ADAPTIVE FINITE ELEMENT METHODS FOR ELLIPTIC PDE

RICARDO H. NOCHETTO

ABSTRACT. In the 80's and 90's a great deal of effort was devoted to the design of a posteriori error estimators for a variety of PDE. These are computable quantities, depending on the discrete solution(s) and data, that can be used to assess the quality of the approximation and improve it adaptively. Despite their practical success, adaptive processes have been shown to converge, and to exhibit optimal complexity, only recently and just for linear elliptic PDE. This course presents this new theory and discusses extensions and open questions.

CONTENTS

1. Lecture 1. The Finite Element Method: Properties and Error Analysis

This lecture introduces the concepts of weak solutions of elliptic partial differential equations (PDE) and Galerkin approximation, and presents the finite element method along with its basic properties and error analysis.

1.1. **Adaptive Approximation.**

We start with a simple motivation in 1D for the use of adaptive procedures. Given $\Omega = (0,1)$, a partition $\mathcal{T}_N = \{x_i\}_{n=0}^N$ of Ω

$$0 = x_0 < x_1 < \cdots < x_n < \cdots < x_N = 1$$

and a continuous function $u : \Omega \to \mathbb{R}$, we consider the problem of approximating u by a *piecewise constant* function u_N over $\mathcal{T}_N$. To quantify the difference between u and u_N we resort to the *maximum norm* and study two cases depending on the regularity of u.

Case 1: Smooth u. Suppose that u is Lipschitz in $[0,1]$. We consider the approximation

$$u_N(x) := u(x_{n-1}) \qquad \forall x_{n-1} \le x < x_n.$$

Since

$$|u(x) - u_N(x)| = |u(x) - u(x_{n-1})| = \Big| \int_{x_{n-1}}^x u'(t)dt \Big| \le h_n \|u'\|_{L^\infty(x_{n-1},x_n)}$$

we conclude that

$$\|u - u_N\|_{L^\infty(\Omega)} \le \frac{1}{N} \|u'\|_{L^\infty(\Omega)}, \tag{1.1}$$

provided the local meshsize h_n is about constant (*quasi-uniform* mesh), and so proportional to N^{-1} (the reciprocal of the number of degrees of freedom). A natural querstion arises: *Is it possible to approximate rough (or singular) functions u with a decay estimate similar to* (1.1)*?*

Case 2: Rough u. To answer this question, we suppose $\|u'\|_{L^1(\Omega)} = 1$ and consider the non-decreasing function

$$\phi(x) := \int_0^x |u'(t)|dt$$

which satisfies $\phi(0) = 0$ and $\phi(1) = 1$. Let $\mathcal{T}_N = \{x_i\}_{n=0}^N$ be the partition given by

$$\int_{x_{n-1}}^{x_n} |u'(t)|dt = \phi(x_n) - \phi(x_{n-1}) = \frac{1}{N}.$$

Then, for $x \in [x_{n-1}, x_n]$,

$$|u(x) - u(x_{n-1})| = \Big| \int_{x_{n-1}}^x u'(t)dt \Big| \le \int_{x_{n-1}}^x |u'(t)|dt \le \int_{x_{n-1}}^{x_n} |u'(t)|dt = \frac{1}{N},$$

whence

$$\|u - u_N\|_{L^\infty(\Omega)} \le \frac{1}{N} \|u'\|_{L^1(\Omega)}.$$

A relevant example, which mimics corner singularities in higher dimensions, is the function $u(x) := x^r$ with $0 < \gamma < 1$. It is easy to see that

$$x_n = \Big(\frac{n}{N}\Big)^{\frac{1}{\gamma}}, \qquad \forall\ 0 \le n \le N,$$

is the resulting *nonuniform mesh.* We thus conclude that we could achieve the same rate of convergence N^{-1} for rough functions with just $\|u'\|_{L^1(\Omega)} < \infty$ provided the partition is designed to *equidistribute* the error. However, such a partition may not be adequate for another function with the same basic regularity as u. We point out that such a regularity is Hölder γ, namely γ derivatives in $L^\infty(\Omega)$, while it increases to one full derivative if measure in $L^1(\Omega)$. This trade-off between differentiability and integrability is at the heart of the matter and is known as *nonlinear approximation theory* [20]

The function u_N may be the result of a minimization process. If we wish to minimize the norm $\|u-v\|_{L^2(\Omega)}$ within the space $\mathbb{V}_N$ of piecewise constant functions over $\mathcal{T}_N$, then it is easy to see that the solution u_N satisfies the orthogonality relation

$$u_N \in \mathbb{V}_N: \quad \langle u - u_N, v\rangle = 0 \qquad \forall\, v \in \mathbb{V}_N \tag{1.2}$$

and is given by the explicit local expression

$$u_N(x) = \frac{1}{h_n}\int_{x_{n-1}}^{x_n} u \qquad \forall\, x_{n-1} < x < x_n.$$

The previous comments apply to this u_N as well even though u_N coincides with u at an unknown point in each interval $[x_{n-1}, x_n]$.

The following issues arise and will be discussed in this course:

- PDE: The function u is not directly accessible but rather it is the solution of an elliptic PDE. We thus have to derive regularity and approximation properties of u. This is discussed in Lectures 1 and 4.
- FEM: We need a numerical method to approximate u which is sufficiently flexible to handle both geometry and accuracy (local mesh refinement) such as the finite element method (FEM). We then derive approximation properties of FEM via polynomial interpolation theory in the spirit of (1.2). This is explained in Lecture 1 for the energy norm and in Lecture 5 for the maximum norm.
- A posteriori error estimation: We need a practical procedure to estimate the local error and equidistribute it. This is explained in Lecture 2.
- Adaptivity: This is a concept associated with iterative loops of the form

$$\mathsf{SOLVE} \to \mathsf{ESTIMATE} \to \mathsf{MARK} \to \mathsf{REFINE}.$$

  Their convergence and complexity is the main subject of these lectures, and is addressed in Lectures 2, 3, and 4.

1.2. **Variational Formulation and Galerkin Method.** Let Ω be a polyhedral bounded domain in $\mathbb{R}^d$, $(d = 2, 3)$. We consider a homogeneous Dirichlet boundary value problem for a general second order elliptic partial differential equation (PDE)

$$\mathcal{L}u = -\operatorname{div}(\mathbf{A}\nabla u) + \mathbf{b}\cdot\nabla u + c\,u = f \quad \text{in } \Omega, \tag{1.3}$$

$$u = 0 \quad \text{on } \partial\Omega; \tag{1.4}$$

the choice of boundary condition is made for easy of presentation, since similar results are valid for other boundary conditions. We also assume

- $\mathbf{A}: \Omega \mapsto \mathbb{R}^{d\times d}$ is Lipschitz and symmetric positive definite with smallest eigenvalue a_- and largest eigenvalue a_+, i.e.,

$$a_-(x)\,|\xi|^2 \le \mathbf{A}(x)\xi\cdot\xi \le a_+(x)\,|\xi|^2, \quad \forall \xi \in \mathbb{R}^d,\ x\in\Omega; \tag{1.5}$$

- $\mathbf{b} \in [L^\infty(\Omega)]^d$ is divergence free ($\operatorname{div}\mathbf{b} = 0$ in Ω);
- $c \in L^\infty(\Omega)$ is nonnegative ($c \ge 0$ in Ω);
- $f \in L^2(\Omega)$.

For an open set $G \subset \mathbb{R}^d$ we denote by $H^1(G)$ the usual Sobolev space of functions in $L^2(G)$ whose first derivatives are also in $L^2(G)$, endowed with the norm

$$\|u\|_{H^1(G)} := \Big(\|u\|_{L^2(G)} + \|\nabla u\|_{L^2(G)}\Big)^{1/2};$$

we use the symbols $\|\cdot\|_{H^1}$ and $\|\cdot\|_{L^2}$ when $G = \Omega$. Moreover, we denote by $H^1_0(G)$ the space of functions in $H^1(G)$ with vanishing trace on the boundary. We set $\mathbb{V} := H^1_0(\Omega)$ and denote the norm $\|\cdot\|_{\mathbb{V}}$.

A weak solution of (1.3) and (1.4) is a function u satisfying

$$u \in \mathbb{V}: \qquad \mathcal{B}[u, v] = \langle f, v\rangle \qquad \forall v \in \mathbb{V}, \tag{1.6}$$

where $\langle u, v\rangle := \int_\Omega uv$ for any $u, v \in L^2(\Omega)$, and the bilinear form is defined on $\mathbb{V}\times\mathbb{V}$ as

$$\mathcal{B}[u,v] := \langle \mathbf{A}\nabla u, \nabla v\rangle + \langle \mathbf{b}\cdot\nabla u + c\,u, v\rangle\,. \tag{1.7}$$

By the Cauchy-Schwarz inequality one can easily show the *continuity* of the bilinear form

$$|\mathcal{B}[u,v]| \le C_B \, \|u\|_{\mathbb{V}} \, \|v\|_{\mathbb{V}}\,, \tag{1.8}$$

where C_B depends only on the data. Combining Poincaré inequality with the divergence free condition $\operatorname{div}\mathbf{b}=0$, one has *coercivity* in $\mathbb{V}$

$$\mathcal{B}[v,v] \ge \int_\Omega a_- \, |\nabla v|^2 + cv^2 \ge c_B \, \|v\|_{\mathbb{V}}^2\,, \tag{1.9}$$

where c_B depends only on the data. The bilinear form $\mathcal{B}$ induces the so-called *energy norm*:

$$|\!|\!| v |\!|\!| := \mathcal{B}[v,v]^{1/2} \qquad \forall\ v \in \mathbb{V}. \tag{1.10}$$

Existence and uniqueness of (1.6) thus follows from Lax-Milgram theorem [25]. The next critical issue is regularity of u. This is illustrated with the following 2D examples, all leading to point singularities.

Example 1.1 (Reentrant Corner)**.** Let Ω be a polygonal doamin with a reentrant corner $\omega > \pi$ at the origin, let $\gamma := \omega/\pi < 1$, and let u be the exact solution of the Dirichlet problem $\Delta u = 0$ in Ω be (in polar coordinates)

$$u(r,\theta) = r^\gamma \sin(\gamma\theta).$$

If $D^{1+s}u$ denotes a (formal) fractional derivative of u with $s < 1$, we note that $D^{1+s}u \in L^p(\Omega)$ provided $s < \frac{2}{p} + \gamma - 1$, because $\int_0^1 r^{p(\gamma-s-1)+1} dr < \infty$ in this case. We conclude that $u \in W_p^{1+s}(\Omega)$ for $1 \le p \le 2$ and, in particular, that $u \notin H^2(\Omega)$. If $-\Delta u = f$, then Grisvard shows that $\|u\|_{W_p^2(\Omega)} \le C_p \|f\|_{L^p(\Omega)}$ for all $1 < p < 4/3$ regardless of the size of corner angles [28]; p=4/3 corresponds to the crack problem $\omega = 2\pi$ for which $\gamma = 1/2$.

Example 1.2 (Mixed Boundary Condition)**.** Let Ω be a polygonal domain with flat boundary at the origin and let u be the exact solution $\Delta u = 0$, with mixed boundary condition $u = 0$ for $\theta = 0$ and $\partial_\nu = 0$ for $\theta = \pi$ (in polar coordinates), given by

$$u(r,\theta) = r^{\frac{1}{2}} \sin(\theta/2).$$

We infer that $\in W_p^{1+s}(\Omega)$ provided $s < \frac{2}{p} - \frac{1}{2}$, but $u \notin H^2(\Omega)$. The same asymptotic behavior of u is to be expected for a domain with a crack, namely $\omega = 2\pi$, and Dirichlet boundary condition.

Example 1.3 (Discontinuous $\mathbf{A}$)**.** We recall the formulas derived by Kellogg [29] to construct an exact solution of $\operatorname{div}(\mathbf{A}\nabla u) = f$ with piecewise constant coefficients $\mathbf{A}$ and vanishing right-hand side f. We now write these formulas in the particular case $\Omega = (-1,1)^2$, $\mathbf{A} = a_1\mathbf{I}$ in the first and third quadrants, and $\mathbf{A} = a_2\mathbf{I}$ in the second and fourth quadrants. An exact weak solution u is given (in polar coordinates) by $u(r,\theta) = r^\gamma\phi(\theta)$, where

$$\phi(\theta) = \begin{cases} \cos((\pi/2-\sigma)\gamma)\cdot\cos((\theta-\pi/2+\rho)\gamma) & \text{if } 0 \le \theta \le \pi/2, \\ \cos(\rho\gamma)\cdot\cos((\theta-\pi+\sigma)\gamma) & \text{if } \pi/2 \le \theta \le \pi, \\ \cos(\sigma\gamma)\cdot\cos((\theta-\pi-\rho)\gamma) & \text{if } \pi \le \theta < 3\pi/2, \\ \cos((\pi/2-\rho)\gamma)\cdot\cos((\theta-3\pi/2-\sigma)\gamma) & \text{if } 3\pi/2 \le \theta \le 2\pi, \end{cases}$$

and the numbers γ, ρ, σ satisfy the nonlinear relations

$$\begin{cases} R := a_1/a_2 = -\tan((\pi/2-\sigma)\gamma)\cdot\cot(\rho\gamma), \\ 1/R = -\tan(\rho\gamma)\cdot\cot(\sigma\gamma), \\ R = -\tan(\sigma\gamma)\cdot\cot((\pi/2-\rho)\gamma), \\ 0 < \gamma < 2, \\ \max\{0, \pi\gamma-\pi\} < 2\gamma\rho < \min\{\pi\gamma, \pi\}, \\ \max\{0, \pi-\pi\gamma\} < -2\gamma\sigma < \min\{\pi, 2\pi-\pi\gamma\}. \end{cases} \tag{1.11}$$

If we choose $\gamma = 0.1$, which produces a very singular solution u that is barely in H^1, and then solve (1.11) for R, ρ, and σ using Newton's method, we obtain

$$R = a_1/a_2 \cong 161.4476387975881, \qquad \rho = \pi/4, \qquad \sigma \cong -14.92256510455152,$$

and finally choose $a_1 = R$ and $a_2 = 1$. A smaller γ would lead to a larger ratio R, but in principle γ may be as close to 0 as desired. We see that again $u \in W_p^{1+s}(\Omega)$ for $s < \frac{2}{p} + \gamma - 1$.

Consider now a finite dimensional subspace $\mathbb{V}_N$ of $\mathbb{V}$ of dimension N. We formulate the *Galerkin method* upon restricting (1.6) to $\mathbb{V}_N$:

$$u_N \in \mathbb{V}_N : \qquad \mathcal{B}[u_N, v] = \langle f, v\rangle \qquad \forall v \in \mathbb{V}_N. \tag{1.12}$$

Existence and uniqueness of (1.12) follows also from the Lax-Milgram theorem on $\mathbb{V}_N$. In addition, if $\{\phi_j\}_{j=1}^N$ is a basis of $\mathbb{V}_N$, then we can write $u_N(x) = \sum_{j=1}^N U_j \phi_j(x)$ and (1.12) is equivalent to the linear system of equations

$$\sum_{j=1}^N U_j \mathcal{B}[\phi_j, \phi_i] = \langle f, \phi_i\rangle \qquad \forall 1 \le i \le N.$$

The ensuing matrix $(\mathcal{B}[\phi_j, \phi_i])_{i,j=1}^N$ is positive definite and non-symmetric unless $\mathbf{b} = 0$. The error function $u - u_N$ satisfies the following crucial property, usually called *Galerkin orthogonality*:

$$\mathcal{B}[u - u_N, v] = 0 \qquad \forall\ v \in \mathbb{V}_N. \tag{1.13}$$

This is a fundamental property for our error analysis and design of adaptive FEM (AFEM), which is not valid for finite differences.

1.3. **The Finite Element Method.** Let $\mathcal{T}_h$ be a partition of Ω into triangles (for $d = 2$) and tetrahedra (for $d = 3$) T of size $h_T = \operatorname{diam}(T)$; hereafter h denotes the piecewise constant function defined by $h|_T = h_T$ for $T \in \mathcal{T}_h$. We assume that $\mathcal{T}_h$ is *conforming*, namely the intersection of distinct elements T is either an edge, face, or vertex. Let $\{\mathcal{T}_h\}$ be a *shape-regular* family of nested conforming meshes over Ω: that is there exists a constant γ^* such that

$$\frac{h_T}{\rho_T} \le \gamma^* \qquad \forall\ T \in \bigcup_h \mathcal{T}_h, \tag{1.14}$$

where ρ_T is the diameter of the biggest ball contained in T. Let $\mathcal{N}_h = \{x_j\}_{j=1}^N$ be the set of internal nodes (or vertices) of $\mathcal{T}_h$. See Figure 1.1.

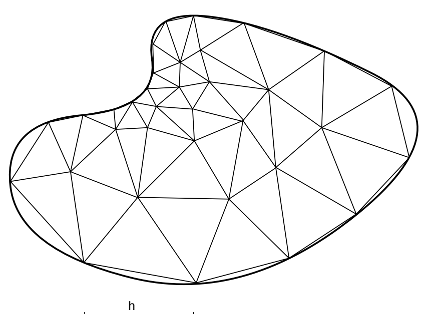

FIGURE 1.1. Conforming triangulation of a 2D domain Ω with local meshsize h

We consider now a subspace $\mathbb{V}_h$ of $\mathbb{V}$ of continuous piecewise polynomial functions over the mesh $\mathcal{T}_h$. If the polynomial degree is one, then the so-called *hat functions* are defined by its nodal values $\phi_i(x_j) = \delta_{ij}$ and are a basis of $\mathbb{V}_h$. The *finite element method* (FEM) is a Galerkin method with finite dimensional subspace $\mathbb{V}_h$:

$$u_h \in \mathbb{V}_h : \qquad \mathcal{B}[u_h, v] = \langle f, v\rangle \qquad \forall v \in \mathbb{V}_h. \tag{1.15}$$

If $\{\phi_i\}_{i=1}^N$ is chosen as a basis of $\mathbb{V}_h$, then the resulting matrix is sparse. We refer to [8, 15] for details about FEM, especially a discussion of higher order elements, unisolvence, and the solution of the resulting sparse linear systems of equations. We turn our attention to the error analysis.

1.4. **Error Analysis.** We now estimate the error $u-u_h$ in the energy norm $\mathbb{V}$. There are two distinct, but related, type of estimates depending on whether the continuous solution u or the discrete solution u_h occurs in the estimate. The first bound is called *a priori* error estimate whereas the second one is called *a posteriori* error estimate.

Lemma 1.4 (A Priori Error Estimate)**.** *Let $c_B \leq C_B$ be the constants in* (1.8) *and* (1.9)*. Then*

$$\inf_{v\in\mathbb{V}_h} \|u-v\|_{\mathbb{V}} \leq \|u-u_h\|_{\mathbb{V}} \leq \frac{C_B}{c_B} \inf_{v\in\mathbb{V}_h} \|u-v\|_{\mathbb{V}}. \tag{1.16}$$

Proof. It suffices to obtain the rightmost estimate. We then obtain for all $v \in \mathbb{V}_h$

$$\begin{aligned}
c_B\|u-u_h\|_{\mathbb{V}}^2 &\leq \mathcal{B}[u-u_h, u-u_h] && \text{(coercivity (1.9))}\\
&= \mathcal{B}[u-u_h, u-v] && \text{(orthogonality (1.13))}\\
&\leq C_B\|u-u_h\|_{\mathbb{V}}\|u-v\|_{\mathbb{V}}, && \text{(continuity (1.8))}
\end{aligned}$$

which implies the assertion. □

This says that the finite element solution is almost the *best approximation* to u within $\mathbb{V}_h$ in the norm of $H_0^1(\Omega)$. It thus motivates the study of piecewise polynomial approximation, which we discuss in section 1.5, but it does not provide quantitative information about the actual size of $\|u-u_h\|_{H^1}$. We point out though that if the error is measured in the energy norm (1.10), and $\mathbf{b}=0$ and so $\mathcal{B}$ is symmetric, then u_h is the best approximation to u within $\mathbb{V}_h$:

$$|\!|\!| u-u_h |\!|\!| = \inf_{v\in\mathbb{V}_h} |\!|\!| u-v |\!|\!| .$$

Quantitative error information is crucial to assess whether or not a given discretization yields a desired accuracy. To this end, we introduce the *residual* $\mathcal{R}(u_h) \in \mathbb{V}^* = H^{-1}(\Omega)$

$$\langle \mathcal{R}(u_h), v\rangle := \langle f, v\rangle - \mathcal{B}[u_h, v] \qquad \forall\ v \in \mathbb{V}_h, \tag{1.17}$$

along with its norm

$$\|\mathcal{R}(u_h)\|_{\mathbb{V}^*} := \sup_{v\in\mathbb{V}^*} \frac{\langle \mathcal{R}(u_h), v\rangle}{\|v\|_{\mathbb{V}}}. \tag{1.18}$$

The residual depends solely on data and the discrete solution u_h. The notation $\mathcal{R}(u_h)$ is meant as a reminder of the explicit dependence on u_h. The following simpe result illustrates the connection between residual and error.

Lemma 1.5 (A Posteriori Error Estimate)**.** *Let $c_B \leq C_B$ be the constants in* (1.8) *and* (1.9)*. Then*

$$c_B\|u-u_h\|_{\mathbb{V}} \leq \|\mathcal{R}(u_h)\|_{\mathbb{V}^*} \leq C_B\|u-u_h\|_{\mathbb{V}}. \tag{1.19}$$

Proof. We simply observe the crucial relation between error and residual

$$\mathcal{B}[u-u_h, v] = \langle f, v\rangle - \mathcal{B}[u_h, v] = \langle \mathcal{R}(u_h), v\rangle,$$

and use (1.9) to derive the lower bound and (1.8) to prove the upper bound. □

In order to be able to estimate $\|\mathcal{R}(u_h)\|_{\mathbb{V}^*}$, instead of $\|u-u_h\|_{\mathbb{V}}$, we need a practical way to deal with the negative norm in $\mathbb{V}^*$. This will be accomplished in sections 2.1.1 and 2.1.2. We draw the analogy between this framework and that in linear algebra: if $A \in \mathbb{R}^{N\times N}$, $\mathbf{x} \in \mathbb{R}^N$, and $\mathbf{b} \in \mathbb{R}^N$ satisfy $A\mathbf{x} = \mathbf{b}$, then for any $\mathbf{y} \in \mathbb{R}^N$ the residual is defined by $\mathbf{r} = \mathbf{b} - A\mathbf{y}$ whence the relation between error $\mathbf{e} = \mathbf{x} - \mathbf{y}$ and residual reads $A\mathbf{e} = \mathbf{r}$. We usually deduce information of $\mathbf{e}$ upon manipulating $\mathbf{r}$.

1.5. **Polynomial Interpolation in Sobolev Spaces.** We start with the definition of the Sobolev spaces

$$W_p^m(\Omega) := \{v \in L^p(\Omega) : \quad D^\alpha v \in L^p(\Omega) \quad \forall |\alpha| \le m\}$$

along with the corresponding *Sobolev number*

$$\mathrm{sob}(W_p^m) := m - \frac{d}{p}.$$

It is well known that if $m > k$ and $\mathrm{sob}(W_p^m) > \mathrm{sob}(W_q^k)$ then the embbeding $W_p^m(\Omega) \subset W_q^k(\Omega)$ is compact [25]. In addition, if $0 = \mathrm{sob}(L^\infty(\Omega)) < \gamma = \mathrm{sob}(W_p^m) < 1$, then $W_p^m(\Omega) \subset C^{0,\gamma}(\bar{\Omega})$ and so functions from $W_p^m(\Omega)$ are continuous. On the other hand, if $\mathrm{sob}(W_p^m) \le 0$ then pointvalues are not always well defined; this is the case of $H^1(\Omega) = W_2^1(\Omega)$ for $d \ge 2$.

The Sobolev number plays not only a crucial role in functional analysis but also in polynomial interpolation theory. This is due to the fact that it helps quantify the effect of scaling. Suppose that we change variables $x = h\hat{x}$, and the domain change from ω to $\hat{\omega}$. Then a simple calculation reveals

$$\|D^\alpha \hat{v}\|_{L^p(\hat{\omega})} = h^{\mathrm{sob}(W_p^m)} \|D^\alpha v\|_{L^p(\omega)} \qquad \forall\ |\alpha| = m. \tag{1.20}$$

In particular, invoking the Poincaré inequality in $\hat{\omega}$ we easily deduce

$$\|v\|_{L^2(\omega)} \le Ch \|\nabla v\|_{L^2(\omega)} \qquad \forall\ v \in H_0^1(\omega). \tag{1.21}$$

We now turn our attention to piecewise polynomial interpolation over a mesh $\mathcal{T}_h$ of Ω. If $v \in W_p^m(\Omega)$ is continuous, namely $\mathrm{sob}(W_p^m) > 0$, then we can define the Lagrange interpolation operator on the nodal values of v. However, if $\mathrm{sob}(W_p^m) \le 0$ the construction of an interpolation operator with good stability and approximability properties is much less obvious. Hereafter, we resort to the so-called *Clement interpolation operator* I_h, which is defined by local averaging at the element level [15]. What is relevant for us of I_h is the following result.

Proposition 1.6 (Clement interpolation operator). *Let $k \ge 1$ be the polynomial degree. Then for all $v \in W_p^m(\Omega)$ and $T \in \mathcal{T}_h$, we have*

$$\|D^t(v - I_h v)\|_{L^p(T)} \le C h_T^{s-t} \|D^s v\|_{L^p(N(T))}, \tag{1.22}$$

where $0 \le t \le s \le \min(m, 1+k)$ and $N(T)$ is a discrete neighborhood of T that includes all elements of $\mathcal{T}_h$ with nonempty intersection with T. In addition, if $m \ge 1$ and the trace of v is zero, then I_h can also be defined with vanishing trace and (1.22) *is valid up to the boundary.*

We point out that, upon taking $s = t = 0$ we readily see that I_h is stable in L^p, namely,

$$\|I_h v\|_{L^p(T)} \le C \|D^s v\|_{L^p(N(T))}.$$

Another interplation operator, due to Scott and Zhang [8], is able to deal with more general boundary values but requires more regularity than mere membership in $L^p(N(T))$ to be defined. In this course we will use the Clement operator exclusively. A simple global consequence of (1.22) can be written as

$$\|h^{-1}(v - I_h v)\|_{L^p(\Omega)} \le C \|Dv\|_{L^p(\Omega)}. \tag{1.23}$$

Suppose now that we have a function $f \in L^2(\Omega)$ that is orthogonal to all continuous piecewise linear functions over a shape-regular mesh $\mathcal{T}_h$ that vanish at the boundary. Then, we can estimate its norm in $H^{-1}(\Omega)$ as

$$\|f\|_{H^{-1}(\Omega)} \le C \|hf\|_{L^2(\Omega)}, \tag{1.24}$$

because, in view of (1.23) for $p = 2$,

$$\langle f, v\rangle = \langle f, v - I_h v\rangle \le \|hf\|_{L^2(\Omega)} \|h^{-1}(v - I_h v)\|_{L^2(\Omega)} \le C \|hf\|_{L^2(\Omega)} \|Dv\|_{L^2(\Omega)}.$$

This will be instrumental for our discussion of sections 2.1.1 and 2.1.2.

We consider now a couple of examples in 2D ($d = 2$) with polynomial degree $k = 1$.

Example 1.7 (Smooth Solution)**.** Suppose $u \in H^2(\Omega)$, that is u has two weak derivatives in $L^2(\Omega)$. Suppose $\mathcal{T}_h$ is a quasi-uniform mesh with meshsize h, namely $h_T \approx h$ for all $T \in \mathcal{T}_h$. The *number of degrees of freedom* N in $\mathbb{R}^2$ is then

$$N \approx h^{-2}.$$

Therefore, using Lemma 1.4 and (1.6), we see that

$$\|u - u_h\|_{\mathbb{V}} \le Ch\|D^2u\|_{L^2(\Omega)} = CN^{-1/2} \tag{1.25}$$

is the optimal decay achievable with piecewise linear finite elements for smooth functions.

Example 1.8 (Rough Solution)**.** Suppose now that $\Omega \subset \mathbb{R}^2$ is a polygonal domain with a reentrant corner $\omega > \pi$, and let $\gamma = \pi/\omega < 1$. In view of Example 1.1, the solution u behaves like

$$u(r, \theta) = r^\gamma \phi(\theta)$$

in polar coordinates close to the corner. If $D^{1+s}u$ denotes a (formal) fractional derivative of u with $s < 1$, we note that $D^{1+s}u \in L^2(\Omega)$ provided

$$s < \gamma \ (< 1),$$

because $\int_0^1 r^{2(\gamma-s-1)+1} dr < \infty$ in this case. Therefore, Lemma 1.4 and (1.6) imply

$$\|u - u_h\|_{\mathbb{V}} \le Ch^s\|D^{s+1}u\|_{L^2(\Omega)} = CN^{s/2}. \tag{1.26}$$

This estimate is *suboptimal* because $s < 1$. However, this is due to lack of regularity of u and not inability of the FEM to approximate u. We may thus wonder, as we did in section 1.1, whether we can still achieve (1.25) by properly designing the mesh $\mathcal{T}_h$.

We would like to minimize the interpolation error $\|uI_hu\|_{\mathbb{V}}$ for a prescribed number of degrees of freedom N. This can be formulated as the constrained minimization problem

$$\min_{\int_\Omega h^{-2}dx = N} \int_\Omega h^2|D^2u|^2 dx.$$

The optimality condition is

$$h^4|D^2u|^2 = \text{constant}. \tag{1.27}$$

Since $|T| \approx h_T^2$, this means that the contribution to to the error per triangle, namely

$$h_T^2|D^2u|^2|T| \approx \int_T h^2|D^2u|^2$$

must be constant, say Λ. If $T \in \mathcal{T}_h$ is a triangle at distance r_T to the corner (say the origin), then

$$\Lambda = \int_T h^2|D^2u|^2 \approx h_T^4 r_T^{2(\gamma-2)}.$$

This implies

$$h_T = \Lambda^{\frac{1}{4}} r_T^{1-\frac{\gamma}{2}},$$

which corresponds to amesh *graded* towards the origin. Moreover,

$$N = \int_\Omega h^{-2}dx = \Lambda^{-\frac{1}{2}} \int_0^1 r^{-1+\gamma}dr = C\Lambda^{-\frac{1}{2}},$$

whence

$$h_T = Cr_T^{1-\frac{\gamma}{2}} N^{-\frac{1}{2}}.$$

Consequently, we can still recover the optimal bound (1.25)

$$\|u - u_h\|_{\mathbb{V}}^2 = \int_\Omega h^2|D^2u|^2dx = \Lambda \int_\Omega h^{-2}dx = \Lambda N = CN^{-1}$$

with the same number of degrees of freedom N. This clearly shows that it is possible to compensate for singularities a priori provided that we know the solution local behavior. We would like, however, to devise a technique that would be able to perform the above minimization without our knowledge of the regulalrity of u. This way we could extend the technique to more complex situations. This is the purpose of sections 2.1.1 and 2.1.2.

Example 1.9 (Error Equidistribution). We point out that in Examples 1.1, 1.2, and 1.3 the solution $u \in W_1^2(\Omega)\backslash H^2(\Omega)$. This is a typical situation for elliptic PDE in 2D. We now assume, as suggested by Example 1.8, that the error in H^1 is equidistributed in $\mathcal{T}_h$, namely,

$$\|u - I_h u\|_{H^1(T)} = \Lambda \qquad \forall\, T \in \mathcal{T}_h$$

where Λ is a constant. To find the value of Λ we resort to the relation between $H^1(T)$ and $W_1^2(T)$. Since these spaces have the same Sobolev number, that is $\mathrm{sob}(H^1) = 1-2/2 = 0 = 2-2/1 = \mathrm{sob}(W_1^2)$, the scaling of norms does not have any power of the meshsize h_T; hence

$$\|u - I_h u\|_{H^1(T)} \le C\|D^2(u - I_h u)\|_{L^1(T)} = C\|D^2 u\|_{L^1(T)}.$$

Therefore

$$N\Lambda^2 = \sum_{T\in\mathcal{T}_h} \|u - I_h u\|_{H^1(T)}^2 \le C\Lambda \sum_{T\in\mathcal{T}_h} \|D^2 u\|_{L^1(T)} = C\Lambda\|D^2 u\|_{L^1(\Omega)},$$

which implies $\Lambda \le CN^{-1}\|D^2 u\|_{L^1(\Omega)}$ as well as

$$\|u - I_h u\|_{H^1(\Omega)} \le CN^{-\frac{1}{2}}\|D^2 u\|_{L^1(\Omega)}.$$

We see again the trade-off of differentiability and integrability alluded to before in section 1.1; see [20].

1.6. **Numerical Experiments.** Here we present numerical experiments on uniform meshes for the Dirichlet problem $-\Delta u = f$ with exact solution (in polar coordinates)

$$u(r,\theta) = r^{\frac{2}{3}}\sin(2\theta/3) - r^2/4,$$

on an L-shaped domain Ω (see Example 1.1). In Figure 1.6 we depict the sequence of *uniform* meshes, and in Table 1.1 we report the order of convergence for polynomial degree $k = 1,2,3$. The asymptotic rate is about $s = 2/3$ regardless of k and is consistent with the estimate (1.26). We will show numerical experiments for *graded* meshes along with optimal rates in section 2.6.

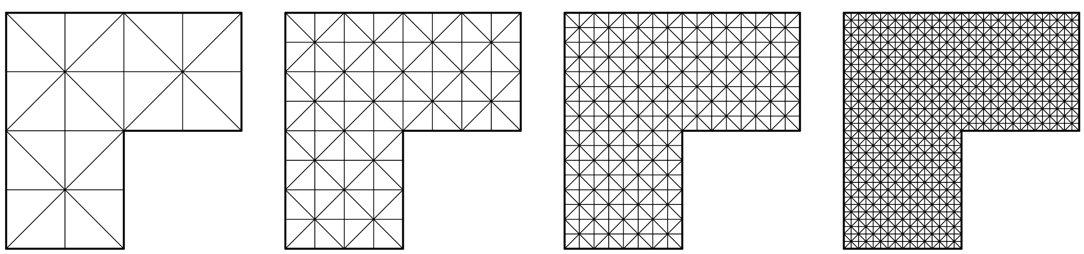

FIGURE 1.2. Sequence of uniform meshes for L-shaped domain Ω

h	linear ($k=1$)	quadratic ($k=2$)	cubic ($k=3$)
1/4	1.14	9.64	9.89
1/8	0.74	0.67	0.67
1/16	0.68	0.67	0.67
1/32	0.66	0.67	0.67
1/64	0.66	0.67	0.67
1/128	0.66	0.67	0.67

TABLE 1.1. The asymptotic rate of convergence is about $s = 2/3$ irrespective of the polynomial degree k as predicted by (1.26).

1.7. **Exercises.**

1.7.1. *Exercise: Third Boundary Value Problem.* Find a variational formulation which amounts to solving

$$-\Delta u = f \quad \text{in } \Omega, \qquad \partial_\nu u + pu = g \quad \text{on } \partial\Omega,$$

where $f \in L^2(\Omega), g \in L^2(\partial\Omega)$ and $0 < P_1 \leq p \leq P_2$ on $\partial\Omega$.
(a) Show that Lax-Milgram theorem applies and conclude that there exists a unique solution $u \in H^1(\Omega)$.
(b) Suppose that $p = \epsilon^{-1} \to \infty$. What it the boundary value problem satisfied by $u_0 = \lim_{\epsilon \to 0} u_\epsilon$? Can you derive an error estimate for $\|u_\epsilon - u_0\|_{H^1(\Omega)}$?

1.7.2. *Exercise: Minimization and Euler-Lagrange Equations.* Suupose that $\mathbf{b} = 0$ in (1.3) and so that the bilinear form $\mathcal{B}$ is symmetric. Show that the weak solution u is also the minimizer of an energy $I(v)$ defined over $H_0^1(\Omega)$. Show that the discrete solution $u_h \in \mathbb{V}_h$ is also a minimizer

1.7.3. *Exercise: Scaling.* Show (1.20) and (1.21).

1.7.4. *Exercise: Clement Interpolant.* Let $\omega_z = \operatorname{supp}(\phi_z)$ be the support of a piecewise linear basis function ϕ_z where $z \in \mathcal{N}_h$ is a node. Let $I_h : L^1(\Omega) \to \mathbb{V}_h$ be defined as follows:

$$v_z = \frac{1}{|\omega_z|} \int_{\omega_z} v, \quad \forall z \in \mathcal{N}_h, \qquad I_h v = \sum_{z \in \mathcal{N}_h} v_z \phi_z.$$

(a) Show that I_h is 1st order accurate in all $W_p^1(\Omega)$ for all $1 \leq p \leq \infty$.
(b) Show that I_h is not second order accurate in any $W_p^2(\Omega)$ for any $1 \leq p \leq \infty$.

1.7.5. *Exercise: Optimality Condition.* Prove (1.27).

1.7.6. *Exercise: Pointwise Values.* Let ρ be a smooth function defined for $0 < r \leq 1$ satisfying

$$\int_0^1 |\rho'(r)| r^{n-1} dr < \infty.$$

Define f on $\Omega = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| < 1\}$ via $f(x) = \rho(|\mathbf{x}|)$ for $n \geq 2$. Show that the weak derivative $D^\alpha f$ for $|\alpha| = 1$ is given by

$$g(\mathbf{x}) = \rho'(|\mathbf{x}|) \frac{\mathbf{x}^\alpha}{|\mathbf{x}|}.$$

Use this to prove that $f(\mathbf{x}) = \log \big| \log |\mathbf{x}/2| \big| \in W_p^1(\Omega)$ for all $p \leq n$. This shows that $W_n^1(\Omega)$ in NOT contained in $L^\infty(\Omega)$ for $n \geq 2$.

1.7.7. *Exercise: Equivalent Norms (Deni-Lions).* Consider the Sobolev space $W_p^{k+1}(\Omega)$ with $k \geq 0, 1 \leq p \leq \infty$ and a Lipschitz domain Ω in $\mathbb{R}^d$. Let $\{f_i\}_{i=1}^N$ be linear continuous functionals in $W_p^{k+1}(\Omega)$ such that for any polynomial $v \in \mathbb{P}_k$ of degree $\leq k$:

$$f_i(v) = 0 \quad \forall 1 \leq i \leq N \qquad \Longleftrightarrow \qquad v = 0.$$

Show that $\|v\|_{W_p^{k+1}(\Omega)}$ is equivalent to the seminorm

$$|v|_{W_p^{k+1}(\Omega)} + \sum_{i=1}^N |f_i(v)|.$$

Hint: Proceed by contradiction assuming that there is a sequence $\{v_n\} \subset W_p^{k+1}(\Omega)$ such that $\|v_n\|_{W_p^{k+1}(\Omega)} = 1$ but the latter seminorm tends to 0. Use that $W_p^{k+1}(\Omega)$ is compactly imbedded in $W_p^k(\Omega)$ (Rellich Theorem), namely that each bounded sequence in $W_p^{k+1}(\Omega)$ admits a convergence subsequence in $W_p^k(\Omega)$.