integrable function, we say that u possesses the (weak) derivative $v = D^\alpha u$ in $L^2(\Omega)$ provided that $v \in L^2(\Omega)$ and

$$\langle v, \phi \rangle = (-1)^{|\alpha|} \langle u, D^\alpha \phi \rangle \quad \forall\ \phi \in C_0^\infty(\Omega) \tag{2.4}$$

in the sense of distributions.

The Sobolev space of order m, denoted by $H^m(\Omega)$, is defined as the space consisting of those functions in $L^2(\Omega)$ that, together with all their *weak partial derivatives* up to order m, belong to $L^2(\Omega)$:

$$H^m(\Omega) = \{u : D^\alpha u \in L^2(\Omega)\ \forall\ |\alpha| \leq m\}. \tag{2.5}$$

We introduce the Sobolev inner product $(\cdot,\cdot)_{H^m(\Omega)}$ to make $H^m(\Omega)$ an inner product space:

$$(u,v)_{H^m(\Omega)} = \sum_{|\alpha|\leq m} (D^\alpha u, D^\alpha v)_{L^2(\Omega)} \text{ for } u,\, v \in H^m(\Omega). \tag{2.6}$$

The inner product space induces the Sobolev norm $\|\cdot\|_{H^m(\Omega)}$ which is defined as

$$\|u\|^2_{H^m(\Omega)} = (u,u)_{H^m(\Omega)} = \sum_{|\alpha|\leq m} (D^\alpha u, D^\alpha u)_{L^2(\Omega)}. \tag{2.7}$$

We note that the Hilbert space $H^0(\Omega) = L^2(\Omega)$. The family of semi-norms $|\cdot|_{H^m(\Omega)}$ on the Hilbert space are defined as

$$|u|^2_{H^m(\Omega)} = \sum_{|\alpha|=m} (D^\alpha u, D^\alpha u)_{L^2(\Omega)}. \tag{2.8}$$

The distinction between the norm and the semi-norm is that the semi-norm does not satisfy the axiom of positive-definiteness; semi-norms are only positive-semidefinite. The relevance of semi-norms stem from the fact that $|\cdot|_{H^m(\Omega)}$ is a norm on $H^m_0(\Omega)$ and is equivalent to the H^m-norm, which is a consequence of the Poincaré-Friedrichs inequality (Ciarlet, 1978).

As a final remark on Sobolev spaces, we mention the Sobolev space $H^m_0(\Omega)$ which is the appropriate space for test functions (and also trial functions for the homogeneous Dirichlet problem) in Galerkin approximations of linear elliptic PDEs of order $2m$. For a rigorous and complete definition of $H^m_0(\Omega)$ that is based on the trace operator, see Ciarlet (1978). For the homogeneous Dirichlet problem, $H^m_0(\Omega)$ is a subspace of $H^m(\Omega)$ functions that, together with their derivatives of order less than m vanish on the boundary $\Gamma = \partial\Omega$. In the context of this work, the cases $m = 1$ and $m = 2$ are of particular interest. For scalar-valued functions, the Sobolev spaces $H_0^1(\Omega)$ and $H_0^2(\Omega)$ are defined as:

$$H_0^1(\Omega) = \{v : v \in H^1(\Omega),\ v = 0 \text{ on } \Gamma\}, \tag{2.9a}$$

$$H_0^2(\Omega) = \{v : v \in H^2(\Omega),\ v = 0,\ \partial v/\partial n = 0 \text{ on } \Gamma\}, \tag{2.9b}$$

where n is the outward normal direction to Ω.

2.1.2 Linear Functionals

We consider linear functionals $l : V \to \mathbb{R}$ that map real-valued functions from a Hilbert space V to real numbers in $\mathbb{R}$. The functional l is defined by its action on an arbitrary member $v \in V$ by

$$\langle l, v\rangle = (u, v)_{L^2(\Omega)} \tag{2.10}$$

for a given u and arbitrary v. The notation $\langle\cdot,\cdot\rangle$ is used to denote the duality pair. The space of bounded linear functionals on V is denoted by the dual space V', and for $l \in V'$ the norm $\|l\|_{V'}$ is defined as

$$\|l\|_{V'} = \sup_{v\in V\setminus\{0\}} \frac{\langle l, v\rangle}{\|v\|_V}. \tag{2.11}$$

The continuity of l is met if there exists a positive constant M such that

$$|\langle l, v\rangle| \leq M\|v\|_V. \tag{2.12}$$

The above equation is satisfied for $M = \|u\|_V$, a consequence of using the Cauchy-Schwartz inequality in Eq. (2.10), and it is easily shown that $\|l\|_{V'} = \|u\|_V$. The norm $\|l\|_{V'}$ is of relevance in ascertaining if the solution is continuously dependent on the data.

2.1.3 Bilinear Forms

Let us consider the bilinear form $a(u, v) : U \times V \to \mathbb{R}$ that maps pair of functions from Hilbert spaces U and V to real numbers in $\mathbb{R}$. The operator $a(\cdot, \cdot)$ is linear in both slots; hence the term bilinear. We outline some of the most important properties of bilinear forms that are germane to the existence and uniqueness of solutions for linear elliptic PDEs.

Continuity Consider Hilbert spaces U and V. A bilinear form $a(u, v) : U \times V \to \mathbb{R}$ is continuous if

$$|a(u, v)| \leq K \|u\|_U \|v\|_V \quad \forall\ u \in U,\ v \in V \tag{2.13}$$

for a positive constant K.

Ellipticity Let $v \in V$, where V is a Hilbert space. The V-ellipticity (coercivity) condition points to the positive-definiteness of the bilinear form $a(\cdot, \cdot)$:

$$a(v, v) \geq \alpha \|v\|_V^2 \tag{2.14}$$

for a constant $\alpha > 0$. If v is a finite-dimensional approximation and the bilinear form $a(v, v) = \mathbf{v}^T \mathbf{A} \mathbf{v}$, then the ellipticity condition indicates the existence of the inverse matrix $\mathbf{A}^{-1}$ and the fact that $\|\mathbf{A}^{-1}\|$ is bounded away from zero. In linear elasticity, the ellipticity of the continuous bilinear operator is a consequence of the positive-definiteness of the strain (stored) energy.

2.2 Variational Boundary Value Problem

Consider a Hilbert space V, and let $a(u, v) : V \times V \to \mathbb{R}$ be a continuous, V-elliptic bilinear form on V. In addition, let $l : V \to \mathbb{R}$ be a continuous linear functional on V. We consider the abstract variational boundary value problem:

$$a(u, v) = \langle l, v \rangle \quad \forall\, v \in V. \tag{2.15}$$

The well-posedness of the variational problem—existence and uniqueness of u with u being continuously dependent on the data, is a consequence of the Lax-Milgram theorem (Ciarlet, 1978). From a physical viewpoint, any linear problem that arises from a variational principle which involves the minimization (critical point) of a quadratic functional (potential energy functional in linear elasticity) has the above form. The minimizing problem is posed as:

$$\inf_{u \in V} J(u) = \inf_{u \in V} \frac{1}{2} a(u, u) - \langle l, u \rangle. \tag{2.16}$$

The weak form, however, is more general with no dependence on the existence of a variational principle. The required regularity (smoothness) of the solution $u \in H^s(\Omega)$

is also inferred on the basis of the condition that the integrands in the weak formulation are square-integrable. The variational problem posed in Eq. (2.15) can be obtained by an extremal variational principle (compressible elasticity) or by a stationary functional as in a saddle-point principle (incompressible elasticity).

Let us consider $2m$-order elliptic boundary value problems, where $2m$ is the order of the highest derivative of the primary variable in the partial differential equation. We partition the set of boundary conditions into two subsets: essential boundary conditions which are of order $< m$ and natural boundary conditions which are of order $\geq m$. This allows one to formulate the variational problem such that the solution is required only in a subspace of $H^m(\Omega)$ (Strang and Fix, 1973).

In order to illustrate the derivation of the weak formulation, we consider the Poisson equation with Dirichlet boundary conditions:

$$-\nabla^2 u = f \quad \text{in } \Omega, \tag{2.17a}$$

$$u = 0 \quad \text{on } \Gamma, \tag{2.17b}$$

where $\Omega \subset \mathbb{R}^2$ is an open bounded domain and the data f is assumed to be sufficiently smooth. Since $m = 1$ for the Poisson equation and only homogeneous Dirichlet boundary conditions are prescribed, we choose $V = H_0^1(\Omega)$. Now, on multiplying the above equation by an arbitrary function $v \in V$ and integrating over Ω, we obtain

$$-\int_\Omega (\nabla^2 u) v \, d\Omega = \int_\Omega f v \, d\Omega. \tag{2.18}$$

By Green's formulas, we have the identities:

$$\int_\Omega v\nabla^2 u \, d\Omega = -\int_\Omega \boldsymbol{\nabla} u \cdot \boldsymbol{\nabla} v \, d\Omega + \int_\Gamma v \frac{\partial u}{\partial n} \, d\Gamma, \tag{2.19a}$$

$$\int_\Omega (u\nabla^2 v - v\nabla^2 u) \, d\Omega = \int_\Gamma \left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n} \right) \, d\Gamma, \tag{2.19b}$$

where n is the outward normal direction to Ω. Since $v = 0$ on Γ by Eq. (2.17b), the boundary integral in Eq. (2.19a) vanishes, and hence Eq. (2.18) can be written as

$$\int_\Omega \boldsymbol{\nabla} u \cdot \boldsymbol{\nabla} v \, d\Omega = \int_\Omega f v \, d\Omega \tag{2.20}$$

which is known as the weak or variational form. Comparing Eq. (2.17) to the abstract variational form in Eq. (2.15), we have $V = H_0^1(\Omega)$ and

$$a(u, v) = \int_\Omega \boldsymbol{\nabla} u \cdot \boldsymbol{\nabla} v \, d\Omega, \quad \langle l, v \rangle = \int_\Omega f v \, d\Omega. \tag{2.21}$$

It is noted that by re-casting Eq. (2.17) in its weak form, the Poisson system in Eq. (2.17) is satisfied in the distributional sense—see Section 2.1.1. The bilinear form in Eq. (2.21) is symmetric, and the space of admissible trial functions has been enlarged from $H^2(\Omega)$ to $H^1(\Omega)$ by transforming the strong form to its associated weak or variational form. Thus, as the name suggests, the weak form in Eq. (2.21) permits weak solutions to the Poisson equation.

2.3 Galerkin Method

In the variational boundary value problem stated in Eq. (2.15), V is an infinite-dimensional subspace of the Hilbert space H. In order to present an approximate solution, we choose a finite-dimensional subspace $V^h \subset V$ which is spanned by the shape functions ϕ_I, ϕ_2, ... , ϕ_M:

$$V^h = \text{span } \{\phi_I\}_{I=1}^{M}, \tag{2.22}$$

where h refers to a measure of the distance between nodes or the element size in case of finite elements. If the number of shape functions gets larger ($M \to \infty$), then $h \to 0$. The variational form for the discrete system can now be written as:

$$a(u^h, v^h) = \langle l, v^h \rangle \quad \forall\ v^h \in V^h. \tag{2.23}$$

The above equation is the statement of the Galerkin method. In a Bubnov-Galerkin or Ritz-Galerkin method, both the trial function u^h and the arbitrary test function v^h are expressed as linear combinations of the same shape functions. In Petrov-Galerkin methods, u^h and v^h are chosen from different finite-dimensional subspaces. We can express u^h and v^h for a Bubnov-Galerkin method as

$$u^h = \sum_{I=1}^{M} \phi_I u_I, \quad v^h = \sum_{I=1}^{M} \phi_I v_I. \tag{2.24}$$

Substituting Eq. (2.24) in Eq. (2.23) and noting that a is bilinear and l is linear, we obtain

$$\sum_{I=1}^{M}\sum_{J=1}^{M} a(\phi_I,\phi_J)u_I v_J = \sum_{J=1}^{M}\langle l,\phi_J\rangle v_J \tag{2.25}$$

which can be written as

$$\sum_{J=1}^{M} v_J \left(\sum_{I=1}^{M} K_{IJ}u_I - f_J \right) = 0, \tag{2.26a}$$

where

$$K_{IJ} = a(\phi_I,\phi_J), \quad f_J = \langle l,\phi_J\rangle. \tag{2.26b}$$

Since the coefficients in Eq. (2.26a) are arbitrary, the term within the parentheses must be identically equal to zero, which results in the following discrete matrix system:

$$\mathbf{Ku} = \mathbf{f}, \tag{2.27}$$

where the symmetry of $\mathbf{K}$ has been used to arrive at the above equation. In solid mechanics, the matrix $\mathbf{K}$ is referred to as the stiffness matrix and $\mathbf{f}$ is the external force vector. For displacement-based finite element methods or for the natural element method, $\mathbf{u}$ is the nodal displacement vector, while for meshless methods such as EFG, $\mathbf{u}$ is the vector of nodal coefficients. In a Bubnov-Galerkin method, the stiffness matrix is symmetric and also positive definite when rigid-body translation and rotation modes are removed. For Petrov-Galerkin methods, a non-symmetric $\mathbf{K}$

matrix is obtained. It is noted that, in general, only few shape functions are non-zero in the summation in Eq. (2.24). For instance, in finite elements, for a point $\mathbf{x}$ in the interior of an element E, only the shape functions associated with the nodes in the connectivity of E are non-zero at $\mathbf{x}$.

2.3.1 Approximability, Stability, and Convergence

Having outlined the Galerkin method for the numerical solution of elliptic PDEs, the crux that remains is in ascertaining the approximability, stability, and convergence properties of the numerical method. The presentation that ensues closely follows that pursued by Arnold (1990).

Let $u \in V = H^m(\Omega)$ and u^h, $v^h \in V^h \subset V$ for a Dirichlet problem. We consider a suitable norm such as the energy norm $\|v\|_V := \big(a(v,v)\big)^{1/2}$. By virtue of continuity and ellipticity of the bilinear form, we have $\alpha\|v\|_V^2 \leq a(v,v) \leq K\|v\|_V^2$, and hence the energy norm is an equivalent norm to the H^m-norm in V. The approximability and convergence of a Galerkin method are measured by the quantities

$$\inf_{v^h \in V^h} \|u - v^h\| \quad \text{and} \quad \|u - u^h\|, \tag{2.28}$$

respectively. In order to quantify stability, the dual norm indicated in Eq. (2.11) is used:

$$\|l\|_{V'} = \sup_{v^h \in V^h \setminus \{0\}} \frac{\langle l, v \rangle}{\|v^h\|_V} \tag{2.29}$$

for $l : V^h \to \mathbb{R}$. Then, the stability constant for the Galerkin method is given by

$$M = \sup_{l:V^h \to \mathbb{R}} \frac{\|u^h\|_V}{\|l\|_{V'}} \tag{2.30}$$

which indicates that the stability constant is the largest value attained by the ratio of the solution to the discrete problem to that of the dual norm generated by the data. The stability constant $M = 1/\alpha$ where α is defined in Eq. (2.14). For the discrete problem, the stability constant is just the norm of the inverse matrix. The fundamental relation between approximability, stability, and convergence is given by Céa's lemma (Ciarlet, 1978):

$$\boxed{\|u - u^h\| \leq C \inf_{v^h \in V^h} \|u - v^h\|.} \tag{2.31}$$

Hence if the subspaces V^h of V are such that, for each $u \in V$

$$\lim_{h \to 0} \inf_{v^h \in V^h} \|u - v^h\| = 0, \tag{2.32}$$

then convergence is attained and u^h converges to u at the same rate as the best approximation. By the property of continuity and ellipticity of the bilinear form $a(\cdot,\cdot)$, an estimate $C = K/\alpha$ is obtained; if the bilinear form is symmetric, an even tighter bound $C = \sqrt{(K/\alpha)}$ is realized (Ciarlet, 1978), where K and α are defined in Eqs. (2.13) and (2.14). As seen above, good approximability in tandem with stability leads to approximate solutions that converge to the exact solution at the same rate as the best approximation. The sufficiency condition in Eq. (2.31) indicates that the error estimate $\|u - u^h\|$ is reduced to finding the distance $d(u, V^h) = \inf_{v^h \in V^h} \|u - v^h\|$

between a function $u \in V$ and a subspace $V^h \subset V$ (Fig. 2.1). Typically, the distance is in itself bounded by a constant times an exponent of the nodal spacing h. If the bilinear form is smooth and satisfies the continuity and ellipticity conditions given in Eqs. (2.13) and (2.14), then for linear elliptic problems of order $2m$, we have the estimate (Strang and Fix, 1973)

$$\|u - u^h\|_s \leq \begin{cases} Ch^{k-s}\|u\|_k & \text{if } s \geq 2m-k, \\ Ch^{2(k-m)}\|u\|_k & \text{if } s \leq 2m-k, \end{cases} \tag{2.33}$$

where C is a constant independent of h. In the above equation, the trial space is assumed to possess $k-1$ completeness ($k = 2$ for an approximation space with linear completeness), and the exponent of h is known as the rate of convergence. For linear elliptic PDEs of order $2m$, mth order derivatives appear in the weak form. Clearly, in this case, optimal convergence in H^s-norms ($s \leq m$) is attained if and only if $k > m$.

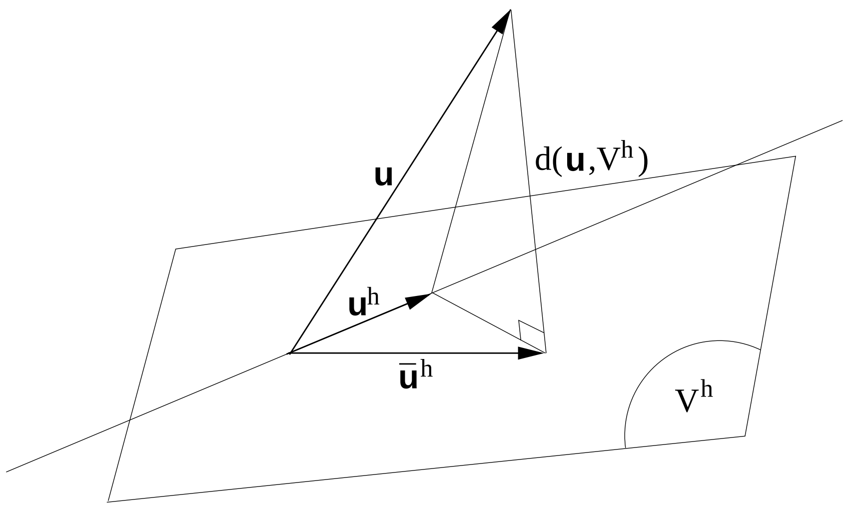

Figure 2.1: Schematic representation of convergence in Galerkin methods.

For displacement-based Galerkin methods in linear elasticity, the bilinear form is symmetric and V-elliptic, and hence stability and convergence are readily established (Ciarlet, 1978). As opposed to displacement methods, in mixed methods based on incompressible (or near incompressible) elasticity, the bilinear form is not coercive, and hence a stability condition termed as inf-sup or LBB (Ladyzhenskaya, 1969; Babuška, 1971; Brezzi, 1974) must be satisfied for convergence.