Chapter 5

Adaptive Methods

5.1 Introduction

As our trial functions - set of basis functions - we have thus far utilized piecewise linear functions. We noted that this amounted to justifying that the function for which we sought a solution, varied only linearly over each element. In effect, by describing our function within our linear subspace, we were approximating the actual behavior of the function by truncating the degree of the polynomial representation we used. Whenever our element fails to approximate the behavior of the actual solution because of either truncating the polynomial we are using to represent the physical quantity *and/or* beause we have not used an appropriate number of elements to capture the physical phenomena, we have committed a *discretization error*. To obtain a physical feeling for the errors encountered from discretization, consider the following heuristic example.

Take the approximation of our field $u(x, y)$, by a two dimensional Taylor's series expansion about a point (x, y):

$$u(x+h, y+k) = u(x,y) + [h\frac{\partial u(x,y)}{\partial x} + k\frac{\partial u(x,y)}{\partial y}] + \frac{1}{2!}[h^2\frac{\partial^2 u(x,y)}{\partial^2 x} + 2hk\frac{\partial^2 u(x,y)}{\partial x \partial y} + k^2\frac{\partial^2 u(x,y)}{\partial^2 y}] + \ldots \tag{5.1}$$

Let us neglect the second order term and using the Taylor's expansion, estimate the nodal values of the function on the triangle shown in Figure 5.1. Here $h = \max|x_i - x_j, x_i - x_m, x_m - x_j|, k = \max|y_i - y_j, y_i - y_m, y_m - y_j|$, are the maximum x and y distances respectively. If we estimate the function at the jth and mth nodes in terms of the function at the ith node using the Taylor's series, we obtain:

$$u_j = u_i + [x_j\frac{\partial u(x_i, x_j)}{\partial x} + y_m\frac{\partial u(x_i, x_j)}{\partial y}] \tag{5.2}$$

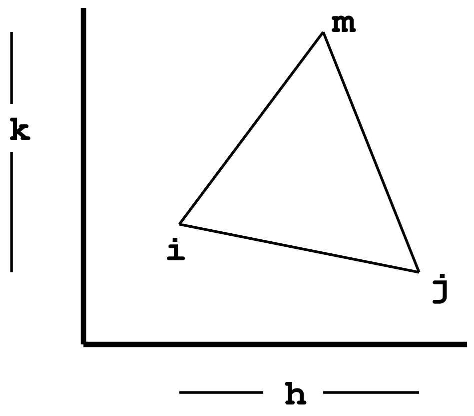

Figure 5.1: A triangular element.

$$u_m = u_i + [x_m \frac{\partial u(x_i, x_j)}{\partial x} + y_j \frac{\partial u(x_i, x_j)}{\partial y}] \tag{5.3}$$

solving for $\frac{\partial u(x_i, x_j)}{\partial x}$, one obtains the standard linear interpolation function for a triangle:

$$\frac{\partial u(x_i, x_j)}{\partial x} = \frac{1}{2A}[u_i(y_j - y_m) + u_m(y_i - y_j) + u_j(y_m - y_i) \tag{5.4}$$

where A is the area of the triangle. Likewise, one could calculate the interpolant for the other two nodes and discover that:

$$\frac{\partial u(x_i, y_i)}{\partial x} = \frac{\partial u(x_j, y_j)}{\partial x} = \frac{\partial u(x_m, y_m)}{\partial x} \tag{5.5}$$

is constant over the triangle (and thus so is the gradient in y as well). Thus we can derive the standard linear interpolation formulas on a triangle which represent the first two terms of the Taylor's series expansion. This means that the error due to discretization (from using linear elements) is proportional to the third term of the Taylor's expansion:

$$\epsilon \approx \frac{1}{2!}[h^2 \frac{\partial^2 u(x, y)}{\partial^2 x} + 2hk \frac{\partial^2 u(x, y)}{\partial x \partial y} + k^2 \frac{\partial^2 u(x, y)}{\partial^2 y}] \tag{5.6}$$

where u is the exact solution and again, h and k are the maximum distances in the x and y directions. Thus the error is proportional to the second derivative. If u is a linear function over the element, then the first derivative is a constant and the second derivative is zero and there is no error due to discretization. This implies that the gradient must be constant over each element. If the function is not linear, or the gradient is not constant over an element, the second derivative will not be zero and is proportional to the error incurred due to discretization. It is noted that as one decreases the constants h and k to zero the error tends to zero as well, thus decreasing the mesh size in places of high errors due to high gradients, decreases the error. As an aside, we note that if one divides the previous equation by hk, one can also express the error in terms of the elemental aspect ratio $\frac{h}{k}$, which is a measure of the relative shape of the element, as is easily seen, one must be careful to maintain an aspect ratio as close to unity as possible (say $\frac{h}{k} < 4 - 5$ or less).

The problem with the preceding heuristic argument is that one had to know the exact solution *a priorily* before one could estimate the error. One can, however derive an estimate of the upper bound of the mean square norm of the interpolation error and use this error estimate to adaptively refine the mesh to reduce the error due to discretization to within an *a priorily* defined error tolerance as we shall see later on.

5.1.1 Convergence of a Sequence of Approximate Solutions

Let's try and quantitate our error a bit further. When considering the preceeding example, it seems to make sense that if we increase the number of degrees of freedom we used to approximate our function, the accuracy must approach the true solution. That is, we would hope that the sequence of approximate solutions will *converge* to the exact solution as the number of degrees of freedom (DOF) increases indefinitely.

$$\|u(x) - \tilde{u}_n(x)\| \to 0 \quad \text{as } n \to \infty \tag{5.7}$$

This is a statement of *pointwise convergence*. It describes the approximate solution as approaching arbitrarily close to the exact solution at each point in the domain as the number of DOF increases.

Measures of convergence often depend on how the *closeness* of measuring the distance between functions is defined. Other common descriptions of measuring convergence are, *uniform convergence* which requires that the maximum value of $\|u(x) - \tilde{u}_n(x)\|$ in the domain vanish as $n \to \infty$. This is stronger than pointwise convergence as it requires a uniform rate of convergence at every point in the domain. Two other commonly used measures are *convergence in energy* and *convergence in mean* which involve measuring an *average* of a function of the pointwise error over the domain.

In general proving pointwise convergence is very difficult except in the simplest cases, while proving the convergence of an averged value, such as energy is often easier. Of course, scientists and engineers are often much more interested in assuring that their answers are accurate in a pointwise sense rather than an energy sense because they typically want to know values of the solution $u(x)$ and graidents, $\nabla u(x)$ at specific places.

One intermediate form of convergence is called the *Cauchy convergence*. Here, we require the sequences of two different approximate solutions to approach arbitrarily close to each other,

$$\|u_m(x) - \tilde{u}_n(x)\| \to 0 \quad \text{as } m, n \to \infty \tag{5.8}$$

While the pointwise convergence expression would imply the previous equation, it is important to note that the Cauchy convergence does not imply pointwise convergence as the functions could converge an answer other than the true solution.

While we cannot be assured pointwise convergence of these functions for all but the simplist cases, there do exist theorems that ensure that a sequence of approximate solutions must converge to the exact solution (assuming no computational errors) if the basis functions staisfy certain conditions. The theorems can only ensure convergence in an average sense over the entire domain, but it is usually the case that if the solution converges in an averge sense (energy, etc.), then it will converge in the pointwise sense as well.

5.2 Energy Norms

The error in energy, measured by the *energy norm* is defined in general as [33, 34, 35],

$$\|e\| = (\int_\Omega e^T L e \, d\Omega)^{\frac{1}{2}} \tag{5.9}$$

where $e = u(x) - \tilde{u}_n(x)$ and L is the differential operator for the governing differential equation (i.e. it contains the derivatives operating on $u(x)$ and any function multiplying $u(x)$). For physical problems this is often associated with the energy density.

Another common measure of convergence utilizes the L_2 norm. This can be termed the average error and can be associated with errors in any quantity. The L_2 norm is defined as,

$$\|e\|_{L_2} = (\int_\Omega e^T e \, d\Omega)^{\frac{1}{2}} \tag{5.10}$$

While the norms given above are defined on the whole domain, one can note that the square of each can be obtained by summing element contributions,

$$\|e\|^2 = \sum_{i=1}^{M} \|e\|_i^2 \tag{5.11}$$

where i represents an element contribution and m the total element number. Often for an *optimal* finite element mesh, one trys to make the contributions to this square of the norm equal for all elements.

While the absolute values given by the energy or L_2 norms have little value, one can construct a relative percentage error that can be more readily interpreted,

$$\eta = \frac{\|e\|}{\|u\|} \times 100 \tag{5.12}$$

this quantity, in effect, represents a weighted RMS error. The analysis can be determined for the whole domain or for element subdomains. This can be used in an adaptive algorithm by checking element errors against some predefined tolerance, η_0 and increasing the DOF only those areas which are above the predefined tolerance.

As we noted before, the error in the scalar field (i.e. temparture, voltage, etc.), e_u, is defined as the difference between the exact potential, u, and the calculated potential, $\tilde{u}$.

$$e_u = u - \tilde{u} \tag{5.13}$$

Similarly, the error in the gradient of the potential, q, is

$$e_q = q - \tilde{q} \tag{5.14}$$

where $q = \nabla u$ and $\tilde{q} = \nabla \tilde{u}$. The error norm is defined as:

$$\|e_q\| = (\int_\Omega (\nabla u - \nabla \tilde{u})^T (\nabla u - \nabla \tilde{u}) d\Omega)^{\frac{1}{2}}. \tag{5.15}$$

Zienkiewicz [36, 37] has shown that

$$\int_\Omega (\nabla u)^T (\nabla \tilde{u}) d\Omega = \int_\Omega (\nabla \tilde{u})^T (\nabla u) d\Omega = \int_\Omega (\nabla \tilde{u})^T (\nabla \tilde{u}) d\Omega. \tag{5.16}$$

Using this result, (5.15) becomes

$$\|e_q\|^2 = \int_\Omega (\nabla u)^T (\nabla u) d\Omega - \int_\Omega (\nabla \tilde{u})^T (\nabla \tilde{u}) d\Omega \tag{5.17}$$

or equivalently,

$$\|e_q\|^2 = \|q\|^2 - \|\tilde{q}\|^2 \tag{5.18}$$

Here, $\|q\|^2$ is a measure of the total energy in the domain. Using this, the percentage error, η, can be defined to be

$$\eta = \frac{\|e_q\|}{\|q\|} \cdot 100\% \tag{5.19}$$

This error measure gives a ratio of the energy difference and the total energy which is in effect a percentage error for each element.

The FEM we implemented used linear basis functions to describe the variation of the potential in each element. At every point in the mesh, the function is uniquely specified. However, the gradient is not specified at every point in space but is defined to be constant over each element and discontinuous across element boundaries. Zienkiewicz showed that when the vector field arising from the gradients are globally smoothed, they provide a more accurate estimate of the energy than the energy using the constant vector fields [36, 37]. Hence, (5.19) can be calculated using the energy due to the smoothed currents and the energy due to the constant vectors. The smoothed current densities are now defined at the same nodal locations as the scalar values and are continuous across element boundaries.

The smoothing process used the Galerkin technique [38] involves minimizing the difference between the two vector fields,

$$(\nabla \hat{u} - \nabla \tilde{u}) \tag{5.20}$$

where $\nabla \hat{u}$ is the smoothed gradient representing the exact gradient and $\nabla \tilde{u}$ is the constant gradient from the FEM solution. This method produces a system of equations in the following form:

$$(\int_{\Omega_e} \Phi_i \Phi_j d\Omega_e)(\nabla \hat{u}_i) = (\int_{\Omega_e} \nabla \tilde{u}_i \Phi_i d\Omega_e). \tag{5.21}$$

where Φ_i is a linear basis function defined at node i, $\nabla \hat{u}_i$ is the value of the smoothed gradient at that point, and $\nabla \tilde{u}_i$ is the constant gradient resulting from the finite element solution. In equation (5.21), the vector field, $\nabla \tilde{u}_i$ is the unknown that is solved for at each node, i of the grid. Thus we calculate a new estimate, $\hat{q}$, which can be used in place of the exact gradient, q. The vector field now varies linearly in each element and is continuous everywhere.

Once the energy error has been computed, the mesh refinement can begin. The error must be related to some parameter which guides the mesh refinement. We can utilize a spacing function, h_e, which controls the size and number of elements [38]. This spacing function, h_e is defined to be the linear interpolation of the spacing values at the nodes of the element. A variable ξ is defined as the ratio of the element error to the average of the element errors:

$$\xi_e = \frac{\|e_q\|_e}{\|\tilde{e}_q\|_e}. \tag{5.22}$$

If $\xi_e > 1$, then a new spacing function can be defined as

$$h^* = \frac{h}{\xi_e}. \tag{5.23}$$

Thus, the spacing values at each node of the element are reduced. This new spacing function will then increase the number of points in the regions of large errors. This whole process is repeated until the global error estimate falls below the specified level.

The algorithm does not depend on how the modification was made such that a different error estimate other than the one just described can be used by only modifying the spacing values in some fashion. This allows any number of error estimators to be used to guide the adaptive mesh refinement process.

For the convergence of the above norms to be valid, we need to make sure our basis functions meet certain criteria, namely those of *completeness* and *continuity.* These can be stated as:

Completeness condition. The element trial solution, $\tilde{u}^{(e)}$ and any of its derivaties up to order m appearing in the integrals of the Galerkin (or variational - weak) formuation should be able to assume any constant value within an element when, in the limit, the size of the elemtn decrease to zero.

Continuity condition. At interelement boundaries, the element trial solutions should be C^{m-1}-continuous; that is, $\tilde{u}^{(e)}$ and its derivatives up to order $m-1$ should be continuous.

If the element basis functions, $\Phi_i^{(e)}$ are constructed so that $\tilde{u}^{(e)}$ satifies both conditions, then a sequence of approximate solutions, corresponding to a sequence of successively refined meshes, will converge energywise to the exact solution (again, assuming no other errors) in the limit as $h_{\max} \to 0$. Where h represents the size of the element. This is often refered to as *h-convergence.*

The completeness condition is necessary for convergence to the exact solution. An incomplete element will generally converge to some other limiting function. The continuity condition is not necessary for convergence. The purpose of the continuity condition is to ensure that discontinuities at the interelement boundaries are not severe enough to introduce errors in addition to the discretization error.

If continuity and completeness conditions are satisfied, convergence is assured. An element which satisfies both conditions is called a *conforming* or *compatible* element. An element which satifies completeness but not continuity is called a *nonconforming* or *incompatible* element.

5.3 A priori and A posteriori Error Estimates

First, we will prove that the Galerkin finite element approximation is the best approximation to the true solution in the discrete domain with respect to the energy norm. Then, we will derive some *a posteriori* error estimators.

Theorem: Assume that u satifies the Poisson equation and u_h is the Galerkin finite element approximation in Ω_h. Then

$$\|\nabla(u-u_h)\| \leq \|\nabla(u-v)\| \ \ \forall v \in \Omega_h \tag{5.24}$$

Proof: Using the Galerkin orthogonality with $u_h - v \in \Omega_h$, we can write

$$\|\nabla e\|^2 = (\nabla e, \nabla(u-u_h)) = (\nabla e, \nabla(u-u_h)) + (\nabla e, \nabla(u_h - v)). \tag{5.25}$$

Adding the terms involving u_h on the right, whereby u_h drops out, and using Cauchy's inequality, we get

$$\|\nabla e\|^2 = (\nabla e, \nabla(u-v)) \leq \|\nabla e\| \|\nabla(u-v))\|, \tag{5.26}$$

which proves the theorem after dividing by $\|\nabla e\|$.

The following theorem summarizes the interpolation error estimates that we need in the analysis of the finite element approximation for Poisson's equation. This result estimates the error of the continuous piecwise linear interpolant of a function over the entire domain in the L_2 norm.

Theorem: There are interpolation error constant C_i, depending only on the minimum angle in the mesh T and the order of the estimate m, such that the piecewise linear nodal interpolant, $\pi_h w \in \Omega_h$ of a function w satisfies for $m = 0$ and $m = 1$,

$$\|D^m(w - \pi_h w)\| \leq C_i \|h^{2-m} D^2 w\| \tag{5.27}$$

where $D^1 w = Dw = \nabla w$, and

$$\|h^{-2+m} D^m(w - \pi_h w)\| + (\sum_{K \in T_h} h_K^{-3} \|w - \pi_h w\|_{\partial K}^2)^{1/2} \leq C_i \|h^{2-m} D^2 w\|. \tag{5.28}$$

Furthermore, thre is an interpolant $\pi_h w_h$ in Ω_h of w such taht for $m = 0$ and 1 we have

$$\|h^{-1+m} D^m(w - \pi_h w)\| + (\sum_{K \in T_h} h_K^{-1} \|w - \pi_h w\|_{\partial K}^2)^{1/2} \leq C_i \|h^{2-m} D w\|. \tag{5.29}$$

Using the interpolation results from the previous theorem and choosing $v = \pi_h u$, we get the following concrete quantitative a priori error estimate

$$\|\nabla(u - u_h)\| \leq C_i \|h D^2 u\| \tag{5.30}$$

5.4 $H^k(\Omega)$ Norms

Let's now study the error, $u - u_h$ in a little more detail. Remembering that the Galerkin formulation for the original differential equation is,

$$(u', v') = (f, v) \quad \forall v \in \Omega \tag{5.31}$$

and our finite element approximation is,

$$(u_h', v') = (f, v) \quad \forall v \in \Omega_h \tag{5.32}$$

Subtracting the to we obtain

$$((u - u_h)', v') = 0 \quad \forall v \in \Omega_h \tag{5.33}$$

In order to work with the preceeding equation, we need a couple of definitions. First define the scalar product, denoted by $(.,.)$ as given by the L_2 norm,

$$\|v\| = (\int_\Omega v^2 d\Omega)^{\frac{1}{2}} \tag{5.34}$$

and the norms with subscripts, $\|\cdot\|_k$ be defined in the Sobolev space, $H^k(\Omega)$ defined as,

$$\|v\|_k = (\sum_{|\alpha|\leq k} \|D^\alpha v\|_0^2)^{\frac{1}{2}} \tag{5.35}$$

where α is a multi-index, that is $\alpha = (\alpha_1, \alpha_2)$, with α_1 and α_2 nonnegative integers, $|\alpha| = \alpha_1 + \alpha_2$ and

$$D^\alpha v = \frac{\partial^{|\alpha|} v}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2}}. \tag{5.36}$$

The space $H^k(\Omega)$ consists off all functions v on Ω that, together with the partial derivatives $D^\alpha v$ or order $|\alpha|$ at most k, belong to $L_2(\Omega)$. For example, v in the $H^2(\Omega)$ can be expressed as

$$\|v\|_2 = (\|v\|_0^2 + \|\frac{\partial v}{\partial x_1}\|_0^2 + \|\frac{\partial v}{\partial x_2}\|_0^2 + \|\frac{\partial^2 v}{\partial x_1^2}\|_0^2 + \|\frac{\partial^2 v}{\partial x_2^2}\|_0^2 + \|\frac{\partial^2 v}{\partial x_1 \partial x_2}\|_0^2 + \|\frac{\partial^2 v}{\partial x_2 \partial x_1}\|_0^2 \tag{5.37}$$

Where we have used the notation $\|\cdot\|_0$ to indicate an L_2 norm. $H^k(\Omega)$ is a Hilbert space with the indicated norm and corresponding scalar product. The spaces $H^k(\Omega)$ are examples of so-called *Sobolev spaces* named after the Russian mathematician S. L. Sobolev. From a mathematical point of view, results for the finite elemtn method are most easily derived using $H^k(\Omega)$ norms. From a computational point of view, the pointwise estimates are more interesting but harder to obtain.

Now we just need to state Cauchy's inequality and we're ready to roll.

$$|(v, w)| \leq \|v\| \cdot \|w\| \tag{5.38}$$

Theorem 1. For any $v \in \Omega$,

$$\|(u - u_h)'\| \leq \|(u - v)'\| \tag{5.39}$$

This is a statement that u_h is the best approximation in Ω_h (in the sense of the particular norm we have chosen) of the function u. In other words, u_h is the orthogonal projection of u onto Ω_h [39].

Proof. Let $v \in \Omega_h$ be an arbitrary function and set $w = u_h - v$. Then $w \in \Omega_h$ and using $((u - u_h)', v') = 0$ with v replaced by w, we obtain, using Cauchy's inequality,

$$\begin{gathered} \|(u - u_h)'\|^2 = ((u - u_h)', (u - u_h)') + ((u - u_h)', w') = \\ ((u - u_h)', (u - u_h + w)') = ((u - u_h)', (u - v)') \leq \|(u - u_h)'\| \|(u - v)\|. \end{gathered} \tag{5.40}$$

by dividing by $\|(u - u_h)'\|$ we obtain our original statement.

From Theorem 1 we can obtain a quantitative estimate for the error, $\|(u - u_h)'\|$ by estimating $\|(u - \tilde{u}_h)'\|$ where $\tilde{u}_h$ is defined to be the *interpolant* of u (i.e. $\tilde{u}_h$ interpolates u at the nodes, $\tilde{u}_h(x_j) = \Phi(x_j)$. If $\tilde{u}_h$ is chosen in this way, then we can invoke error estimates used for interpolation theory to bound our finite element solution. For the interpolant, $\tilde{u}_h$ we have,

$$|u'(x) - \tilde{u}_h'(x)| \leq h \max |u''(y)| \tag{5.41}$$

and

$$|u(x) - \tilde{u}_h(x)| \leq \frac{h^2}{8} \max |u''(y)| \tag{5.42}$$

using Theorem 1 along with the preceeding two equations, we can obtain for the following error estimtes for the finite element error,

$$\|(u - u_h)'\| \leq h \max |u''(y)| \tag{5.43}$$

and

$$\|(u(x) - u_h(x)\| \leq h \max |u''(y)| \tag{5.44}$$

To see how we obtain the bounds in the preceeding equations, consider the case of linear interpolation. Let x_0 and x_1 be two real numbers and suppose $f(x)$ is defined at these points. Then the linear function $P(x)$ which satifies $P(x_0) = f(x_0)$ and $P(x_1) = f(x_1)$ is given by,

$$P(x) = \frac{x - x_1}{x_0 - x_1} f(x_0) + \frac{x - x_0}{x_1 - x_0} f(x_1) \tag{5.45}$$

then $P(x) \approx f(x)$ and is said to the the linear interpolant of $f(x)$. The truncation error for linear interpolation is defined as,

$$\frac{(x - x_0)(x - x_1) f''(\epsilon)}{2} \tag{5.46}$$

where $\min(x_0, x_1, 1) < \epsilon < \max(x_0, x_1, x)$. Suppose that $(x_0 < x < x_1)$ and $|f''(x)| < M$ on $[x_0, x_1]$, then the truncation error is bounded by,

$$\max_{x_0 \leq x \leq x_1} |\frac{(x - x_0)(x - x_1)}{2}| \max_{x_0 \leq \epsilon \leq x_1} |f''(\epsilon)|. \tag{5.47}$$

The first maximum occurs at $(x_0 + x_1)/2$. Using this along with the expression for the interpolation of $f(x)$,

$$f(x) = P(x) + \frac{x - x_0(x - x_1) f''(\epsilon)}{2} \tag{5.48}$$

we obtain

$$|f(x) - P(x)| \leq \frac{M(x_1 - x_0)^2}{8} \tag{5.49}$$

if we have equally spaced points at a, $a + h$, $a + 2h$, $\ldots$, $a + nh = b$ where $h = (b - a)/h$ then we have the bound,

$$\frac{h^2}{8} \max_{a \leq \epsilon \leq b} |f''(\epsilon)| \tag{5.50}$$

Take as a simple example the approximation of $f(x) = sin(x)$. Given $f(.1) = .09983$, $f(.2) = .19867$, find the approximate value for $f(.16)$ and estimate the error.

$$P(.16) = \frac{.16 - .2}{.1 - .2}(0.09983) + \frac{.16 - .1}{.2 - .1}(.19876) \approx .15913 \tag{5.51}$$

The error is thus,

$$|f(.16) - P(.16)| < (.19876)|\frac{(.16 - .1)(.16 - .2}{2}| \approx 0.00024 \tag{5.52}$$

where the value of (.19876) comes from noting that $f''(x) = -sin(x)$ and the maximum value of $|-sin(x)|$ on $[.1, .2]$ is given by $sin(.2) = .19876$.

The nice thing about the previous example is that it was one of those simple cases that we could get exact pointwise estimates (as we already new the form of the function - $sin(x)$). For *real* problems, we can't obtain such estimates and have to end up using approximate estimates. To continue off from where we were before, we have noted that the finite element solution was bounded from above by the interpolant. We can also show that the interpolant is bounded from above by another quantity,

$$|u - u_h|_{H^1(\Omega)} \ \leq \ |u - \tilde{u}_h|_{H^1(\Omega)} \ \leq \ C[\sum_{K \in T_n} (h_k |u|_{H^2(K)})^2]^{\frac{1}{2}} \tag{5.53}$$

or, in the $L_2(\Omega)$ norm,

$$\|u - u_h\|_{L_2(\Omega)} \ \leq \ \|u - \tilde{u}_h\|_{L_2(\Omega)} \ \leq \ Ch^2 |u|_{H^2(\Omega)}. \tag{5.54}$$

We can also provide error estimates for the potential gradients [40, 39, 41]

$$\|\nabla u - \nabla u_h\|_{L_\infty(\Omega)} \leq C\|\nabla u - \nabla \tilde{u}_h\|_{L_\infty(\Omega)} \leq C \max[h_K \max_{\alpha=2} \|D^\alpha u\|_{L_\infty(\Omega)}]. \tag{5.55}$$

We can now use the above estimates to create an adaptive algorithm. Suppose we want the accuracy of our finite element approximation to be within some tolerance, δ, of the true solution, i.e.

$$|u - u_h|_{H^1(\Omega)} \ \leq \ \delta \tag{5.56}$$

then, from the estimates above, we have

$$\sum_{K \in T_h} (h_k |u|_{H^2(K)})^2 \ \approx \ \frac{\delta^2}{C^2}. \tag{5.57}$$

Thus we redefine our mesh until,

$$\sum_{K \in T_h} (h_k |u_h|_{H^2(K)})^2 \ \leq \ \frac{\delta^2}{C^2} \tag{5.58}$$

and/or

$$h_K \max_{\alpha=2} \|D^\alpha u\|_{L_\infty(\Omega)} \leq \frac{\delta}{C}. \tag{5.59}$$

Here we sum over all the K elements in our tessellated geometry, T_h, and test to see if the error is less than our *normalized* tolerance. If the value of the error estimate exceeds the tolerance, then the element is "flagged" for further refinement.

5.5 Adaptive Error Control

While one goal of the error bounds is to understand the accuracy of our approximate solution, another goal is to use the error bound to develop an a posteriori error estimate that can be used to optimize the computational work needed to reach a certain level of accuracy (here we follow the development from Eriksson et al. [42].

More quantitatively, the basic goal of adaptive error control is, for a given tolerance δ, find a triangulation (or mesh, in general) T_h that requires the least amount of computational work to achieve

$$\|\nabla u - \nabla u_h\| \leq \delta \tag{5.60}$$

where $u_h \in \Omega_h$ is the finite element approximation corresponding to T_h. Measuring the computational work in terms of the number of nodes of the triangulation T_h and estimating the unknown error by the computable a posteriori error bound, leads to the problem of finding the triangulation T_h with the least number of nodes such that the corresponding finite element approximation u_h satisfies the stopping criteria

$$C_i \|hR(u_h)\| \leq \delta. \tag{5.61}$$

This comes from the following theorem

Theorem: There is an interpolation constant C_i that only depends on the minimal angle τ such that the error of the Galerkin finite element approximation u_h of the solution of the Poisson equation satisfies

$$\|\nabla u - \nabla u_h\| \leq C_i \|hR(u_h)\|, \tag{5.62}$$

where $R(u_h) = R_1(u_h) + R_2(u_h)$ with

$$R_1(u_h) = |f + \Delta u_h| \;\text{ on } K \in T_h \tag{5.63}$$

and

$$R_2(u_h) = \frac{1}{2} \max_{S \subset K} h_K^{-1} |\partial_S u_h| \;\text{ on } K \in T_h. \tag{5.64}$$

Here, $R_1(u_h)$ is the contribution to the total residual from the interior of the elements K and is defined on Ω_h by setting $R_1(u_h) = |f + \Delta u_h|$. $R_2(u_h)$ is the contribution to the residual from the jump of the normal derivative of u_h across edges.

The formula in (5.61) represents a nonlinear constrained optimization problem with u_h depending on T_h. We cannot usually solve this minimization problem analyitically, thus a solution has to be sought by an iterative method in which we start with a coarse initial mesh and successively modify the mesh by seeking to satisfy the stopping criterion in equation (5.61) with a minimal number of elements. To achieve this, we use the following adaptive algorithm:

1. Choose an initial triangulation $T_{h^{(0)}}$.

2. Given the j^{th} triangulation $T_{h^{(j)}}$ with mesh function $h^{(j)}$, compute the corresponding finite element approximation $u_h^{(j)}$.

3. Compute the corresponding residuals $R_1(u_h^{(j)})$ and $R_2(u_h^{(j)})$ and check whether or not equation (5.61) holds. If it does, stop.

4. Find a new triangulation $T_{h^{(j+1)}}$ with mesh function $h^{(j+1)}$ and with the minimal number of nodes such that equation (5.62) is satisfied and then proceed to step 2.

Not surprisingly, the success of this iterative strategy hinges on the mesh modification strategy used to perform step number 4. A natural strategy for error control based on the L_2 norm uses the *principle of equi-distribution* of the error in which we try to equalize the contribution from each element to the integral defining the L_2 norm. The rationale is that refining an element with large contribution to the error norm gives a large pay-off in terms of error reduction per new degree of freedom.

Quantitively, this means that the approximation computed on optimal mesh T_h in terms of computational work satisfies

$$\|\nabla e\|^2_{L_2(K)} \approx \frac{\delta^2}{M} \ \forall\ K \in T_h, \tag{5.65}$$

where M is the number of elements in T_h. Based on equation (5.61), we would therefore like to compute the triangulation at step number 4 so that

$$C_i^2(\|h^{(j+1)}R(u_h^{(j+1)})\|^2_{L_2(K)} \approx \frac{\delta^2}{M^{(j+1)}} \ \forall\ K \in T_{h^{(j+1)}}, \tag{5.66}$$

where $M^{(j+1)}$ is the number of elements in $T_{h^{(j+1)}}$. However, equation (5.66) is a nonlinear equation since we don't know $M^{(j+1)}$ and $u_h^{(j+1)}$ until we have chosen the triangulation. Hence, we replace (5.66) with

$$C_i^2(\|h^{(j+1)}R(u_h^{(j)})\|^2_{L_2(K)} \approx \frac{\delta^2}{M^{(j)}} \ \forall\ K \in T_{h^{(j+1)}}, \tag{5.67}$$

and use this formula to compute the new mesh size $h^{(j+1)}$.

There are lots of questions regarding the effect of replacing equation (5.62) with (5.63). How much larger is the right-hand side of equation (5.61) than the left hand side? Does the iterative process 1-4 converge to a solution of the minimization problem? How should the initial triangulation $T_{h^{(0)}}$ be chosen and how does this affect the convergence of the adaptive procedure? Is the approximation from equation (5.67) justified?

5.5.1 Example: the L-problem

We examine the performance of the adaptive control algorithm on the L-problem. We solve Laplace's equation with Dirichlet boundary conditions that are zero at the sides meeting at the origin and are non-zero on the other sides. So that we can see how well our adaptive algorithm does, we choose the non-zero boundary condition so that the exact solution is $u(r,\theta) = r^{2/3} sin(2\theta/3)$ in polar coordinates in which (r,θ) is centered at the origin. Thus we have an exact solution in which to compare our adaptive solution.

Since we know the exact solution, we can use an a priori error estimate to determine a mesh that gives the desired accurracy. We do this by combining the a priori error estimate in equation (5.30) and the principle of equidistribution of error to determine $h(r)$ so that $C_i\|hD^2u\| = \delta$ while keeping h as large as possible (and keeping the number of elements at a minimum). Since $D^2u(r) \approx r^{-4/3}$ as $h \leq r$ (that is, up to the elements touching the corner, we find that

$$(hr^{-4/3})^2h^2 \approx \frac{\delta^2}{M} \tag{5.68}$$

or

$$h^2 = \delta M^{-1/2} r^{4/3} \tag{5.69}$$

where M is the number of elements and h^2 measures the element area. To compute M from this relation, we note that

$$M \approx \int_\Omega h^{-2} dx, \tag{5.70}$$

since the number of elements per unit area is $O(h^{-2})$, which gives

$$M \approx M^{1/2}\delta^{-1} \int_\Omega r^{-4/3} dx, \tag{5.71}$$

One can prove that this integral is convergent (left to the reader to show), it then follows that $M \propto \delta^{-2}$, which implies that $h(r) \propto r^{1/3}\delta$.

To show that this relation agrees with our adaptive solution - we'll have to complete part of Assignment number 6 (of course).