Notes for Math 663
Spring 2016

Alan Demlow

Chapter 1

Introduction and preliminaries

1.1 Introduction

In this chapter we will introduce the main topics of the course. We first briefly define the finite element method and state a basic error estimate.

Let $\Omega \subset \mathbb{R}^d$ ($d = 2, 3$) be a polygonal or polyhedral domain, and let $f \in L_2(\Omega)$. Consider the boundary value problem

$$\begin{aligned} -\Delta u &= f \text{ in } \Omega, \\ u &= 0 \text{ on } \partial\Omega. \end{aligned}$$

Rewriting in weak form, we seek $u \in H_0^1(\Omega)$ such that

$$a(u, v) := \int_\Omega \nabla u \nabla v \, \mathrm{d}x = \int_\Omega f v \, \mathrm{d}x =: (f, v), \quad v \in H_0^1(\Omega).$$

Here we define

$$H^m(\Omega) = \{u : D^\alpha u \in L_2(\Omega), \text{ all multiindices } \alpha \text{ with } |\alpha| \leq m\}$$

and

$$H_0^1(\Omega) = \{u \in H^1(\Omega) : \operatorname{tr} u = 0 \text{ on } \partial\Omega\}.$$

Here tr is the trace operator.

In order to define a finite element method, let $\mathcal{T}_h$ be a *regular* simplicial decomposition of Ω. That is, $\Omega = \cup_{T \in \mathcal{T}_h} T$, and the intersection of any two members of $\mathcal{T}_h$ is either empty or consists of a common lower-dimensional simplicial boundary component of each member (i.e., a face, edge, or vertex if $d = 3$, or an edge or vertex if $d = 2$). We additionally assume that for each $T \in \mathcal{T}_h$, there are balls $b_T \subset T \subset B_T$ such that $\operatorname{diam} b_T \simeq \operatorname{diam} B_T \simeq \operatorname{diam} T$. Here we use the shorthand $a \simeq b$ to mean that $a \leq c_1 b \leq c_2 a$ with constants c_1, c_2 independent of essential quantities. For the time being, we also assume that $\mathcal{T}_h$ is *quasi-uniform*, that is, $\operatorname{diam} T \simeq \operatorname{diam} T'$ for all $T, T' \in \mathcal{T}_h$.

Next let $V_h = \{u \in C(\Omega) \cap H_0^1(\Omega) : u|_T \in \mathbb{P}^r\}$ for some fixed $r \geq 1$. $r = 1$ corresponds to the case of piecewise linear finite element spaces, $r = 2$ to piecewise quadratics, etc. We then seek $u_h \in V_h$ such that

$$a(u_h, v_h) = (f, v_h), \quad v_h \in V_h.$$

Due to the coercivity of the form $a(\cdot,\cdot)$ on $H_0^1(\Omega)$, we may apply Céa's Lemma to find that

$$\|u-u_h\|_{H_0^1(\Omega)} = \inf_{\chi \in V_h} \|u-\chi\|_{H_0^1(\Omega)}.$$

Application of a suitable interpolation operator then implies that

$$\inf_{\chi \in V_h} \|u-\chi\|_{H_0^1(\Omega)} \leq Ch^r |u|_{H^{r+1}(\Omega)},$$

assuming that $u \in H^{r+1}(\Omega)$. That is, the finite element error decreases with order h^r:

$$\|u-u_h\|_{H^1(\Omega)} \leq Ch^r |u|_{H^{r+1}(\Omega)}.$$

From approximation theory we can conclude that this is the best possible order of convergence when using Lagrange elements. This error estimates thus tells us that in a certain (rather basic) sense, we are getting what we paid for out of our finite element method–*assuming that u is sufficiently smooth*.

The question naturally arises, then, whether the assumption $u \in H^{r+1}(\Omega)$ is realistic. If the boundary $\partial\Omega$ is smooth, then a *shift theorem* holds:

$$\|u\|_{H^{m+2}(\Omega)} \leq C\|\Delta u\|_{H^m(\Omega)}, \ m \geq -1, \tag{1.1}$$

whenever $-\Delta u = f \in H^m(\Omega)$. When $\partial\Omega$ is not smooth, such an estimate does not hold in general. We then may ask some natural questions: For what combinations of domains Ω, parameters r, and data functions f can we expect that $u \in H^{r+1}(\Omega)$? What happens if $u \notin H^{r+1}(\Omega)$? Can we somehow fix the error estimate or our method to ensure that we can still get optimal-order error decrease from our method in this situation?

We also recall a standard L_2 error estimate and outline its proof. Assume now that (1.1) holds with $m = 0$, and let $z \in H_0^1(\Omega)$ solve

$$a(v, z) = (v, u-u_h), \quad v \in H_0^1(\Omega).$$

Then for any $z_h \in V_h$,

$$\|u-u_h\|^2_{L_2(\Omega)} = a(u-u_h, z) = a(u-u_h, z-z_h),$$

since $a(u,\chi) = (f,\chi) = a(u_h,\chi)$ and so $a(u-u_h,\chi) = 0$, $\chi \in V_h$ (Galerkin orthogonality). Then by using Cauchy-Schwarz, choosing z_h so that $\|z-z_h\|_{H^1(\Omega)} \leq Ch|z|_{H^2(\Omega)}$, and employing (1.1), we have

$$\|u-u_h\|^2_{L_2(\Omega)} \leq \|u-u_h\|_{H^1(\Omega)}\|z-z_h\|_{H^1(\Omega)} \leq Ch\|u-u_h\||z|_{H^2(\Omega)} \leq Ch\|u-u_h\|^2_{L_2(\Omega)}.$$

Thus assuming that $u \in H^{r+1}(\Omega)$, we have

$$\|u-u_h\|_{L_2(\Omega)} \leq Ch\|u-u_h\|_{H^1(\Omega)} \leq Ch^{r+1}|u|_{H^{r+1}(\Omega)}.$$

We note that (1.1) holds for $m = 0$ whenever Ω is *convex*, and in particular when Ω is convex and polygonal. However, it does not generally hold for $m > 0$ even if Ω is convex, though there are some convex polygonal domains for which (1.1) does hold for $m = 1$, such as squares and rectangles.

If Ω is a non-convex polygonal or polyhedral domain, then (1.1) *never* holds for $m = 0$. Thus the above L_2 error estimate can not be assumed to be valid in such cases as the above poof technique does not apply, and indeed a reduction in order of convergence generally can be observed in such cases.

In this course, we will study finite element error behavior in the presence of singularities, which are points where the solution u or one of its derivatives fails to be defined. In such cases the range of values of r for which $u \in H^{r+1}(\Omega)$ as required above is generally restricted. Our main focus will be singularities caused by non-smoothness of the boundary $\partial\Omega$, and in particular on polygonal and polyhedral domains Ω. As we noted above, convexity or nonconvexity of the domain already has a pronounced effect on the regularity of solutions to Poisson's problems. The techniques we use will allow us to make finer distinctions on the smoothness of such solutions based on geometric properties of the domain. In addition, these regularity properties in turn affect the provable convergence rates of the FEM in the energy norm as well as other norms such as the L_2 norm. We will delve further into this issue as well.

The twin tasks of understanding error behavior and devising ways to improve numerical performance in the presence of such singularities have drawn attention from finite element researchers since at least the 1970's. Early efforts in this direction tended to focus on a priori error estimation and manual construction of graded (locally refined) meshes which could appropriately resolve singularities having known properties. While helpful in understanding basic method and error behavior, this approach is practically limited. For linear problems, singularity locations can generally be predicted based on problem data such as f and elliptic coefficients. However, the singularity strength can depend on these data, and even for the Laplacian it is somewhat involved to meaningfully predict singularity properties strengths especially in three space dimensions. For nonlinear problems, the situation is even more complex as singularity location also cannot generally be predicted a priori.

For these reasons, self-adaptive algorithms are a popular tool for efficiently resolving singularities and other significant local variations in the solution u. These algorithms automatically detect singularities and direct efficient local mesh refinement. Adaptive finite element methods (AFEM) have also been in existence for many years (at least since the 1980's). Meaningful convergence analysis of such algorithms is however relatively recent, with a mature theory for basic elliptic problems occurring only within the past decade. Convergence theory for adaptive finite element algorithms continues to be a very active area of research.

The main goal for the course, broadly speaking, is to understand the interplay between regularity of solutions to boundary value problems and convergence rates of finite element methods. More specifically, our goals are to:

1. Understand the basic structure of singularities arising in elliptic boundary value problems on polygonal and polyhedral domains.

2. Learn various ways of measuring the smoothness (regularity) of solutions to PDE. Examples may include W_p^m (Sobolev) spaces, fractional-order Sobolev spaces, weighted Sobolev spaces, Besov spaces, and Hölder spaces.

3. Gain a brief understanding of a priori mesh grading techniques for resolving singularities, and the effects of systematic grading on a priori error estimates.

4. Understand various notions of optimality of finite element methods.

5. Gain an in-depth understanding of adaptive convergence theory and the ability of adaptive methods to resolve singularities.

6. As time and interests of the audience allow, study finite element error behavior (a priori and a posteriori) in nonstandard norms, including local energy norms and the maximum norm.

There are a number of extensions of these topics that we may either cover briefly or which would make good student projects. These include convergence of adaptive algorithms for other types of elliptic problems and FEM (Stokes', Maxwell's equations, elliptic equations in mixed form), convergence of adaptive *hp*-FEM, singularities and adaptivity for eigenvalue problems, anisotropic mesh refinement, and others.

1.2 Computational prologue

In order to illustrate some of the topics mentioned above, we return to the finite element error estimate

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch^r |u|_{H^{r+1}(\Omega)}. \tag{1.2}$$

when the finite element method uses polynomials of degree r. This estimate seems to imply that it is more efficient to use higher polynomial degrees of finite element spaces. To test this assumption, we discuss what actually happens in computational practice.

Also, We first reorient our definition of order of convergence by expressing it with respect to the number of degrees of freedom in the finite element space instead of h. Assume that we have a quasi-uniform grid. Shape-regularity implies that each $T \in \mathcal{T}_h$ has volume equivalent to h^d. Thus the number of elements in the grid is equivalent (up to some fixed constant) to volume$(\Omega)/h^d \simeq h^{-d}$. Let DOF ("degrees of freedom") be the size of the finite element system. We then have for quasi-uniform grids that

$$DOF \simeq h^{-d}, \quad h \simeq DOF^{-1/d}.$$

On a quasi-uniform grid, (1.2) then translates to

$$\|u - u_h\|_{H^1(\Omega)} \lesssim DOF^{-r/d} |u|_{H^{r+1}(\Omega)}. \tag{1.3}$$

By $a \lesssim b$ we mean $a \leq Cb$ for some nonessential constant C, and by $a \simeq b$ we mean that $a \lesssim b$ and $b \lesssim a$. The advantage of expressing rates of convergence using numbers of degrees of freedom instead of h is that this notion applies and is meaningful even if the grid is not quasi-uniform. The number of degrees of freedom is also a reasonable measure of the computational cost of the finite element algorithm, since with proper solver technology the overall cost of the algorithm increases roughly linearly with DOF.

We take the right-hand-side data (forcing function) to be $f(x) = 1$ and solve $-\Delta u = f$, $u = 0$ on $\partial\Omega$ for a few different choices of Ω. Our actual computations mostly assume $r = 1$ (piecewise linear elements), but we also discuss what happens for different choices of r. We don't know u in these computations, but we can still measure $\|u - u_h\|_{H^1(\Omega)}$ by employing an a posteriori error estimator η which satisfies $\eta \simeq \|u - u_h\|_{H^1(\Omega)}$. (More precisely, we use a residual-type error estimator. As we will learn later in the course, for general f we have equivalence between error and estimator only up to *data oscillation*, but for $f \equiv 1$ data oscillation is zero and so the estimator is always equivalent to the error.)

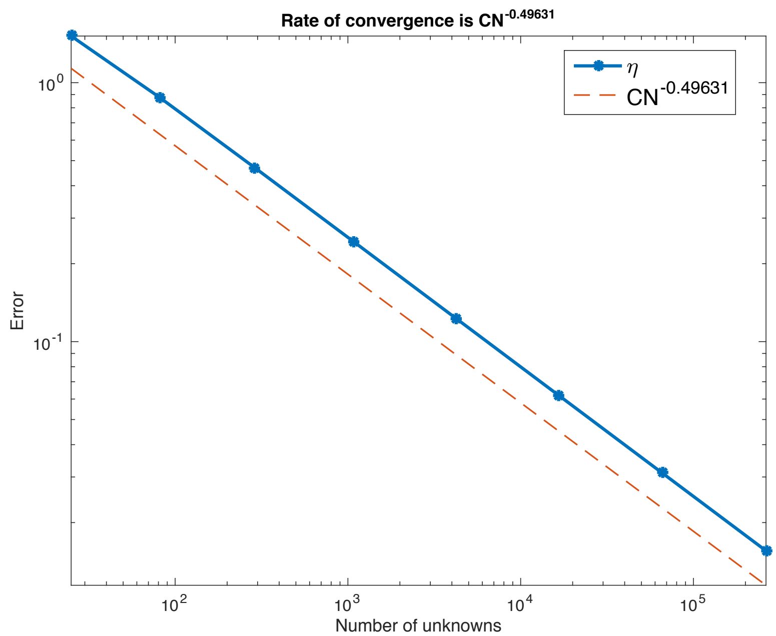

Figure 1.1: Error reduction on a square domain.

First we investigate convergence on the square $\Omega = (-1,1) \times (-1,1)$. This is a convex domain, so we have $u \in H^2(\Omega)$ and $\|u\|_{H^2(\Omega)} \lesssim \|f\|_{L_2(\Omega)} \lesssim 1$. We then expect that for a uniform sequence of mesh refinements,

$$\|u - u_h\|_{H^1(\Omega)} \lesssim h|u|_{H^2(\Omega)} \simeq DOF^{-1/2},$$

since $h \simeq DOF^{-r/d} = DOF^{-1/2}$. Computations are performed using the MATLAB toolbox iFEM reference?, and rates of convergence are estimated using a log-least squares fit of the estimator. The estimated rate of convergence (which is very close to the theoretical one of 0.5) is plotted in Figure 3.1 and one of the computational meshes and the solution in Figure 1.2.

We next work on an "L-shaped domain", realized here as $\Omega = (-1,1) \times (-1,1) \setminus [0,1) \times (-1,0]$. Because the domain is not convex, we don't expect the H^2 regularity result $\|u\|_{H^2(\Omega)} \lesssim \|f\|_{L_2(\Omega)}$ to hold, and in fact it does not. At least our proof that $\|u - u_h\|_{H^1(\Omega)} \lesssim DOF^{-1/2}$ does not hold here. In Figure 1.3 we display the calculated rate of convergence on a quasi-uniform mesh. There we obtain roughly that $\|u - u_h\|_{H^1(\Omega)} \lesssim DOF^{-.37}$, More careful numerical calculation (or theory) would show a convergence rate of $DOF^{-1/3}$ for sufficiently large DOF (note that the observed convergence line is slightly more shallow than the reference line for high DOF in Figure 1.3). This is a worse convergence rate than we experienced when computing on the unit square, so in some sense we aren't getting what we paid for out of our piecewise linear elements.

We finally perform the same computation on the "crack domain", which is a square with a slit removed (in this case the square with vertices $(1,0)$, $(0,1)$, $(-1,0)$, and $(0,-1)$, minus the nonnegative real axis). Figure 1.5 and **??** for the observed convergence rate and solution. The observed convergence rate is about $DOF^{-.262}$, which is expected to settle down to $DOF^{-.25}$ for

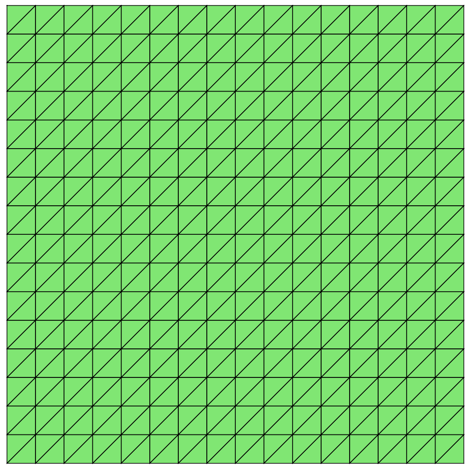

Figure 1.2: Mesh and solution on square domain.

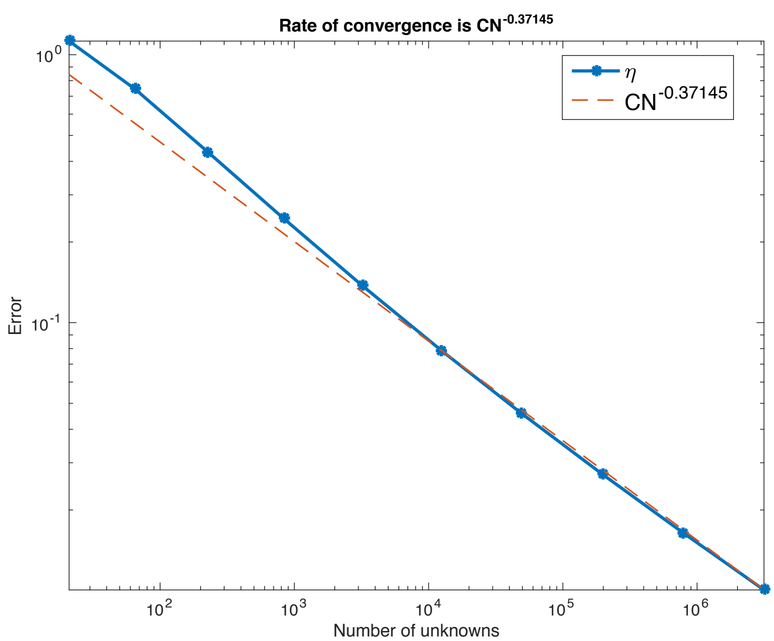

Figure 1.3: Error reduction on an L-shaped domain.

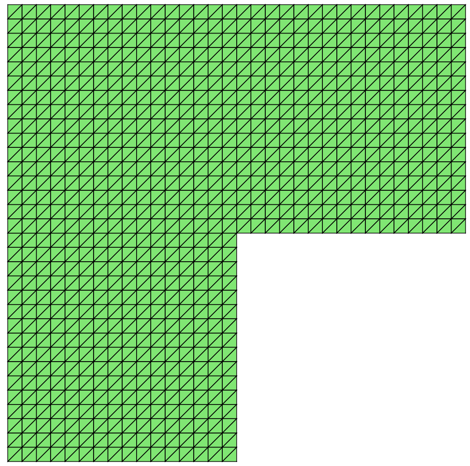

Figure 1.4: Mesh and solution on an L-shaped domain.

sufficiently large DOF. Thus we see a reduction in convergence rate from $DOF^{-.5}$ to $DOF^{-.33}$ to $DOF^{-.25}$ as we move from the square to the L-shaped to the crack domain. This indicates a reduction in the regularity of the boundary value problem as we make the domain less convex, since all other problem parameters are the same here. We note also that while (1.3) indicates that we might expect a better rate of convergence if we increase the polynomial degree, that will not happen at all on the L-shaped or crack domains if we use only quasi-uniform meshes (it will at least to an extent on the square domain).

Because quasi-uniform meshes don't give us what we have paid for on non-convex domains even for piecewise linear elements, we now consider the use of solution-adapted meshes. That is, we will locally refine (grade) the mesh in such a way that the solution is resolved more efficiently. Here we use an adaptive finite element method to accomplish this, but in these simple two-dimensional cases essentially the same results can be obtained using manual (a-priori) mesh grading. We run our adaptive method on the unit square, L-shaped domain, and crack domain as above. On the unit square (Figure 1.6 and 1.7), we obtain optimal-order convergence rate as with uniform mesh refinement. There is moderate grading in the solution-adapted mesh, and perhaps a slight reduction in the error reached for a given number of degrees of freedom, but generating a solution-adapted mesh is not worth the trouble in this particular case. (It might be for the same problem if one wished to use higher polynomial degrees.) On the L-shaped and crack domains, however, we once again achieve a rate of convergence of about $DOF^{-.5}$. Thus using solution-adapted meshes lets us get what we paid for from our algorithm. The obtained meshes (Figures 1.9 and Figure 1.11) show that mesh elements are concentrated near the reentrant (non-convex) corner of both of the domains. This occurs because strong singularities are generated there. They are the reason that $u \notin H^2(\Omega)$ for these problems and the use of solution-adapted meshes is thus necessary to recover optimal convergence rates. We note also that by using adaptivity, for these two-dimensional problems one can always recover the energy-norm convergence rate $DOF^{-r/2}$.

The situation in three space dimensions is somewhat different. In two dimensions, we indicated above that adaptivity generally allows us to recover the "optimal" convergence rate $DOF^{-r/d} =$

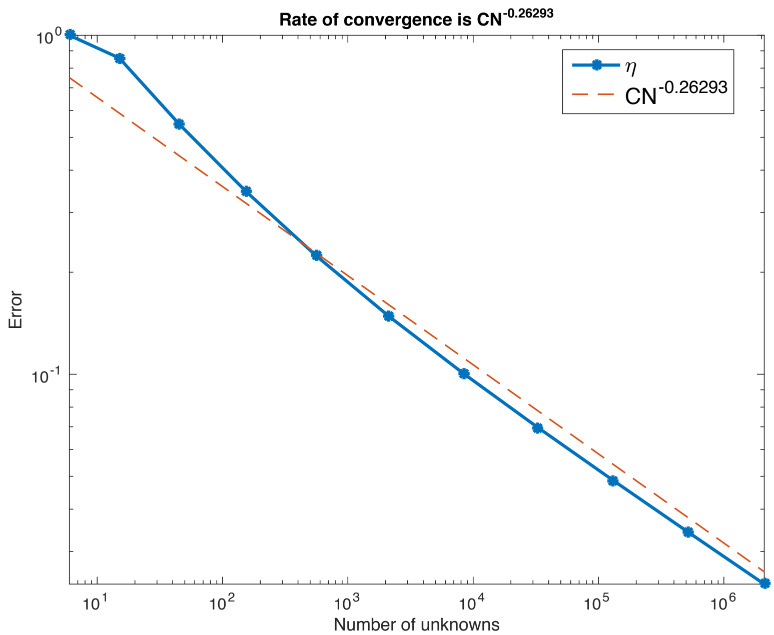

Figure 1.5: Error reduction on a crack domain.

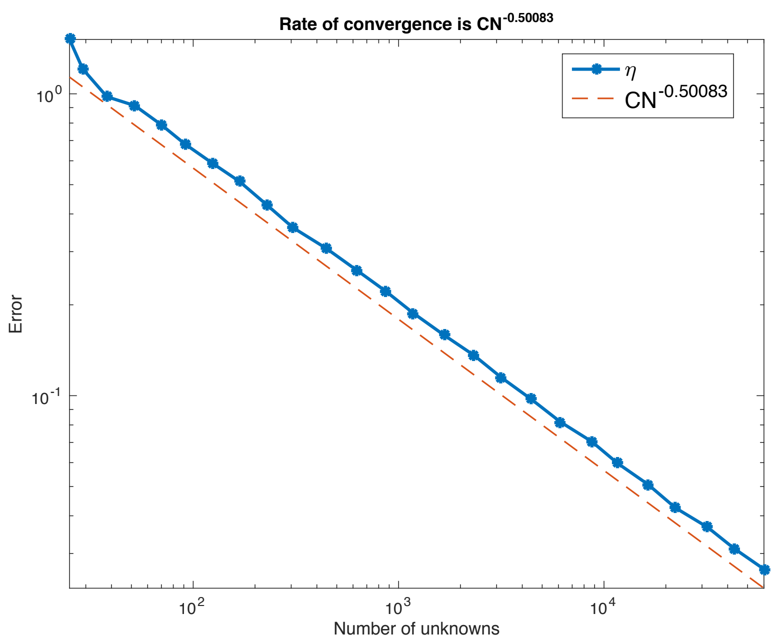

Figure 1.6: Adaptive error reduction on a square domain.

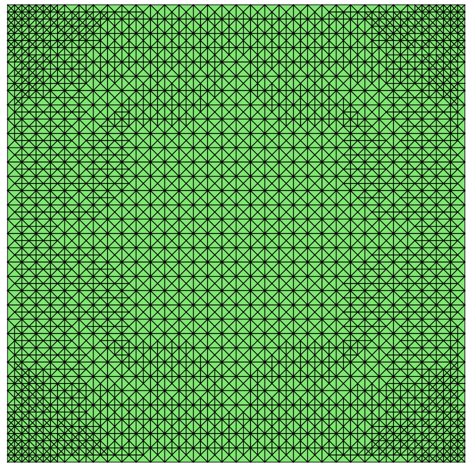

Figure 1.7: Adaptive mesh on a square domain.

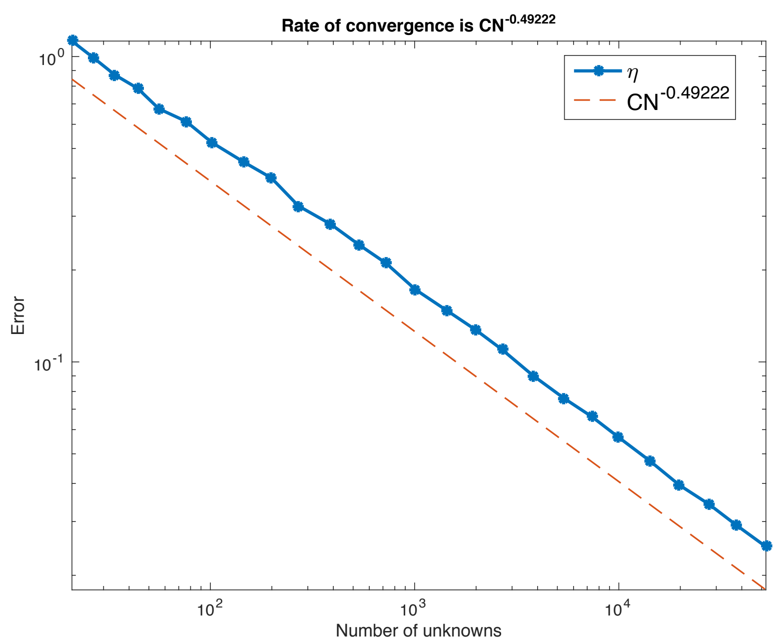

Figure 1.8: Adaptive error reduction on an L-shaped domain.

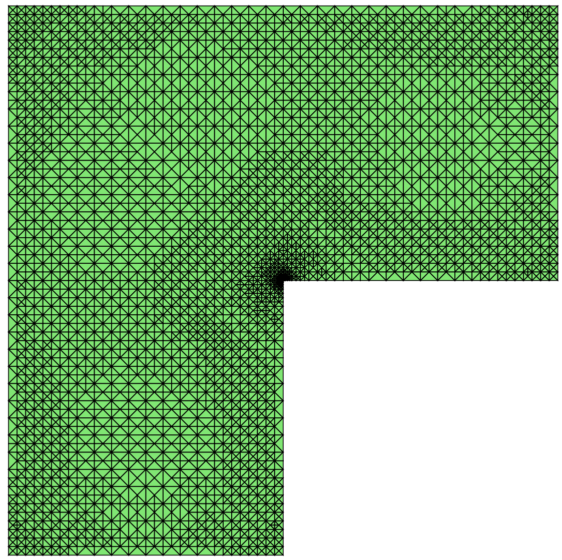

Figure 1.9: Adaptive mesh on an L-shaped domain.

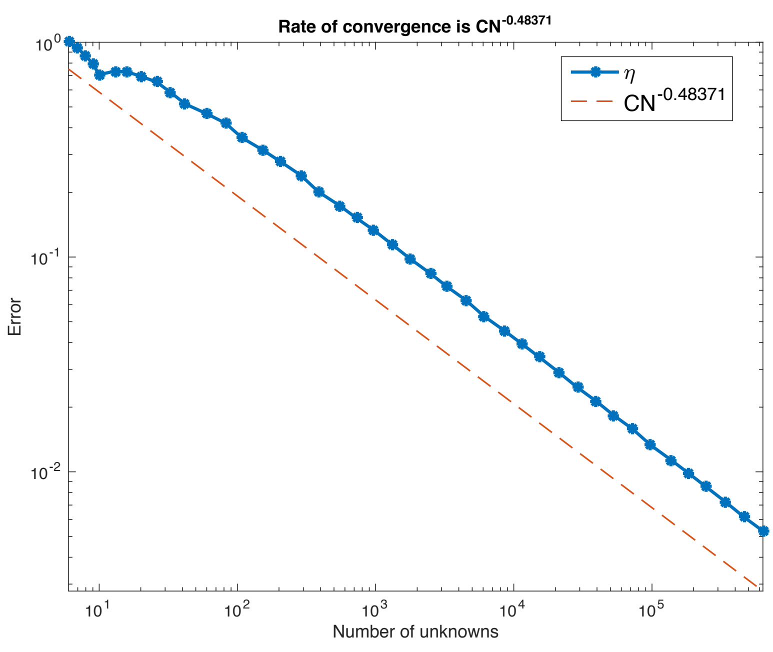

Figure 1.10: Adaptive error reduction on a crack domain.

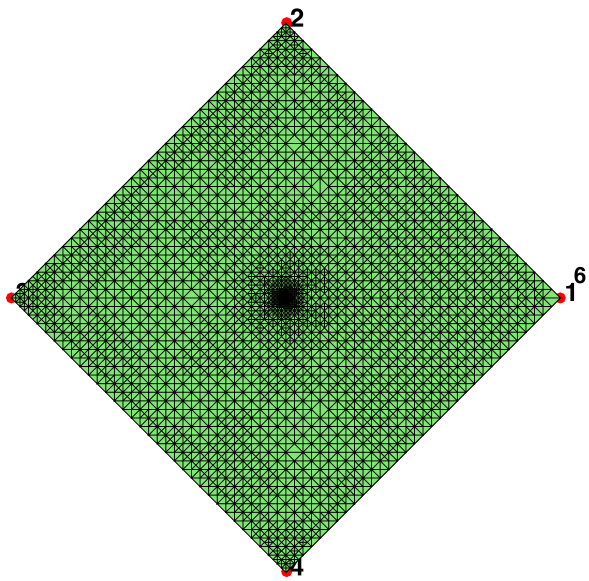

Figure 1.11: Adaptive mesh on a crack domain.

$DOF^{-r/2}$. To test whether this is also true in 3D, consider the domain Ω pictured in Figure 1.12. This is a convex polyhedral domain, so for the case $r = 1$ we expect optimal-order convergence (now $DOF^{-1/3}$). For $r = 2$ and $r = 3$, we can also recover optimal-order convergence rates $DOF^{-2/3}$ and DOF^{-1} by using adaptivity. However, for $r = 4$ (quartic elements), even with adaptive meshes we only recover a convergence rate of about $DOF^{-8/7}$ (the line "H1refinement/H1 estimator", having logarithmic slope -1.12). Using adaptivity for any $r > 4$ will lead to similar results. On non convex domains a similar barrier in convergence rates will be reached for lower polynomial degrees. In summary, we make the following observations:

1. In general for polygonal and polyhedral domains, he more highly non-convex the domain is, the worse the convergence rate on quasi-uniform meshes.

2. When solving Poisson's problem on 2D polygonal domains with smooth f, we can always recover optimal convergence rates by using solution-adapted meshes.

3. Adaptivity helps in three space dimensions, but does not always allow us to recover optimal convergence rates.

Much of the course will be spent explaining these observations. That is, we ask and answer the following questions:

1. What causes increasing reduction in convergence rates on quasi-uniform meshes for increasingly non-convex domains? (We answer this question by explaining the singular structure of solutions on polyhedral domains.)

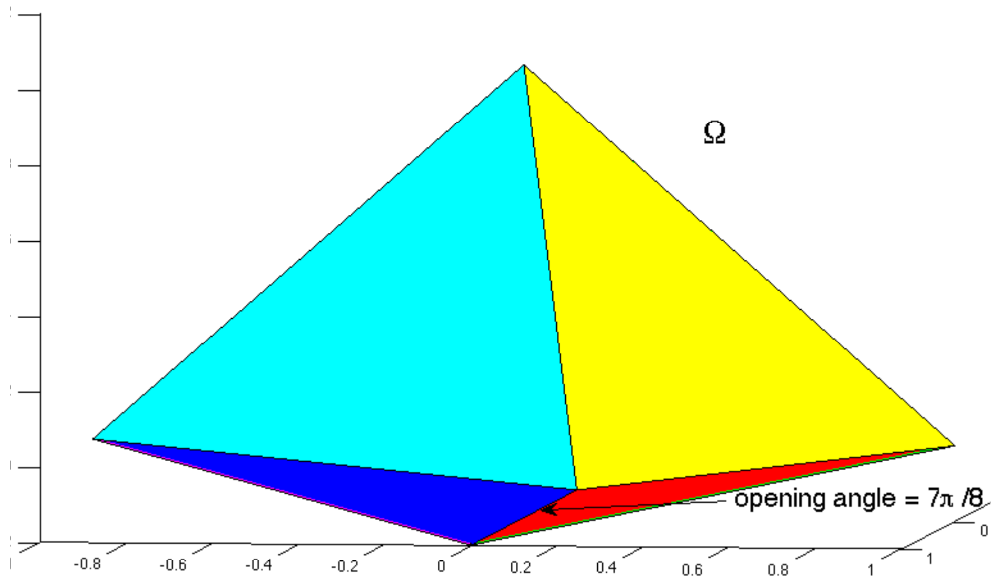

Figure 1.12: 3D convex polyhedral domain.

2. How do we adapt meshes to efficiently resolve singular solutions? (We will give two different answers to this question, one involving a priori "by hand" mesh gradings and the other involving automatic adaptivity.)

3. What's the difference between two and three space dimensions? Why can we generally recover optimal convergence rates in 2D by using adaptivity, but not always in 3D? (The answer requires a good understanding of the involved singularities in each case, and the notion of *anisotropic functions* which which have different regularity in different directions.)

1.3 Analytical preliminaries

1.3.1 Domain properties

Classification of domains plays an important role in analyzing boundary value problems and numerical methods for solving them. By *domain* we mean a nonempty open set in $\mathbb{R}^d$. We generally assume that Ω is in addition bounded.

We first define C^k domains. We say that the boundary $\partial\Omega$ of Ω is C^k if for each $x_0 \in \partial\Omega$ there exists $r > 0$ and a C^k function $\gamma : \mathbb{R}^{d-1} \to \mathbb{R}^d$ such that (upon reorienting the coordinate axes if necessary)

$$\Omega \cap B_r(x_0) = \{x \in B_{x_0}(r) : x_d > \gamma(x_1, ..., x_{d-1})\}. \tag{1.4}$$

We generally denote by ν or $\vec{n}$ the outward-pointing unit normal vector to $\partial\Omega$.

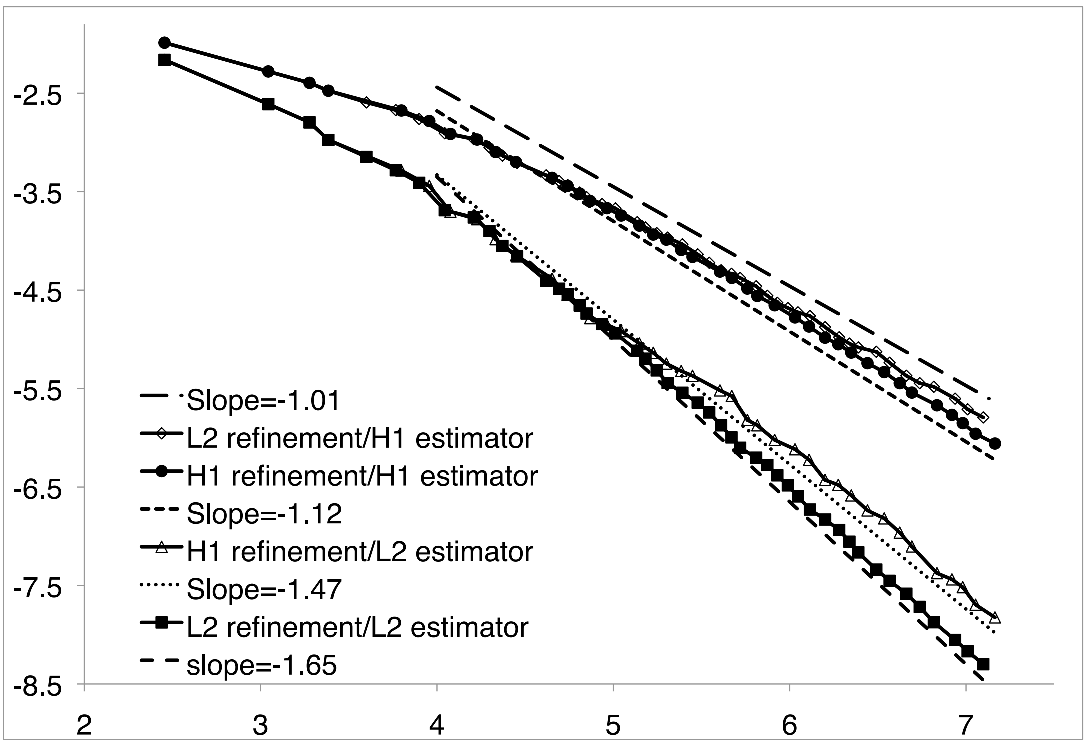

Figure 1.13: Adaptive convergence rates on 3D convex polyhedral domain.

Let $\gamma : \Omega \to \mathbb{R}$. γ is Lipschitz if $|\gamma(x) - \gamma(y)| \le L|x-y|$ for all $x, y \in \Omega$. L is then the Lipschitz constant of γ. (Note that all C^1 functions are Lipschitz, but not vice-versa.) γ is Hölder continuous with exponent α ($0 < \alpha < 1$) if $|\gamma(x) - \gamma(y)| \le C_\alpha |x-y|^\alpha$, $x, y \in \Omega$. A domain Ω has *Lipschitz boundary* if in the definition of C^k domain we replace the requirement that the functions γ be C^k with the requirement that they be Lipschitz. We can also extend the definition of C^k boundaries to $C^{k,\alpha}$ boundaries by requiring that the functions γ in the above definition be k times differentiable with all k-th derivatives Hölder continuous with exponent α.

We say that a domain $\Omega \subset \mathbb{R}^2$ is polygonal if its boundary can be broken into a set of edges consisting of open line segments and vertices. A three dimensional domain $\Omega \subset \mathbb{R}^3$ is polyhedral if its boundary can be decomposed into open polygonal faces, open edges, and vertices. The definition of polyhedral (polygonal) domains can be generalized to the notion of domains having polyhedral (polygonal) structure. There the definition of edges and faces above is generalized to include smooth curve segments in the case of edges and the images of polygons under smooth mappings in the case of faces. While we only consider polyhedral domains below, domains with polyhedral structure are important as they describe objects such as aircraft wings that have both curved smooth portions and edges (and possibly vertices). Much of what we will learn about boundary value problems on polyhedral domains applies as well to domains with polyhedral structure.

Finally, we say that Ω satisfies the *cone condition* if there exists a finite cone C such that each $x \in \Omega$ is the vertex of a finite cone C_x contained in Ω and congruent to C.

It is important to understand the relationship between various notions of smoothness of domain. All C^k domains ($k \ge 1$) are Lipschitz. There is not an inclusion in either direction between the sets of Lipschitz and polyhedral domains. That is, not all polyhedral domains are Lipschitz, and vice-vera. All polyhedral domains do however satisfy the cone condition. Examples of domains not satisfying the cone condition include those having cusps on the boundary.

Over the past several decades much effort has been devoted to the study of elliptic boundary value problems on domains with nonsmooth boundaries. There is a large body of literature devoted to the study of BVP's on Lipschitz domains, and there are many analytical results that are useful to numerical analysts that have been proved under the assumption that Ω is Lipschitz. Such results can also often be useful when considering boundary value problems on polyhedral domains since most (but not all) interesting polyhedral domains are Lipschitz. There are however important counterexamples. These include domains with cracks, as in the computational example in the previous subsection. Another well-known counterexample is the "two-brick" domain (Figure 1.14), which fails to be Lipschitz because the normals on faces near the origin (denoted by the arrow in the illustration) point in opposite directions. One might think that boundary value problems on non-Lipschitz domains might behave worse than on Lipschitz domains, but in the case of polyhedral domains that is not really the case. Polyhedral structure is special and worth studying in its own right in the context of boundary value problems, and the non-Lipschitz nature of these counterexamples does little to change the properties of Lipschitz boundary value problems posed on them.

As a final note, there are also more general classes of non smooth domains that have appeared in the analysis literature and which may be useful for analyzing finite element methods. For example, domain decomposition algorithms can lead to consideration of "jagged" subdomains consisting of somewhat arbitrary unions of mesh elements. Such domains might not have uniformly Lipschitz character, so they have sometimes instead been characterized as *John domains*.

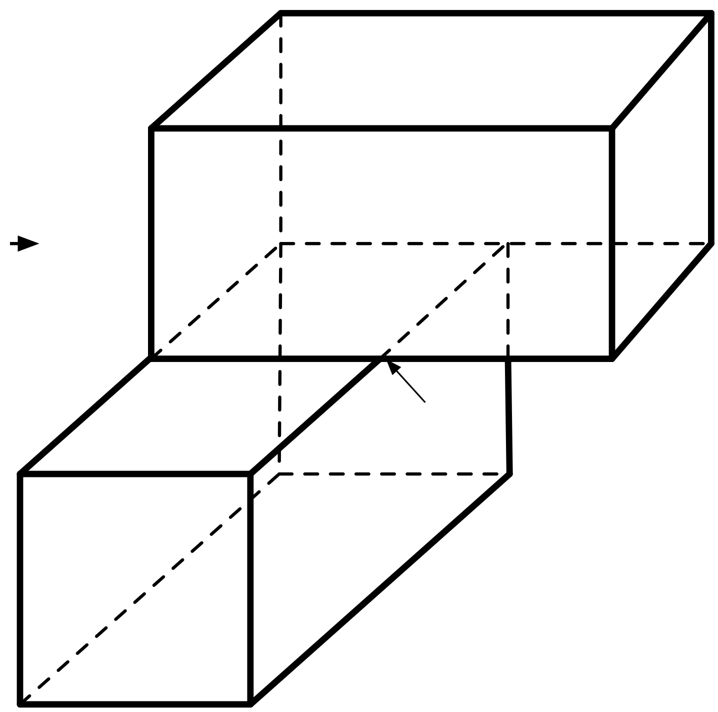

Figure 1.14: Non-Lipschitz “two-brick” domain.

1.3.2 Function spaces

Here we discuss Sobolev spaces (mainly); cf. the well-known books by Adams and Fournier [1, 2] for more details.

Given a domain Ω, $L_p(\Omega)$ is the space of p-integrable measurable functions on Ω ($1 \le p < \infty$) and $L_\infty(\Omega)$ the space of essentially bounded functions. A function $u \in L_1(\Omega)$ has a weak derivative $D^\alpha u$ (with α a multi index) if for all $\phi \in C_0^\infty(\Omega)$,

$$\int_\Omega u D^\alpha \phi \, \mathrm{d}x = (-1)^{|\alpha|} \int_\Omega \phi D^\alpha u \, \mathrm{d}x.$$

We then set for $1 \le p < \infty$

$$W^{p,k}(\Omega) = \{u \in L_p(\Omega) : D^\alpha u \in L_p(\Omega), |\alpha| \le k\},$$
$$\|u\|_{W^{p,k}(\Omega)} = (\sum_{|\alpha| \le k} \|D^\alpha u\|^p_{L_p(\Omega)})^{1/p}. \tag{1.5}$$

For $p = \infty$, we have instead $\|v\|_{W^k_\infty(\Omega)} = \max_{|\alpha|\le k} \|D^\alpha v\|_{L_\infty(\Omega)}$. Note that $W^{p,k}$ is a Banach space (complete linear space), and that when $p = 2$ it is a Hilbert space.

When $1 \le p < \infty$, $C^k(\Omega)$ (the k-times continuously differentiable functions on Ω is dense in $W^{p,k}(\Omega)$. That is, every $u \in W^{p,k}(\Omega)$ can be approximated by a sequence $u_n \in C^k(\Omega)$ such that $\|u - u_n\|_{W^{p,k}(\Omega)} \to 0$ as $n \to \infty$. This is true for any domain Ω (cf. Theorem 3.17 of [2]). Such a result does *not* hold for $p = \infty$. As an interesting historical note, this density result was fist published in a 1964 paper by Meyers and Serrin having the title $H = W$.

We additionally denote by $W_0^{p,k}(\Omega)$ the closure of $C_0^\infty(\Omega)$ in the W_p^k norm. Here $C_0^\infty(\Omega)$ is the set of infinitely differentiable functions having compact support in Ω. Note that in the context of second-order elliptic boundary value problems, we generally do not consider spaces $W_0^{p,k}$ for $k > 1$. This is roughly because we only wish to enforce boundary conditions directly on u itself, while $W_0^{p,k}(\Omega)$ ($k \ge 2$) enforces 0 boundary conditions on all derivates of order $k - 1$ and less. The solution to $-\Delta u = f$, $u = 0$ on $\partial\Omega$ thus may lie in $H_0^1(\Omega) \cap H^2(\Omega)$, but not generally in $H_0^2(\Omega)$.

We also define fractional-order Sobolev spaces. Let $1 \le p < \infty$, and for nonnegative real number s let $s = \bar{s} + \theta$, where $\bar{s}$ is an integer and $0 < \theta < 1$. We then define the seminorm

$$|u|^p_{W^{p,s}(\Omega)} = \sum_{|\alpha|=\bar{s}} \iint_{\Omega\times\Omega} \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x - y|^{n+\theta p}} \, \mathrm{d}x \, \mathrm{d}y \tag{1.6}$$

and the norm

$$\|u\|_{W^{p,s}(\Omega)} = \left(\|u\|^p_{W_p^{\bar{s}}(\Omega)} + |u|^p_{W^{p,s}(\Omega)}\right)^{1/p}; \tag{1.7}$$

cf. [37], Definition 1.3.2.1. When $p = 2$ we set $H^s(\Omega) = W_2^s(\Omega)$, as usual.

1.3.3 Sobolev imbeddings

It is natural (and often of interest in applications) to ask whether a function lying in given Sobolev space is also continuous, or lies in some other Sobolev space which in general has stronger integrability index p but weaker differentiability index k. The concept of Sobolev imbeddings is intended to answer such questions. Our intention is to give a brief introduction to the idea of such imbeddings; see [1, 2] as above for more details.

We first give a standard example. The function $u : \mathbb{R}^2 \to \mathbb{R}$ defined by $u(x) = \log\log|x|$ can be shown to lie in $H^1(\Omega)$ for any finite domain Ω (including ones containing the origin), but it is not continuous at the origin. Thus H^1 functions need not be continuous. This is important because the weak form of Poisson's problem, and second-order elliptic problems in general, is solved over $H^1(\Omega)$ or $H_0^1(\Omega)$. We thus might naturally ask what conditions are necessary to ensure continuous solutions of Poisson's problem.

Given a bounded domain Ω, we let $C_B^k(\Omega)$ denote those functions which are in $C^k(\Omega)$ with uniformly bounded derivatives of all orders up to k. $C_B^k(\Omega)$ is a Banach space with norm $\|u\|_{C_B^k(\Omega)} = \max_{|\alpha|\le k} \|D^\alpha u\|_{L_\infty(\Omega)}$. Note that $C_B^k(\Omega) \subset W_\infty^k(\Omega)$, but not vice-versa. We write $C_B(\Omega) = C_B^0(\Omega)$.

Given Banach spaces X and Y, we say that X is imbedded into Y and write $X \to Y$ if X is a vector subspace of Y and $\|u\|_Y \lesssim \|u\|_X$.

Theorem 1.3.1 *(Sobolev Imbedding Theorem [2]). Suppose that Ω satisfies the cone condition. Then*

1. *If either $kp > d$ or $k = d$ and $p = 1$, then*
$$W^{k,p}(\Omega) \to C_B(\Omega). \tag{1.8}$$
If in addition $1 \le k \le d$, then
$$W^{k,p}(\Omega) \to L_q(\Omega), \quad p \le q \le \infty. \tag{1.9}$$
2. *If $kp = d$, then*
$$W^{k,p}(\Omega) \to L_q(\Omega), \quad p \le q < \infty. \tag{1.10}$$
3. *If $kp < d$ and either $d - kp \le d$ or $p = 1$, then*
$$W^{k,p}(\Omega) \to L_q(\Omega), \quad p \le q \le p_* = \frac{dp}{d-kp}. \tag{1.11}$$

Here the imbedding constants generally depend on the involved parameters and may degenerate (blow up) as limit cases are approached.

We give a couple of examples. Consider (1.8) with $k = 1$. We may conclude that if $u \in W^{1,p}(\Omega)$ with $p > d$, then u is continuous with $\|u\|_{L_\infty(\Omega)} \lesssim \|u\|_{W^{1,p}(\Omega)}$. In particular, in two space dimensions u is continuous if $u \in W^{1,p}(\Omega)$ for any $p > 2$, while when $d = 3$ u is continuous if $u \in W^{1,p}(\Omega)$ for some $p > 3$. We next take $k = 2$, that is, we ask what integrability index the second derivatives of u must possess in order for u to be continuous. For $d = 2$, we have $d = k = 2$ and thus (1.8) holds with $p = 1$. For $d = 3$ we require $2p > 3$, or $p > 3/2$ in order to ensure that u is continuous.

It can be important to note that the constants involved in the Sobolev imbeddings can blow up near limit cases, though often in a known fashion. To illustrate this, we follow the discussion of Lemma 5.4 of [48]. As an example, if $d = 2$ and $u \in H_0^1(\Omega)$ and $2 \le p < \infty$, then

$$\|u\|_{L_p(\Omega)} \lesssim \sqrt{p}\|\nabla u\|_{L_d(\Omega)}.$$

Note that the constant $\sqrt{p}$ blows up as $p \uparrow \infty$. Combining the above estimate for correctly chosen p with finite element inverse estimates yields a *discrete Sobolev inequality* when $d = 2$. In particular, assume that Ω is polyhedral with T_h a quasi-uniform mesh and $V_h \subset H_0^1(\Omega)$ an associated Lagrange finite element space as above. Then for $\chi \in V_h$,

$$\|\chi\|_{L_\infty(\Omega)} \lesssim h^{-2/p} \|\chi\|_{L_p(\Omega)}, \quad \chi \in S_h.$$

We then let $p = \ln \frac{1}{h}$. In this case $h^{-2/p} = (e^{-\ln h})^{2/\ln h} \lesssim 1$, so

$$\|\chi\|_{L_\infty(\Omega)} \lesssim (\ln 1/h)^{1/2} \|\nabla \chi\|_{L_2(\Omega)}.$$

The above inequality does not hold for general functions in H^1, as we have seen above. However, finite element functions (at least those we consider here) are continuous, so we can get away with an "almost"-Sobolev inequality. We pay for it with the factor $\ln \frac{1}{h}$, which blows up as $h \to 0$ (albeit slowly). This discrete Sobolev inequality is well-known and fairly widely used in the finite element literature, though it is mainly useful for analyzing FEM for 2D problems. (A 3D version holds but requires measuring the gradient in L_3 instead of L_2, which is not nearly as interesting because H^1 is the natural energy space and $W^{1,3}$ generally isn't!)

1.3.4 Basic regularity result for elliptic PDE's

We conclude by giving a more precise statement of the shift theorem (1.1); cf. Chapter 6 of the textbook [30] of Evans. The monograph [34] by Gilbarg and Trudinger is also a good resource for basic material about elliptic PDE.

Consider the divergence-form elliptic operator

$$Lu := -\sum_{i,j=1}^d (a^{ij}(x) u_{x_i})_{x_j} + \sum_{i=1}^d b^i(x) u_{x_i} + c(x) u.$$

Assume first that the coefficients a_{ij}, b^i, and c are all in $L_\infty(\Omega)$, and that the bilinear form

$$a(u, v) := \int_\Omega (\sum_{i,j=1}^d a^{ij}(x) u_{x_i} v_{x_j} + v \sum_{i=1}^d b^i(x) u_{x_i} + c(x) uv) \, \mathrm{d}x$$

is coercive and continuous over $H_0^1(\Omega)$ in the sense that

$$\|u\|_{H^1(\Omega)}^2 \lesssim a(u, u), \quad a(u, v) \lesssim \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}, \quad u, v \in H_0^1(\Omega).$$

We also assume that the operator is (strongly) elliptic, that is, that the matrix $A = [a^{ij}]$ of coefficients is uniformly bounded and positive definite in the sense that there are constants $\lambda, \Lambda > 0$ such that

$$\sum_{i,j=1}^d |a^{ij}(x)| \leq \Lambda^2, \quad a^{ij}(x) \xi_i \xi_j > \lambda |\xi|^2, \quad x \in \Omega, \quad \xi \in \mathbb{R}^d.$$

The weak form of the PDE $Lu = f$, $u = 0$ on $\partial\Omega$ is: Find $u \in H_0^1(\Omega)$ such that

$$a(u, v) = (f, v), \quad v \in H^1(\Omega).$$

Then the Lax-Milgram theorem guarantees existence and uniqueness of $u \in H_0^1(\Omega)$ satisfying the above for any linear functional $f \in H^{-1}(\Omega) := \{f : H_0^1(\Omega) \to \mathbb{R} : f(v) \lesssim \|v\|_{H_0^1(\Omega)},\ v \in H_0^1(\Omega)\}$ and $\|u\|_{H_0^1(\Omega)} \lesssim \|f\|_{H^{-1}(\Omega)}$. Such a u is the weak solution to the strong-form problem $Lu = f$ in Ω, $u = 0$ on $\partial\Omega$.

Theorem 1.3.2 *Assume that $a^{ij} \in C^1(\bar{\Omega})$, $b^i, c \in L_\infty(\Omega)$, $1 \le i, j \le d$, that $f \in L_2(\Omega)$, that $u \in H_0^1(\Omega)$ is the unique weak solution of $Lu = f$, $u = 0$ on $\partial\Omega$ as above, and that $\partial\Omega$ is C^2. Then $u \in H^2(\Omega)$, and*

$$\|u\|_{H^2(\Omega)} \lesssim \|f\|_{L_2(\Omega)}$$

with the constant hidden in $\lesssim$ depending only on Ω and the coefficients of L.

The proof is rather long and technical, and outside of the scope of this course. We do however make a couple of notes:

1. The basic approach of [30] to regularity proofs is to establish estimates for difference quotients of ∇u and take a limit as the mesh size approaches 0.

2. Establishing interior regularity results (that is, proving that $u \in H^2(\Omega')$ for domains $\Omega' \subset\subset \Omega$) is pretty straightforward, if technical, under the assumptions in Theorem 1.3.2.

3. To establish that $u \in H^2(\Omega)$ for domains close to the boundary, the general strategy is to flatten the boundary locally. That is, one first proves regularity if Ω has a flat boundary. One then maps portions of $\partial\Omega$ that aren't flat to a domain having flat boundary and then proceeds as before–with the added difficulty of using the chain rule to take the mapping into account. This is the point where regularity of the boundary comes into play in classical proofs.

We finally state a version of Theorem 1.3.2 assuming higher boundary and coefficient regularity.

Corollary 1.3.3 *In addition to the assumptions of Theorem 1.3.2, assume that $a^{ij}, b^i, c \in C^{k+1}(\bar{\Omega})$, $f \in H^m(\Omega)$, and $\partial\Omega$ is C^{m+2}. Then $u \in H^{m+2}(\Omega)$ and*

$$\|u\|_{H^{m+2}(\Omega)} \lesssim \|f\|_{H^m(\Omega)},$$

the constant hidden in $\lesssim$ only depending on Ω, m, and the coefficients of L.

We also briefly mention other types of estimates for elliptic boundary value problems; cf. [34, Chapters 8 and 9].

1. $W^{p,2}$ estimates: Assume that $a^{ij} \in C^1(\bar{\Omega})$, $b^i, c \in L_\infty(\Omega)$ as in Theorem 1.3.2, that $\partial\Omega$ is $C^{1,1}$, and that $f \in L_p(\Omega)$ for some $1 < p < \infty$. Then the problem (written in strong form) of finding u such that $Lu = f$, $u = 0$ on $\partial\Omega$ has a unique weak solution $u \in W^{p,2}(\Omega) \cap W_0^{p,1}(\Omega)$ satisfying
$$\|u\|_{W^{p,2}(\Omega)} \le C_p \|f\|_{L_p(\Omega)}.$$
Note that the above estimate does *not* hold for $p = 1, \infty$, and that C_p blows up as $p \uparrow \infty$ or $p \downarrow 1$. (The constant C_p comes from the Marcinkiewicz interpolation theorem at least for sufficiently smooth domains. I have seen assertions that it behaves as Cp for large p and $\frac{C}{p-1}$ for p close to 1, but I couldn't immediately find a reference.)

2. Maximum principles: The weak maximum principle states that if $u \in W^{2,1}(\Omega)$ satisfies $Lu \leq 0$ ($Lu \geq 0$) in Ω, then
$$\sup_{\Omega} u \leq \sup_{\partial\Omega} u^+ \quad (\inf_{\Omega} u \geq \inf_{\partial\Omega} u^-).$$
The rough intuition of the maximum principle is that $Lu \leq 0$ implies that u is convex, and $Lu \geq u$ that u is concave. This corresponds to our usual intuition from calculus, since $Lu \leq 0$ roughly means that the second derivative of u is positive.

3. Hölder estimates for the first derivatives of u: Following [34, Section 8.11], we now assume that $Lu = \Delta u$ and that $\Delta u = g + \sum_{i=1}^{d} D_{x_i} f^i$. If u is a weak solution with $u \in C^{1,\alpha}(\bar{\Omega})$ and $u|_{\partial\Omega} = 0$, then
$$\|u\|_{C^{1,\alpha}(\Omega)} \lesssim \|u\|_{C^0(\Omega)} + \|g\|_{C^0(\Omega)} + \|f\|_{C^{0,\alpha}(\Omega)}.$$