Chapter 9

The Finite Element Method for 2D Problems

The FE procedure to solve 2D problems is the same as that for 1D problems, as the flow chart below demonstrates.

$$\begin{aligned} &\text{PDE} \longrightarrow \text{Integration by parts} \longrightarrow \text{weak form in } V : a(u,v) = L(v) \\ &\text{or } \min_{v\in V} F(v) \longrightarrow V_h \text{ (finite dimensional space and basis functions)} \\ &\longrightarrow a(u_h, v_h) = L(v_h) \longrightarrow u_h \text{ and error analysis.} \end{aligned}$$

9.1 The second Green's theorem and integration by parts in 2D

Let us first recall the 2D version of the well known Divergence Theorem in Cartesian coordinates.

Theorem 9.1. *If* $\mathbf{F} \in H^1(\Omega) \times H^1(\Omega)$, *then*

$$\iint_\Omega \nabla\cdot\mathbf{F}\, dS = \int_{\partial\Omega} \mathbf{F}\cdot\mathbf{n}\, ds\,, \tag{9.1}$$

where $dS = dx\, dy$ *is a surface element in* Ω, $\mathbf{n}$ *is the unit normal direction pointing outward at the boundary* $\partial\Omega$ *with line element* ds, *and* ∇ *is the gradient operator*

$$\nabla = \left(\frac{\partial}{\partial x}\,,\ \frac{\partial}{\partial y}\right)^T. \tag{9.2}$$

The second Green's theorem is the corollary where $\mathbf{F} = v\,\nabla u = \left[v\frac{\partial u}{\partial x}, v\frac{\partial u}{\partial y}\right]^T$. Thus since

$$\begin{aligned}\nabla\cdot\mathbf{F} &= \frac{\partial}{\partial x}\left(v\frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(v\frac{\partial u}{\partial y}\right)\\ &= \frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + v\frac{\partial^2 u}{\partial x^2} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y} + v\frac{\partial^2 u}{\partial y^2}\\ &= \nabla u\cdot\nabla v + v\,\Delta u\end{aligned}$$

where $\Delta u = \nabla\cdot\nabla u = u_{xx} + u_{yy}$, we obtain

$$\begin{aligned}\iint_\Omega \nabla\cdot\mathbf{F}\,dxdy &= \iint_\Omega (\nabla u\cdot\nabla v + v\,\Delta u)\,dxdy\\ &= \int_{\partial\Omega}\mathbf{F}\cdot\mathbf{n}\,ds\\ &= \int_{\partial\Omega} v\,\nabla u\cdot\mathbf{n}\,ds = \int_{\partial\Omega} v\,\frac{\partial u}{\partial n}\,ds\,,\end{aligned}$$

where writing $\mathbf{n} = (n_x, n_y)$ we have defined, $\frac{\partial u}{\partial n} = n_x\frac{\partial u}{\partial x} + n_y\frac{\partial u}{\partial y}$, the normal derivative derivative of u. This result immediately yields the formula for integration by parts in 2D:

Theorem 9.2. *If $u(x,y) \in H^2(\Omega)$ and $v(x,y) \in H^2(\Omega)$ where Ω is a bounded domain, then*

$$\iint_\Omega v\,\Delta u\,dxdy = \int_{\partial\Omega} v\frac{\partial u}{\partial n}\,ds - \iint_\Omega \nabla u\cdot\nabla v\,dxdy\,. \tag{9.3}$$

Note: the normal derivative $\partial u/\partial n$ is sometimes written more concisely as u_n.

Some important elliptic PDE in 2D Cartesian coordinates are:

$$\begin{aligned} u_{xx} + u_{yy} &= 0, && \text{Laplace equation,}\\ -u_{xx} - u_{yy} &= f(x,y), && \text{Poisson equation,}\\ -u_{xx} - u_{yy} + \lambda u &= f, && \text{generalized Helmholtz equation,}\\ u_{xxxx} + 2u_{xxyy} + u_{yyyy} &= 0, && \text{Bi-harmonic equation.}\end{aligned}$$

When $\lambda > 0$, the generalized Helmholtz equation is easier to solve than when $\lambda < 0$. Incidentally, the expressions involved in these PDE may also be abbreviated using the gradient operator ∇, *e.g.*, $u_{xx} + u_{yy} = \nabla\cdot\nabla u = \Delta u$ as mentioned before. We also recall that the general linear second order elliptic PDE has the form

$$a(x,y)u_{xx} + 2b(x,y)u_{xy} + c(x,y)u_{yy} + d(x,y)u_x + e(x,y)u_y + g(x,y)u = f(x,y)$$

with discriminant $b^2 - ac < 0$. A second order self-adjoint elliptic equation has the form

$$-\nabla\cdot(p(x,y)\nabla u) + q(x,y)u = f(x,y)\,. \tag{9.4}$$

9.1.1 Boundary conditions

In 2D, the domain boundary $\partial\Omega$ is one or several curves. We consider the following various linear BC.

- Dirichlet BC on the entire boundary, *i.e.*, $u(x,y)|_{\partial\Omega} = u_0(x,y)$ is given.
- Neumann BC on the entire boundary, *i.e.*, $\partial u/\partial n|_{\partial\Omega} = g(x,y)$ is given.
  In this case, the solution to a Poisson equation may not be unique or even exist, depending upon whether a compatibility condition is satisfied. Integrating the Poisson equation over the domain, we have

  $$\iint_\Omega f dxdy = \iint_\Omega \Delta u \, dxdy = \iint_\Omega \nabla\cdot\nabla u \, dxdy = \int_{\partial\Omega} u_n \, ds = \int_{\partial\Omega} g(x,y) \, ds = 0\,,$$

  which is the compatibility condition to be satisfied for the solution to exist. If a solution does exist, it is not unique as it is determined within an arbitrary constant.
- Mixed BC on the entire boundary, *i.e.*,

  $$\alpha(x,y)u(x,y) + \beta(x,y)\frac{\partial u}{\partial n} = \gamma(x,y)$$

  is given, where $\alpha(x,y)$, $\beta(x,y)$, and $\gamma(x,y)$ are known functions.
- Dirichlet, Neumann, and Mixed BC on some parts of the boundary.

9.1.2 Weak form of the second order self-adjoint elliptic PDE

Multiplying the self-adjoint equation (9.4) by a test function $v(x,y) \in H^1(\Omega)$, we have

$$\iint_\Omega \{-\nabla\cdot(p(x,y)\nabla u) + q(x,y)u\}\, v \, dxdy = \iint_\Omega f v \, dxdy\,;$$

and on using the formula for integration by parts the left-hand side becomes

$$\iint_\Omega (p\nabla u\cdot\nabla v + quv)\, dxdy - \int_{\partial\Omega} pvu_n \, ds\,,$$

so the weak form is

$$\iint_\Omega (p\nabla u\cdot\nabla v + quv)\ dxdy = \iint_\Omega fv dxdy + \int_{\partial\Omega_N} pg(x,y)v(x,y)\, ds \quad \forall v(x,y) \in V(\Omega)\,.$$

Here $\partial\Omega_N$ is the part of boundary where a Neumann boundary condition is applied; and the solution space resides in

$$V = \big\{v(x,y)\ ,\ v(x,y) = 0\ ,\ (x,y)\in\partial\Omega_D\ ,\ v(x,y)\in H^1(\Omega)\big\}\,, \tag{9.5}$$

where $\partial\Omega_D$ is the part of boundary where a Dirichlet boundary condition is applied.

9.1.3 Verification of the conditions of the Lax-Milgram Lemma

The bilinear form for (9.4) is

$$a(u,v) = \iint_\Omega \left(p\nabla u \cdot \nabla v + quv\right)\, dxdy\,, \tag{9.6}$$

and the linear form is

$$L(v) = \iint_\Omega fv\, dxdy \tag{9.7}$$

for a Dirichlet BC on the entire boundary. As before, we assume that

$$0 < p_{min} \le p(x,y) \le p_{max}\ ,\ 0 \le q(x) \le q_{max}\ ,\ p \in C(\Omega)\ ,\ q \in C(\Omega)\,.$$

We need the Poincaré inequality to prove the V-elliptic condition.

Theorem 9.3. *If $v(x,y) \in H_0^1(\Omega)$, $\Omega \subset R^2$, then*

$$\iint_\Omega v^2 dxdy \le C \iint_\Omega |\nabla v|^2\, dxdy\,, \tag{9.8}$$

where C is a constant.

Now we are ready to check the conditions of the Lax-Milgram Lemma.

1. It is obvious that $a(u,v) = a(v,u)$.
2. It is easy to see that
$$\begin{aligned} |a(u,v)| &\le \max\{p_{max}, q_{max}\} \left| \iint_\Omega \left(|\nabla u \cdot \nabla v| + |uv|\right) dxdy \right| \\ &= \max\{p_{max}, q_{max}\}\, |(|u|, |v|)_1| \\ &\le \max\{p_{max}, q_{max}\}\, \|u\|_1 \|v\|_1\,, \end{aligned}$$
so $a(u,v)$ is a continuous and bounded bilinear operator.
3. From the Poincaré inequality
$$\begin{aligned} |a(v,v)| &= \left| \iint_\Omega p\left(|\nabla v|^2 + qv^2\right) dxdy \right| \\ &\ge p_{min} \iint_\Omega |\nabla v|^2\, dxdy \\ &= \frac{1}{2} p_{min} \iint_\Omega |\nabla v|^2\, dxdy + \frac{1}{2} p_{min} \iint_\Omega |\nabla v|^2\, dxdy \\ &\ge \frac{1}{2} p_{min} \iint_\Omega |\nabla v|^2\, dxdy + \frac{p_{min}}{2C} \iint_\Omega |v|^2\, dxdy \\ &\ge \frac{1}{2} p_{min} \min\left\{1,\ \frac{1}{C}\right\} \|v\|_1^2\,, \end{aligned}$$
therefore $a(u,v)$ is V-elliptic.
4. Finally, we show that $L(v)$ is continuous:
$$|L(v)| = |(f,v)_0| \le \|f\|_0 \|v\|_0 \le \|f\|_0 \|v\|_1\,.$$

Consequently, the solutions to the weak form and the minimization form are unique and bounded in $H_0^1(\Omega)$.

9.2 Triangulation and basis functions

The general procedure of the FE method is the same for any dimension, and the Galerkin FE method involves the following main steps.

- Generate a triangulation over the domain. Usually the triangulation is composed of either triangles or quadrilaterals (rectangles). There are a number of mesh generation software packages available, *e.g.*, the Matlab PDE toolbox from Mathworks, Triangle from Carnegie Mellon University, *etc.* Some are available through the Internet.
- Construct basis functions over the triangulation. We only consider the conforming FE method in this book.
- Assemble the stiffness matrix and the load vector element by element, using either the Galerkin method (the weak form) or the Ritz method (the minimization form).
- Solve the system of equations.
- Do the error analysis.

9.2.1 Triangulation and mesh parameters

Given a general domain, we can approximate the domain by a polygon and then generate a triangulation over the polygon, and we can refine the triangulation if necessary. A simple approach is the mid-point rule by connecting all the middle points of three sides of existing triangles to get a refined mesh.

A triangulation usually has the mesh parameters

$$\begin{aligned}
\Omega_p: &\quad \text{polygonal region} = K_1 \cup K_2 \cup K_3 \cdots \cup K_{nelem}\,,\\
K_j: &\quad \text{are non-overlapping triangles, } j = 1, 2, \cdots nelem\,,\\
N_i: &\quad \text{are nodal points, } i = 1, 2, \cdots nnode\,,\\
h_j: &\quad \text{the longest side of } K_j\,,\\
\rho_j: &\quad \text{the diameter of the circle inscribed in } K_j \text{ (encircle)},\\
h: &\quad \text{the largest of all } h_j, \quad h = \max\{h_j\}\,,\\
\rho: &\quad \text{the smallest of all } \rho_j, \quad \rho = \min\{\rho_j\}\,,
\end{aligned}$$

with

$$1 \geq \frac{\rho_j}{h_j} \geq \beta > 0\,,$$

where the constant β is a measurement of the triangulation quality. The larger the β, the better the quality of the triangulation. Given a triangulation, a node is also the vertex of all adjacent triangles. We do not discuss hanging nodes here.

9.2.2 FE space of piecewise linear functions over a mesh

For linear second order elliptic PDE, we know the solution space is in the $H^1(\Omega)$. Unlike the 1D case, an element $v(x, y)$ in $H^1(\Omega)$ may not be continuous under the Sobolev embedding

theorem. However, in practice most solutions are indeed continuous, especially for second order PDE with certain regularities. Thus, we still look for a solution in the continuous function space $C^0(\Omega)$. Let us first consider how to construct piecewise linear functions over a mesh with the Dirichlet BC

$$u(x,y)|_{\partial\Omega} = 0\,.$$

Given a triangulation, we define

$$\begin{aligned} V_h \quad = \quad & \Big\{ v(x,y) \ \text{ is continuous in } \Omega \text{ and piecewise linear over each } K_j\,, \\ & \qquad v(x,y)|_{\partial\Omega} = 0 \ \Big\}\,. \end{aligned} \tag{9.9}$$

We need to determine the dimension of this space and construct a set of basis functions. On each triangle, a linear function has the form

$$v_h(x,y) = \alpha + \beta x + \gamma y\,, \tag{9.10}$$

where α, β and γ are constants (three free parameters). Let

$$P_k = \{\, p(x,y)\,, \text{ a polynomial of degree of } \ k\,\}\,. \tag{9.11}$$

We have the following theorem.

Theorem 9.4.

1. *A linear function $p_1(x,y) = \alpha+\beta x+\gamma y$ defined on a triangle is uniquely determined by its values at the three vertices.*
2. *If $p_1(x,y) \in P_1$ and $p_2(x,y) \in P_1$ are such that $p_1(A) = p_2(A)$ and $p_1(B) = p_2(B)$, where A and B are two points in the xy-plane, then $p_1(x,y) \equiv p_2(x,y)$, $\forall (x,y) \in I_{AB}$, where I_{AB} is the line segment between A and B.*

Proof: Assume the vertices of the triangle are (x_i, y_i), $i = 1, 2, 3$. The linear function takes the value v_i at the vertices, *i.e.*,

$$p(x_i, y_i) = v_i\,,$$

so we have the three equations

$$\begin{aligned} \alpha + \beta x_1 + \gamma y_1 &= v_1\,, \\ \alpha + \beta x_2 + \gamma y_2 &= v_2\,, \\ \alpha + \beta x_3 + \gamma y_3 &= v_3\,. \end{aligned}$$

The determinant of this linear algebraic system is

$$det \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix} = \pm 2 \text{ area of the triangle} \neq 0 \ \text{ since } \ \frac{\rho_j}{h_j} \geq \beta > 0\,, \tag{9.12}$$

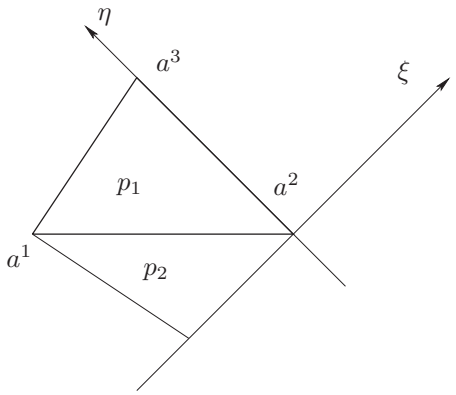


Figure 9.1. *A diagram of a triangle with three vertices a^1, a^2, and a^3; an adjacent triangle with a common side; and the local coordinate system in which a^2 is the origin and a^2a^3 is the η axis.*

hence the linear system of equations has a unique solution.

Now let us prove the second part of the theorem. Suppose that the equation of the line segment is

$$l_1 x + l_2 y + l_3 = 0\,, \quad l_1^2 + l_2^2 \neq 0\,.$$

We can solve for x or for y:

$$x = -\frac{l_2 y + l_3}{l_1} \quad \text{if} \quad l_1 \neq 0\,,$$

$$\text{or} \quad y = -\frac{l_1 x + l_3}{l_2} \quad \text{if} \quad l_2 \neq 0\,.$$

Without loss of generality, let us assume $l_2 \neq 0$ such that

$$\begin{aligned} p_1(x,y) &= \alpha + \beta x + \gamma y \\ &= \alpha + \beta x - \frac{l_1 x + l_3}{l_2}\gamma \\ &= \left(\alpha - \frac{l_3}{l_2}\gamma\right) + \left(\beta - \frac{l_1}{l_2}\gamma\right) x \\ &= \alpha_1 + \beta_1 x\,. \end{aligned}$$

Similarly, we have

$$p_2(x,y) = \bar{\alpha}_1 + \bar{\beta}_1 x\,.$$

Since $p_1(A) = p_2(A)$ and $p_1(B) = p_2(B)$,

$$\begin{aligned} \alpha_1 + \beta_1 x_1 &= p(A)\,, \qquad \bar{\alpha}_1 + \bar{\beta}_1 x_1 = p(A)\,, \\ \alpha_1 + \beta_1 x_2 &= p(B)\,, \qquad \bar{\alpha}_1 + \bar{\beta}_1 x_2 = p(B)\,, \end{aligned}$$

where both of the linear system of algebraic equations have the same coefficient matrix

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \end{bmatrix}$$

that is non-singular since $x_1 \neq x_2$ (because points A and B are distinct). Thus we conclude that $\alpha_1 = \bar{\alpha}_1$ and $\beta_1 = \bar{\beta}_1$, so the two linear functions have the same expression along the line segment, *i.e.*, they are identical along the line segment.

Corollary 9.5. *A piecewise linear function in $C^0(\Omega)$ over a triangulation (a set of non-overlapping triangles) is uniquely determined by its values at the vertices.*

Theorem 9.6. *The dimension of the finite dimensional space composed of piecewise linear functions in $C^0(\Omega) \cap H^1(\Omega)$ over a triangulation for (9.4) is the number of interior nodal points plus the number of nodal points on the boundary where the natural BC are imposed (Neumann and mixed boundary conditions).*

Example 9.1. *Given the triangulation shown in Fig. 9.2, a piecewise continuous function $v_h(x, y)$ is determined by its values on the vertices of all triangles, more precisely, $v_h(x, y)$ is determined from*

$$\begin{aligned}
&(0, 0, v(N_1)), \quad (x, y) \in K_1, \qquad (0, v(N_2), v(N_1)), \quad (x, y) \in K_2, \\
&(0, 0, v(N_2)), \quad (x, y) \in K_3, \qquad (0, 0, v(N_2)), \quad (x, y) \in K_4, \\
&(0, v(N_3), v(N_2)), \quad (x, y) \in K_5, \qquad (0, 0, v(N_3)), \quad (x, y) \in K_6, \\
&(0, v(N_1), v(N_3)), \quad (x, y) \in K_7, \qquad (v(N_1), v(N_2), v(N_3)), \quad (x, y) \in K_8.
\end{aligned}$$

Note that although three values of the vertices are the same, like the values for K_3 and K_4, the geometries are different, hence, the functions will likely have different expressions on different triangles.

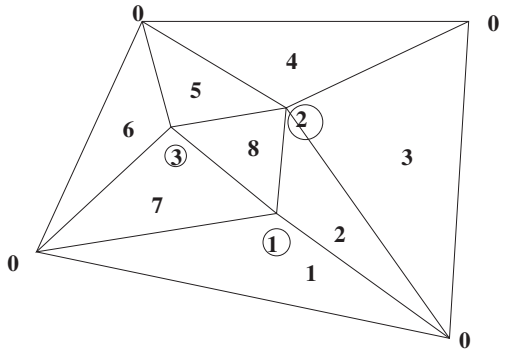


Figure 9.2. *A diagram of a simple triangulation with zero BC.*

9.2.3 Global basis functions

A global basis function in the continuous piecewise linear space is defined as

$$\phi_i(N_j) = \begin{cases} 1 & \text{if } i = j\,, \\ 0 & \text{otherwise}\,, \end{cases} \tag{9.13}$$

where N_j are nodal points. The shape looks like a "tent" without a door, and its support is the union of the triangles surrounding the node N_i, *cf.*, Fig. 9.3, where Fig. 9.3 (a) is the mesh plot of the global basis function, and Fig. 9.3 (b) is the plot of a triangulation and the contour plot of the global basis function centered at a node. The basis function is piecewise linear and it is supported only in the surrounding triangles.

(a) (b)

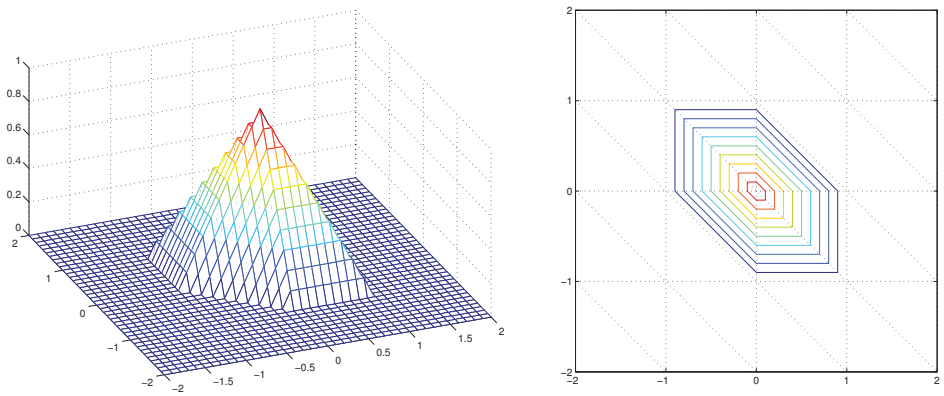


Figure 9.3. *A global basis function ϕ_j. (a) the mesh plot; (b) the triangulation and the contour plot of the global basis function.*

It is almost impossible to give a closed form of a global basis function except for some very special geometries (*cf.*, the example in the next section). However, it is much easier to write down the shape function.

Example 9.2. *Let us consider a Poisson equation and a uniform mesh, as an example to demonstrate the piecewise linear basis functions and the FE method:*

$$-(u_{xx} + u_{yy}) = f(x,y)\,, \quad (x,y) \in [a,b] \times [c,d]\,,$$
$$u(x,y)|_{\partial\Omega} = 0\,.$$

We know how to use the standard central FD scheme with the five point stencil to solve the Poisson equation. With some manipulations, the linear system of equations on using the FE method with a uniform triangulation (*cf.*, Fig. 9.4) proves to be the same as that obtained from the FD method.

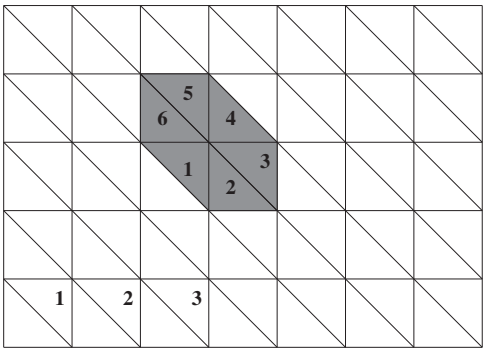


Figure 9.4. *A uniform triangulation defined on a rectangular domain.*

Given a uniform triangulation as shown in Fig. 9.4, if we use row-wise ordering for the nodal points

$$(x_i, y_j)\,, \quad x_i = ih\,, \quad y_j = jh\,, \quad h = \frac{1}{n}\,, \quad i = 1, 2, \cdots, m-1\,, \;\; j = 1, 2, \cdots, n-1\,,$$

then the global basis function defined at $(x_i, y_j) = (ih, jh)$ are

$$\phi_{j(n-1)+i} = \begin{cases} \dfrac{x-(i-1)h+y-(j-1)h}{h} - 1 & \text{Region 1} \\ \dfrac{y-(j-1)h}{h} & \text{Region 2} \\ \dfrac{h-(x-ih)}{h} & \text{Region 3} \\ 1 - \dfrac{x-ih+y-jh}{h} & \text{Region 4} \\ \dfrac{h-(y-jh)}{h} & \text{Region 5} \\ \dfrac{x-(i-1)h}{h} & \text{Region 6} \\ 0 & \text{otherwise}\,. \end{cases}$$

If $m = n = 3$, there are 9 interior nodal points such that the stiffness matrix is a 9×9

matrix:

$$A = \begin{bmatrix} * & * & 0 & * & 0 & 0 & 0 & 0 & 0 \\ * & * & * & o & * & 0 & 0 & 0 & 0 \\ 0 & * & * & 0 & o & * & 0 & 0 & 0 \\ * & o & 0 & * & * & 0 & * & 0 & 0 \\ 0 & * & o & * & * & * & o & * & 0 \\ 0 & 0 & * & 0 & * & * & 0 & o & * \\ 0 & 0 & 0 & * & o & 0 & * & * & 0 \\ 0 & 0 & 0 & 0 & * & o & * & * & * \\ 0 & 0 & 0 & 0 & 0 & * & 0 & * & * \end{bmatrix},$$

where '$*$' stands for the nonzero entries and 'o' happens to be zero. Generally, the stiffness matrix is block tri-diagonal:

$$A = \begin{bmatrix} B & -I & 0 & & & \\ -I & B & -I & & & \\ & & \cdots & \cdots & & \\ & & & \cdots & \cdots & \\ & & & -I & B & -I \\ & & & & -I & B \end{bmatrix}, \text{ where } B = \begin{bmatrix} 4 & -1 & 0 & & & \\ -1 & 4 & -1 & & & \\ & & \cdots & \cdots & & \\ & & & \cdots & \cdots & \\ & & & -1 & 4 & -1 \\ & & & & -1 & 4 \end{bmatrix}$$

and I is the identity matrix. The component of the load vector F_i can be approximated as

$$\iint_D f(x,y)\phi_i dxdy \simeq f_{ij} \iint_D \phi_i \, dxdy = h^2 f_{ij} \,,$$

so after dividing by h^2 we get the same system of equations as in the FD scheme, namely,

$$-\frac{U_{i-1,j} + U_{i+1,j} + U_{i,j-1} + U_{i,j+1} - 4U_{ij}}{h^2} = f_{ij} \,,$$

with the same ordering.

9.2.4 The interpolation function and error analysis

We know that the FE solution u_h is the best solution in terms of the energy norm in the finite dimensional space V_h, *i.e.*, $\|u - u_h\|_a \leq \|u - v_h\|_a$, assuming that u is the solution to the weak form. However, this does not give a quantitative estimate for the FE solution, and we may wish to have a more precise error estimate in terms of the solution information and the mesh size h. This can be done through the interpolation function, for which an error estimate is often available from the approximation theory. Note that the solution information appears as part of the error constants in the error estimates, even though

the solution is unknown. We will use the mesh parameters defined on page 203 in the discussion here.

Definition 9.7. *Given a triangulation of T_h, let $K \in T_h$ be a triangle with vertices a^i, $i = 1, 2, 3$. The interpolation function for a function $v(x, y)$ on the triangle is defined as*

$$v_I(x, y) = \sum_{i=1}^{3} v(a^i)\phi_i(x, y)\,, \tag{9.14}$$

where $\phi_i(x, y)$ is the piecewise linear function that satisfies $\phi_i(a^j) = \delta_i^j$ (with δ_i^j being the Kronecker delta). A global interpolation function is defined as

$$v_I(x, y) = \sum_{i=1}^{nnode} v(a^i)\phi_i(x, y)\,, \tag{9.15}$$

where the a^i are all nodal points and $\phi_i(x, y)$ is the global basis function centered at a^i.

Theorem 9.8. *If $v(x, y) \in C^2(K)$, then we have an error estimate for the interpolation function on a triangle K,*

$$\|v - v_I\|_\infty \leq 2h^2 \max_{|\alpha|=2} \|D^\alpha v\|_\infty\,, \tag{9.16}$$

where h is the longest side. Furthermore, we have

$$\max_{|\alpha|=1} \|D^\alpha (v - v_I)\|_\infty \leq \frac{8h^2}{\rho} \max_{|\alpha|=2} \|D^\alpha v\|_\infty\,. \tag{9.17}$$

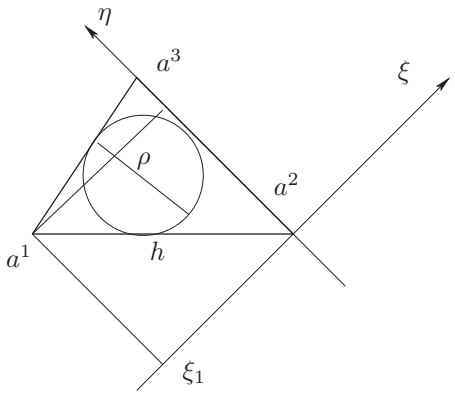


Figure 9.5. *A diagram used to prove Theorem 9.8.*

Proof: From the definition of the interpolation function and the Taylor expansion

of $v(a^i)$ at (x,y), we have

$$\begin{aligned} v_I(x,y) &= \sum_{i=1}^{3} v(a^i)\phi_i(x,y) \\ &= \sum_{i=1}^{3} \phi_i(x,y)\left(v(x,y) + \frac{\partial v}{\partial x}(x,y)(x_i - x) + \frac{\partial v}{\partial y}(x,y)(y_i - y) + \right. \\ &\quad \left. \frac{1}{2}\frac{\partial^2 v}{\partial x^2}(\xi,\eta)(x_i - x)^2 + \frac{\partial^2 v}{\partial x \partial y}(\xi,\eta)(x_i - x)(y_i - y) + \frac{1}{2}\frac{\partial^2 v}{\partial y^2}(\xi,\eta)(y_i - y)^2 \right) \\ &= \sum_{i=1}^{3} \phi_i(x,y)v(x,y) + \sum_{i=1}^{3} \phi_i(x,y)\left(\frac{\partial v}{\partial x}(x,y)(x_i - x) + \frac{\partial v}{\partial y}(x,y)(y_i - y) \right) \\ &\qquad + R(x,y)\,, \end{aligned}$$

where (ξ,η) is a point in the triangle K. It is easy to show that

$$|R(x,y)| \le 2h^2 \max_{|\alpha|=2} \|D^\alpha v\|_\infty \sum_{i=1}^{3} |\phi_i(x,y)| = 2h^2 \max_{|\alpha|=2} \|D^\alpha v\|_\infty\,,$$

since $\phi(x,y) \ge 0$ and $\sum_{i=1}^{3} \phi_i(x,y) = 1$. If we take $v(x,y) = 1$, which is a linear function, then $\partial v/\partial x = \partial v/\partial y = 0$ and $\max_{|\alpha|=2} \|D^\alpha v\|_\infty = 0$. The interpolation is simply the function itself, since it uniquely determined by the values at the vertices of T, hence

$$v_I(x,y) = v(x,y) = \sum_{i=1}^{3} v(a^i)\phi_i(x,y) = \sum_{i=1}^{3} \phi_i(x,y) = 1\,. \tag{9.18}$$

If we take $v(x,y) = d_1 x + d_2 y$, which is also a linear function, then $\partial v/\partial x = d_1$, $\partial v/\partial y = d_2$, and $\max_{|\alpha|=2} \|D^\alpha v\|_\infty = 0$. The interpolation is again simply the function itself, since it uniquely determined by the values at the vertices of K. Thus from the previous Taylor expansion and the identity $\sum_{i=1}^{3} \phi_i(x,y) = 1$, we have

$$v_I(x,y) = v(x,y) = v(x,y) + \sum_{i=1}^{3} \phi_i(x,y)\,(d_1(x_i - x) + d_2(y_i - y)) = v(x,y)\,, \tag{9.19}$$

hence $\sum_{i=1}^{3} \phi_i(x,y)\,(d_1(x_i - x) + d_2(y_i - y)) = 0$ for any d_1 and d_2 — i.e. the linear part in the expansion is the interpolation function. Consequently, for a general function $v(x,y) \in C^2(K)$ we have

$$v_I(x,y) = v(x,y) + R(x,y)\,, \qquad \|v - v_I\|_\infty \le 2h^2 \max_{|\alpha|=2} \|D^\alpha v\|_\infty\,,$$

which completes the proof of the first part of the theorem.

To prove the second part concerning the error estimate for the gradient, choose a point (x_0,y_0) inside the triangle K and apply the Taylor expansion at (x_0,y_0) to get

$$\begin{aligned} v(x,y) &= v(x_0,y_0) + \frac{\partial v}{\partial x}(x_0,y_0)(x - x_0) + \frac{\partial v}{\partial y}(x_0,y_0)(y - y_0) + R_2(x,y), \\ &= p_1(x,y) + R_2(x,y), \qquad |R_2(x,y)| \le 2h^2 \max_{|\alpha|=2} \|D^\alpha v\|_\infty\,. \end{aligned}$$

Rewriting the interpolation function $v_I(x,y)$ as

$$v_I(x,y) = v(x_0,y_0) + \frac{\partial v}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial v}{\partial y}(x_0,y_0)(y-y_0) + R_1(x,y),$$

where $R_1(x,y)$ is a linear function of x and y, we have

$$v_I(a^i) = p_1(a^i) + R_1(a^i),\ i = 1,2,3\,,$$

from the definition above. On the other hand, $v_I(x,y)$ is the interpolation function, such that also

$$v_I(a^i) = v(a^i) = p_1(a^i) + R_2(a^i), \qquad i = 1,2,3\,.$$

Since $p_1(a^i) + R_1(a^i) = p_1(a^i) + R_2(a^i)$, it follows that $R_1(a^i) = R_2(a^i)$, *i.e.*, $R_1(x,y)$ is the interpolation function of $R_2(x,y)$ in the triangle K, and we have

$$R_1(x,y) = \sum_{i=1}^{3} R_2(a^i)\phi_i(x,y)\,.$$

With this equality and on differentiating

$$v_I(x,y) = v(x_0,y_0) + \frac{\partial v}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial v}{\partial y}(x_0,y_0)(y-y_0) + R_1(x,y)$$

with respect to x, we get

$$\frac{\partial v_I}{\partial x}(x,y) = \frac{\partial v}{\partial x}(x_0,y_0) + \frac{\partial R_1}{\partial x}(x,y) = \frac{\partial v}{\partial x}(x_0,y_0) + \sum_{i=1}^{3} R_2(a^i)\frac{\partial \phi_i}{\partial x}(x,y)\,.$$

Applying the Taylor expansion for $\partial v(x,y)/\partial x$ at (x_0,y_0) gives

$$\frac{\partial v}{\partial x}(x,y) = \frac{\partial v}{\partial x}(x_0,y_0) + \frac{\partial^2 v}{\partial x^2}(\bar{x},\bar{y})(x-x_0) + \frac{\partial^2 v}{\partial x \partial y}(\bar{x},\bar{y})(y-y_0)\,,$$

where $(\bar{x},\bar{y})$ is a point in the triangle K. From the last two equalities,

$$\left|\frac{\partial v}{\partial x} - \frac{\partial v_I}{\partial x}\right| = \left|\frac{\partial^2 v}{\partial x^2}(\bar{x},\bar{y})(x-x_0) + \frac{\partial^2 v}{\partial x \partial y}(\bar{x},\bar{y})(y-y_0) - \sum_{i=1}^{3} R_2(a^i)\frac{\partial \phi_i}{\partial x}\right|$$

$$\leq \max_{|\alpha|=2} \|D^\alpha v\|_\infty \left(2h + 2h^2 \sum_{i=1}^{3}\left|\frac{\partial \phi_i}{\partial x}\right|\right).$$

It remains to prove is that $|\partial\phi_i/\partial x| \leq 1/\rho$, $i = 1,2,3$. We take $i = 1$ as an illustration, and use a shift and rotation coordinate transform such that a^2a^3 is the η axis and a^2 is the origin (*cf.* Fig. 9.5):

$$\xi = (x-x_2)\cos\theta + (y-y_2)\sin\theta\,,$$
$$\eta = -(x-x_2)\sin\theta + (y-y_2)\cos\theta\,.$$

Then $\phi_1(x,y) = \phi_1(\xi,\eta) = C\xi = \xi/\xi_1$, where ξ_1 is the ξ coordinate in the (ξ,η) coordinate system, such that

$$\left|\frac{\partial \phi_1}{\partial x}\right| = \left|\frac{\partial \phi_1}{\partial \xi}\cos\theta - \frac{\partial \phi_1}{\partial \eta}\sin\theta\right| \leq \left|\frac{1}{\xi_1}\cos\theta\right| \leq \frac{1}{|\xi_1|} \leq \frac{1}{\rho}\,.$$

The same estimate applies to $\partial\phi_i/\partial x$, $i = 2, 3$, so finally we have

$$\left|\frac{\partial v}{\partial x} - \frac{\partial v_I}{\partial x}\right| \leq \max_{|\alpha|=2} \|D^\alpha v\|_\infty \left(2h + \frac{6h^2}{\rho}\right) \leq \frac{8h^2}{\rho} \max_{|\alpha|=2} \|D^\alpha v\|_\infty ,$$

from the fact that $\rho \leq h$. Similarly, we may obtain the same error estimate for $\partial v_I/\partial y$. □

Corollary 9.9. *Given a triangulation of T_h, we have the following error estimates for the interpolation function:*

$$\|v - v_I\|_{L^2(T_h)} \leq C_1 h^2 \|v\|_{H^2(T_h)} , \ \|v - v_I\|_{H^1(T_h)} \leq C_2 h \|v\|_{H^2(T_h)} , \tag{9.20}$$

where C_1 and C_2 are constants.

9.2.5 Error estimates of the FE solution

Let us now recall the 2D Sturm-Liouville problem in a bounded domain Ω:

$$-\nabla \cdot (p(x,y)\nabla u(x,y)) + q(x,y)u(x,y) = f(x,y) , \quad (x,y) \in \Omega ,$$
$$u(x,y)_{\partial\Omega} = u_0(x,y) ,$$

where $u_0(x,y)$ is a given function, *i.e.*, a Dirichlet BC is prescribed. If we assume that $p, q \in C(\Omega)$, $p(x,y) \geq p_0 > 0$, $q(x,y) \geq 0$, $f \in L^2(\Omega)$ and the boundary $\partial\Omega$ is smooth (in C^1), then we know the weak form has a unique solution and the energy norm $\|v\|_a$ is equivalent to the H^1 norm $\|v\|_1$. Furthermore, we know that the solution $u(x,y) \in C^2(\Omega)$. Given a triangulation T_h with a polygonal approximation to the outer boundary $\partial\Omega$, let V_h be the piecewise linear function space over the triangulation T_h and u_h be the FE solution. Then we have the following theorem for the error estimates.

Theorem 9.10.

$$\|u - u_h\|_a \leq C_1 h \|u\|_{H^2(T_h)} , \qquad \|u - u_h\|_{H^1(T_h)} \leq C_2 h \|u\|_{H^2(T_h)} , \tag{9.21}$$
$$\|u - u_h\|_{L^2(T_h)} \leq C_3 h^2 \|u\|_{H^2(T_h)} , \qquad \|u - u_h\|_\infty \leq C_4 h^2 \|u\|_{H^2(T_h)} , \tag{9.22}$$

where C_i are constants.

Sketch of the proof. Since the FE solution is the best solution in the energy norm, we have

$$\|u - u_h\|_a \leq \|u - u_I\|_a \leq \bar{C}_1 \|u - u_I\|_{H^1(T_h)} \leq \bar{C}_1 \bar{C}_2 h \|u\|_{H^2(T_h)} ,$$

because the energy norm is equivalent to the H^1 norm. Furthermore, because of the equivalence we get the estimate for the H^1 norm as well. The error estimates for the L^2 and L^∞ norm are not trivial in 2D, and the reader may care to consult other advanced textbooks on FE methods.

9.3 Transforms, shape functions, and quadrature formulas

Any triangle with nonzero area can be transformed to the right-isosceles master triangle, or standard triangle $\triangle$, *cf.* the right diagram in Fig. 9.6. There are three nonzero basis functions over this standard triangle $\triangle$, namely,

$$\psi_1(\xi,\eta) = 1-\xi-\eta\,, \tag{9.23}$$

$$\psi_2(\xi,\eta) = \xi\,, \tag{9.24}$$

$$\psi_3(\xi,\eta) = \eta\,. \tag{9.25}$$

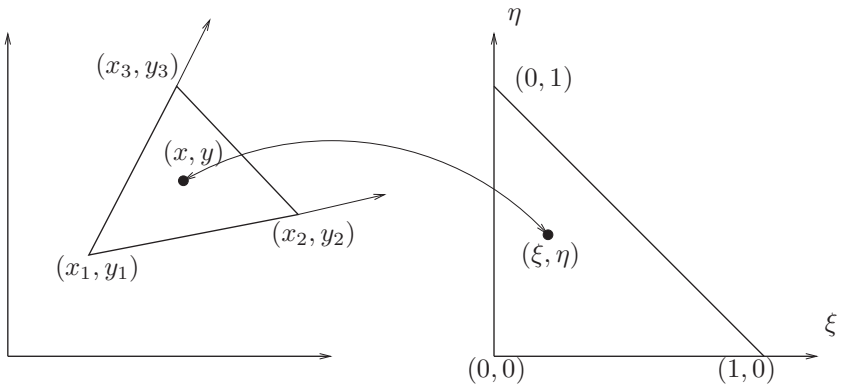


Figure 9.6. *The linear transform from an arbitrary triangle to the standard triangle (master element) and the inverse map.*

The linear transform from a triangle with vertices (x_1, y_1), (x_2, y_2) and (x_3, y_3) arranged in the counter-clockwise direction to the master triangle $\triangle$ is

$$x = \sum_{j=1}^{3} x_j \psi_j(\xi,\eta)\,, \qquad y = \sum_{j=1}^{3} y_j \psi_j(\xi,\eta)\,, \tag{9.26}$$

or

$$\xi = \frac{1}{2A_e}\Big((y_3-y_1)(x-x_1) - (x_3-x_1)(y-y_1)\Big)\,, \tag{9.27}$$

$$\eta = \frac{1}{2A_e}\Big(-(y_2-y_1)(x-x_1) + (x_2-x_1)(y-y_1)\Big)\,, \tag{9.28}$$

where A_e is the area of the triangle that can be calculated using the formula in (9.12).

9.3.1 Quadrature formulas

In the assembling process, we need to evaluate the double integrals

$$\iint_{\Omega_e} q(x,y)\phi_i(x,y)\phi_j(x,y)\,dxdy = \iint_{\triangle} q(\xi,\eta)\,\psi_i(\xi,\eta)\psi_j(\xi,\eta)\left|\frac{\partial(x,y)}{(\partial\xi,\eta)}\right| d\xi d\eta\,,$$
$$\iint_{\Omega_e} f(x,y)\phi_j(x,y)\,dxdy = \iint_{\triangle} f(\xi,\eta)\,\psi_j(\xi,\eta)\,\left|\frac{\partial(x,y)}{(\partial\xi,\eta)}\right| d\xi d\eta\,,$$
$$\iint_{\Omega_e} p(x,y)\nabla\phi_i\cdot\nabla\phi_j\,dxdy = \iint_{\triangle} p(\xi,\eta)\,\nabla_{(x,y)}\psi_i\cdot\nabla_{(x,y)}\psi_j\,\left|\frac{(\partial(x,y)}{\partial\xi,\eta)}\right| d\xi d\eta\,.$$

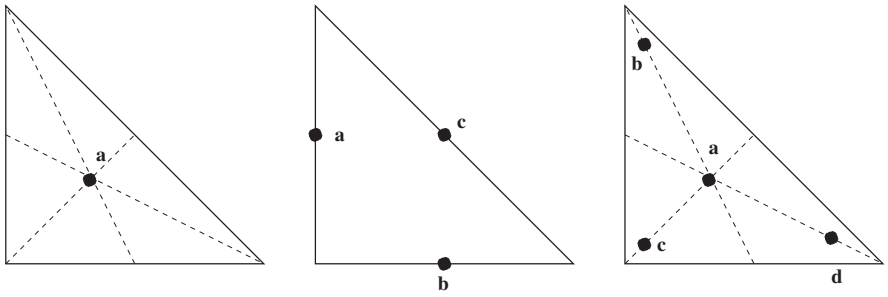


Figure 9.7. *A diagram of the quadrature formulas in 2D with one, three and four quadrature points, respectively.*

A quadrature formula has the form

$$\iint_{S_\triangle} g(\xi,\eta)d\xi d\eta = \sum_{k=1}^{L} w_k\, g(\xi_k,\eta_k)\,, \tag{9.29}$$

where $S_\triangle$ is the standard right triangle and L is the number of points involved in the quadrature. Below we list some commonly used quadrature formulas in 2D using one, three and four points. The geometry of the points are illustrated in Fig. 9.7, and the coordinates of the points and the weights are given in Table 9.1. It is notable that only the three-point quadrature formula is closed, since the three points are on the boundary of the triangle, and the other quadrature formulas are open.

9.4 Some implementation details

The procedure is essentially the same as in the 1D case, but some details are slightly different.

9.4.1 Description of a triangulation

A triangulation is determined by its elements and nodal points. We use the following notation:

Table 9.1. *Quadrature points and weights corresponding to the geometry in Fig. 9.7.*

L	Points	(ξ_k, η_k)	w_k
1	a	$\left(\frac{1}{3}, \frac{1}{3}\right)$	$\frac{1}{2}$
3	a	$\left(0, \frac{1}{2}\right)$	$\frac{1}{6}$
	b	$\left(\frac{1}{2}, 0\right)$	$\frac{1}{6}$
	c	$\left(\frac{1}{2}, \frac{1}{2}\right)$	$\frac{1}{6}$
4	a	$\left(\frac{1}{3}, \frac{1}{3}\right)$	$-\frac{27}{96}$
	b	$\left(\frac{2}{15}, \frac{11}{15}\right)$	$\frac{25}{96}$
	c	$\left(\frac{2}{15}, \frac{2}{15}\right)$	$\frac{25}{96}$
	d	$\left(\frac{11}{15}, \frac{2}{15}\right)$	$\frac{25}{96}$

- Nodal points: N_i, $(x_1, y_1), (x_2, y_2), \cdots, (x_{nnode}, y_{nnode})$, *i.e.*, we assume there are *nnode* nodal points.

- Elements: K_i, $K_1, K_2, \cdots, K_{nelem}$, *i.e.*, we assume there are *nelem* elements.

- A 2D array *nodes* is used to describe the relation between the nodal points and the elements: $nodes(3, nelem)$. The first index is the index of nodal point in an element, usually in the counter-clockwise direction, and the second index is the index of the element.

Example 9.3. *Below we show the relation between the index of the nodal points and elements, and its relations,* cf. *also Fig. 9.8.*

$$
\begin{aligned}
nodes(1,1) &= 5\,, \quad (x_5, y_5) = (0, h)\,,\\
nodes(2,1) &= 1\,, \quad (x_1, y_1) = (0, 0)\,,\\
nodes(3,1) &= 6\,, \quad (x_6, y_6) = (h, h)\,,
\end{aligned}
$$

$$
\begin{aligned}
nodes(1,10) &= 7\,, \quad (x_7, y_7) = (2h, h)\,,\\
nodes(2,10) &= 11\,, \quad (x_{11}, y_{11}) = (2h, 2h)\,,\\
nodes(3,10) &= 6\,, \quad (x_6, y_6) = (h, h)\,.
\end{aligned}
$$

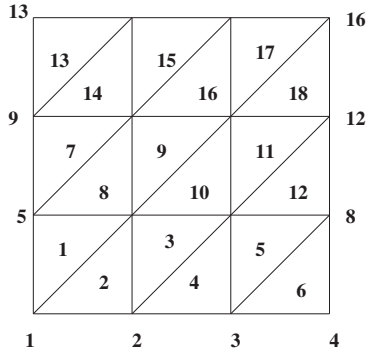


Figure 9.8. *A simple triangulation with the row-wise natural ordering.*

9.4.2 Outline of the FE algorithm using the piecewise linear basis functions

The main assembling process is the following loop.

```
for nel = 1:nelem
   i1 = nodes(1,nel);          % (x(i1),y(i1)), get nodal points
   i2 = nodes(2,nel);          % (x(i2),y(i2))
   i3 = nodes(3,nel);          % (x(i3),y(i3))
     ..............
```

- Computing the local stiffness matrix and the load vector.

```
ef=zeros(3,1);
ek = zeros(3,3);
for l=1:nq              % nq is the number of quadrature points.
  [xi_x(l),eta_y(l)] = getint,     % Get a quadrature point.
  [psi,dpsi] = shape(xi_x(l),eta_y(l));
  [x_l,y_l]  = transform,          % Get (x,y) from (\xi_x(l), \eta_y(l))
  [xk,xq,xf] = getmat(x_l,y_l);    % Get the material
```

```
                                %coefficients at the quadrature point.
    for i= 1:3
       ef(i) = ef(i) + psi(i)*xf*w(l)*J;   % J is the Jacobian
       for j=1:3
          ek(i,j)=ek(i,j)+ (T + xq*psi(i)*psi(j) )*J      % see below
       end
    end
 end
```

Note that *psi* has three values corresponding to three nonzero basis functions; *dpsi* is a 3×2 matrix which contains the partial derivatives $\partial\psi_i/\partial\xi$ and $\partial\psi_i/\partial\eta$. The evaluation of T is

$$\iint_{\Omega_e} p(x,y)\,\nabla\phi_i \cdot \nabla\phi_j \,dx\,dy = \iint_{\Omega_e} p(\xi,\eta)\,\left(\frac{\partial\psi_i}{\partial x}\frac{\partial\psi_j}{\partial x} + \frac{\partial\psi_i}{\partial y}\frac{\partial\psi_j}{\partial y}\right) dx\,dy\,.$$

We need to calculate $\partial\psi_i/\partial x$ and $\partial\psi_i/\partial y$ in terms of ξ and η. Notice that

$$\frac{\partial\psi_i}{\partial x} = \frac{\partial\psi_i}{\partial\xi}\frac{\partial\xi}{\partial x} + \frac{\partial\psi_i}{\partial\eta}\frac{\partial\eta}{\partial x}\,,$$

$$\frac{\partial\psi_i}{\partial y} = \frac{\partial\psi_i}{\partial\xi}\frac{\partial\xi}{\partial y} + \frac{\partial\psi_i}{\partial\eta}\frac{\partial\eta}{\partial y}\,.$$

However, we know that

$$\xi = \frac{1}{2A_e}\Big(\,(y_3-y_1)(x-x_1) - (x_3-x_1)(y-y_1)\Big)\,,$$
$$\eta = \frac{1}{2A_e}\Big(\,-(y_2-y_1)(x-x_1) + (x_2-x_1)(y-y_1)\Big)$$

so

$$\frac{\partial\xi}{\partial x} = \frac{1}{2A_e}(y_3-y_1)\,,\quad \frac{\partial\xi}{\partial y} = -\frac{1}{2A_e}(x_3-x_1)\,,$$

$$\frac{\partial\eta}{\partial x} = -\frac{1}{2A_e}(y_2-y_1)\,,\quad \frac{\partial\eta}{\partial y} = \frac{1}{2A_e}(x_2-x_1)\,.$$

- Add to the global stiffness matrix and the load vector.

```
for i= 1:3
   ig = nodes(i,nel);
   gf(ig) = gf(ig) + ef(i);
   for j=1:3
     jg  = nodes(j,nel);
     gk(ig,jg) = gk(ig,jg) + ek(i,j);
   end
end
```

- Solve the system of equations $\mathrm{gk}\,U = \mathrm{gf}$.
  - Direct method, *e.g.*, Gaussian elimination.
  - Sparse matrix technique, *e.g.*, $A = sparse(M, M)$.

  - Iterative method plus preconditioning, *e.g.*, Jacobi, Gauss-Seidel, SOR, conjugate gradient methods, *etc.*

- Error analysis.
  - Construct interpolation functions.
  - Error estimates for interpolation functions.
  - FE solution is the best approximation in the FE space in the energy norm.

9.5 Simplification of the FE method for Poisson equations

With constant coefficients, there is a closed form for the local stiffness matrix, in terms of the coordinates of the nodal points, so the FE algorithm can be simplified. We now introduce the simplified FE algorithm. A good reference is: *An introduction to the FE method with applications to non-linear problems* by R.E. White, John Wiley & Sons.

Let us consider the Poisson equation problem

$$\begin{aligned} -\Delta u &= f(x,y),\ (x,y) \in \Omega\,, \\ u(x,y) &= g(x,y)\,,\ (x,y) \in \partial\Omega_1\,, \\ \frac{\partial u}{\partial n} &= 0\,,\ (x,y) \in \partial\Omega_2\,, \end{aligned}$$

where Ω is an arbitrary but bounded domain. We can use Matlab PDE Tool-box to generate a triangulation for the domain Ω.

The weak form is

$$\iint_\Omega \nabla u \cdot \nabla v \, dx\, dy = \iint_\Omega f v \, dx\, dy\,.$$

With the piecewise linear basis functions defined on a triangulation on Ω, we can derive analytic expressions for the basis functions and the entries of the local stiffness matrix.

Theorem 9.11. *Consider a triangle determined by* (x_1, y_1), (x_2, y_2) *and* (x_3, y_3). *Let*

$$a_i = x_j y_m - x_m y_j\,, \tag{9.30}$$

$$b_i = y_j - y_m\,, \tag{9.31}$$

$$c_i = x_m - x_j\,, \tag{9.32}$$

where i, j, m *is a positive permutation of* 1, 2, 3, e.g., $i = 1$, $j = 2$ *and* $m = 3$; $i = 2$, $j = 3$ *and* $m = 1$; *and* $i = 3$, $j = 1$ *and* $m = 2$. *Then the corresponding three nonzero basis functions are*

$$\psi_i(x,y) = \frac{a_i + b_i\, x + c_i\, y}{2\Delta}, \quad i = 1,2,3\,, \tag{9.33}$$

where $\psi_i(x_i, y_i) = 1$, $\psi_i(x_j, y_j) = 0$ *if* $i \neq j$, *and*

$$\Delta = \frac{1}{2}\, det \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix} = \pm \textit{ area of the triangle.} \tag{9.34}$$

We prove the theorem for $\psi_1(x,y)$. Thus

$$\psi_1(x,y) = \frac{a_1 + b_1 x + c_1 y}{2\Delta}\,,$$
$$= \frac{(x_2y_3 - x_3y_2) + (y_2 - y_3)x + (x_3 - x_2)y}{2\Delta}\,,$$

so
$$\psi_1(x_2,y_2) = \frac{(x_2y_3 - x_3y_2) + (y_2 - y_3)x_2 + (x_3 - x_2)y_2}{2\Delta} = 0\,,$$
$$\psi_1(x_3,y_3) = \frac{(x_2y_3 - x_3y_2) + (y_2 - y_3)x_3 + (x_3 - x_2)y_3}{2\Delta} = 0\,,$$
$$\psi_1(x_1,y_1) = \frac{(x_2y_3 - x_3y_2) + (y_2 - y_3)x_1 + (x_3 - x_2)y_1}{2\Delta} = \frac{2\Delta}{2\Delta} = 1\,.$$

We can prove the same feature for ψ_2 and ψ_3.

We also have the following theorem, which is essential for the simplified FE method.

Theorem 9.12. *With the same notation as in Theorem 9.11, we have*

$$\iint_{\Omega_e} (\psi_1)^m (\psi_2)^n (\psi_3)^l \, dxdy = \frac{m!\, n!\, l!}{(m+n+l+2)\,!} 2\Delta\,, \tag{9.35}$$

$$\iint_{\Omega_e} \nabla\psi_i \cdot \nabla\psi_j \, dxdy = \frac{b_i b_j + c_i c_j}{4\Delta}\,,$$

$$F_1^e = \iint_{\Omega_e} \psi_1 f(x,y)\, dxdy \simeq f_1 \frac{\Delta}{6} + f_2 \frac{\Delta}{12} + f_3 \frac{\Delta}{12}\,,$$

$$F_2^e = \iint_{\Omega_e} \psi_2 f(x,y)\, dxdy \simeq f_1 \frac{\Delta}{12} + f_2 \frac{\Delta}{6} + f_3 \frac{\Delta}{12}\,,$$

$$F_3^e = \iint_{\Omega_e} \psi_3 f(x,y)\, dxdy \simeq f_1 \frac{\Delta}{12} + f_2 \frac{\Delta}{12} + f_3 \frac{\Delta}{6}\,,$$

where $f_i = f(x_i, y_i)$.

The proof is straightforward since we have the analytic form for ψ_i. We approximate $f(x,y)$ using

$$f(x,y) \simeq f_1\psi_1 + f_2\psi_2 + f_3\psi_3, \tag{9.36}$$

and therefore

$$\begin{aligned} F_1^e &\simeq \iint_{\Omega_e} \psi_1 f(x,y)\, dxdy \\ &= f_1 \iint_{\Omega_e} \psi_1^2 dxdy + f_2 \iint_{\Omega_e} \psi_1\psi_2\, dxdy + f_3 \iint_{\Omega_e} \psi_1\psi_3\, dxdy\,. \end{aligned} \tag{9.37}$$

Note that the integrals in the last expression can be obtained from the formula (9.35). There is a negligible error from approximating $f(x,y)$ compared with the error from the FE approximation when we seek approximate solution only in V_h space instead of $H^1(\Omega)$ space. Similarly we can get approximation F_2^e and F_3^e.

9.5.1 Main pseudo-code of the simplified FE method

Assume that we have a triangulation, *e,g.*, a triangulation generated from Matlab by saving the mesh. Then we have

$$\begin{aligned} &p(1,1),\ p(1,2),\ \cdots,\ p(1,nnode) && \text{as } x \text{ coordinates of the nodal points,}\\ &p(2,1),\ p(2,2),\ \cdots,\ p(2,nnode) && \text{as } y \text{ coordinates of the nodal points;} \end{aligned}$$

and the array t (the nodes in our earlier notation)

$$\begin{aligned} &t(1,1),\ t(1,2),\ \cdots,\ t(1,nele) && \text{as the index of the first node of an element,}\\ &t(2,1),\ t(2,2),\ \cdots,\ t(2,nele) && \text{as the index of the second node of the element,}\\ &t(3,1),\ t(3,2),\ \cdots,\ t(3,nele) && \text{as the index of the third node of the element;} \end{aligned}$$

and the array e to describe the nodal points on the boundary

$$\begin{aligned} &e(1,1),\ e(1,2),\ \cdots,\ e(1,nbc) && \text{as the index of the beginning node of a boundary edge,}\\ &e(2,1),\ e(2,2),\ \cdots,\ e(2,nbc) && \text{as the index of the end node of the boundary edge.} \end{aligned}$$

A Matlab code for the simplified FE method is listed below.

```
% Set-up: assume we have a triangulation p,e,t from Matlab PDE tool box
% already.

    [ijunk,nelem] = size(t);
    [ijunk,nnode] = size(p);

    for i=1:nelem
      nodes(1,i)=t(1,i);
      nodes(2,i)=t(2,i);
      nodes(3,i)=t(3,i);
    end

    gk=zeros(nnode,nnode);
    gf = zeros(nnode,1);

    for nel = 1:nelem,   % Begin to assemble by element.

      for j=1:3,           % The coordinates of the nodes in the
        jj = nodes(j,nel);    % element.
        xx(j) = p(1,jj);
        yy(j) = p(2,jj);
      end

    for nel = 1:nelem,   % Begin to assemble by element.
```

```
      for j=1:3,           % The coordinates of the nodes in the
        jj = nodes(j,nel);    % element.
        xx(j) = p(1,jj);
        yy(j) = p(2,jj);
      end

      for i=1:3,
        j = i+1 - fix((i+1)/3)*3;
        if j == 0
           j = 3;
        end
        m = i+2 - fix((i+2)/3)*3;
        if m  == 0
           m = 3;
        end

        a(i) = xx(j)*yy(m) - xx(m)*yy(j);
        b(i) = yy(j) - yy(m);
        c(i) = xx(m) - xx(j);
      end

      delta = ( c(3)*b(2) - c(2)*b(3) )/2.0;   % Area.

      for ir = 1:3,
        ii = nodes(ir,nel);
        for ic=1:3,
          ak = (b(ir)*b(ic) + c(ir)*c(ic))/(4*delta);
          jj = nodes(ic,nel);
          gk(ii,jj) = gk(ii,jj) + ak;
        end
          j = ir+1 - fix((ir+1)/3)*3;
             if j == 0
                j = 3;
             end
          m = ir+2 - fix((ir+2)/3)*3;
             if m == 0
                m = 3;
             end
          gf(ii) = gf(ii)+( f(xx(ir),yy(ir))*2.0 + f(xx(j),yy(j)) ...
                               + f(xx(m),yy(m)) )*delta/12.0;
      end

    end                  % End assembling by element.

%-------------------------------------------------------
```

```
% Now deal with the Dirichlet BC

    [ijunk,npres] = size(e);
    for i=1:npres,
       xb = p(1,e(1,i));  yb=p(2,e(1,i));
       g1(i) = uexact(xb,yb);
    end

    for i=1:npres,
      nod = e(1,i);
      for k=1:nnode,
         gf(k) = gf(k) - gk(k,nod)*g1(i);
         gk(nod,k) = 0;
         gk(k,nod) = 0;
      end
         gk(nod,nod) = 1;
         gf(nod) = g1(i);
    end

    u=gk\gf;            % Solve the linear system.
    pdemesh(p,e,t,u)    % Plot the solution.

% End.
```

Example 9.4. *We test the simplified FE method for the Poisson equation using the following example:*

- *Domain: Unit square with a hole,* cf. *Fig. 9.9.*
- *Exact solution:* $u(x, y) = x^2 + y^2$, *for* $f(x, y) = -4$.
- *BC: Dirichlet condition on the whole boundary.*
- *Use Matlab PDE Tool-box to generate initial mesh and then* export it.

Fig. 9.9 shows the domain and the mesh is generated by the Matlab PDE Tool-box. The left plot in Fig. 9.10 is the mesh plot for the FE solution, and the right plot is the error plot (the error is $O(h^2)$).

9.6 Some FE spaces in $H^1(\Omega)$ and $H^2(\Omega)$

Given a triangulation (triangles, rectangles, quadrilaterals, *etc.*), let us construct different FE spaces with finite dimensions. There are several reasons to do so, including:

- better accuracy of the FE solution, with piecewise higher order polynomial basis functions; and
- to allow for higher order derivatives in higher order PDE, *e.g.*, in solving the biharmonic equation in H^2 space.

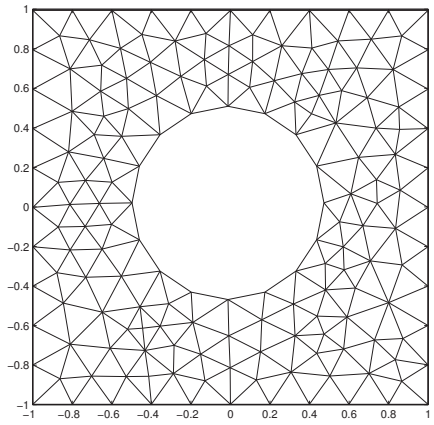


Figure 9.9. *A mesh generated from Matlab.*

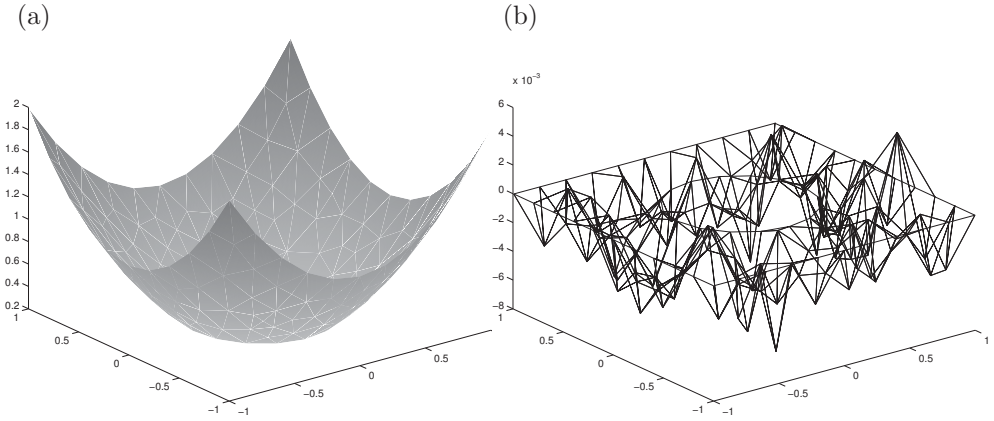


Figure 9.10. *(a) A plot of the FE solution when $f(x,y) = -4$; (b) The corresponding error plot.*

As previously mentioned, we consider conforming piecewise polynomial FE spaces. A set of polynomials of degree k is denoted by

$$P_k = \left\{ \; v(x,y) \; , \quad v(x,y) = \sum_{i,j=0}^{i+j \le k} a_{ij} \, x^i x^j \right\},$$

in the xy-plane. Below we list some examples,

$$\begin{aligned}
P_1 &= \{ \; v(x,y), \quad v(x,y) = a_{00} + a_{10}x + a_{01}y \; \}, \\
P_2 &= \{ \; v(x,y), \quad v(x,y) = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2 \; \}, \\
P_3 &= P_2 + \{ a_{30}x^3 + a_{21}x^2y + a_{12}xy^2 + a_{03}y^3 \}, \\
&\cdots\cdots\cdots\cdots .
\end{aligned}$$

Degree of freedom of P_k. For any fixed x^i, the possible y^j terms of a $p_k(x,y) \in P_k$ are $y^0, y^1, \cdots, y^{k-i}$, *i.e.*, j ranges from 0 to $k-i$. Thus there are $k-i+1$ parameters for a given x^i, and the total degree of freedom is

$$\begin{aligned}\sum_{i=0}^{k}(k-i+1) &= \sum_{i=0}^{k}(k+1) - \sum_{i=0}^{k} i \\ &= (k+1)^2 - \frac{k(k+1)}{2} = \frac{(k+1)(k+2)}{2}\,.\end{aligned}$$

Some degrees of freedom for different k are:

- 3 when $k=1$, the linear function space P_1;
- 6 when $k=2$, the quadratic function space P_2;
- 10 when $k=3$, the cubic function space P_3;
- 15 when $k=4$, the fourth order polynomials space P_4; and
- 21 when $k=5$, the fifth order polynomials space P_5.

Regularity requirements: Generally, we cannot conclude that $v(x,y) \in C^0$ if $v(x,y) \in H^1$. However, if V_h is a finite dimensional space of piecewise polynomials, then that is indeed true. Similarly, if $v(x,y) \in H^2$ and $v(x,y)|_{K_i} \in P_k,\ \forall K_i \in T_h,$ then $v(x,y) \in C^1$. The regularity requirements are important for the construction of the FE space.

As is quite well known, there are two ways to improve the accuracy. One way is to decrease the mesh size h, and the other is to use high order polynomial spaces P_k. If we use the P_k space on a given triangulation T_h for a linear second order elliptic PDE, the error estimates for the FE solution u_h are

$$\|u-u_h\|_{H^1(\Omega)} \le C_1 h^k \|u\|_{H^{k+1}(\Omega)},\ \|u-u_h\|_{L^2(\Omega)} \le C_2 h^{k+1}\|u\|_{H^{k+1}(\Omega)}\,. \quad (9.38)$$

9.6.1 The piecewise quadratic functions space

The degree of the freedom of a quadratic function on a triangle is 6, so we may add 3 auxiliary middle points along the three sides of the triangle.

Theorem 9.13. *Consider a triangle $K=(a^1,a^2,a^3)$, as shown in Fig. 9.11. A function $v(x,y) \in P_2(K)$ is uniquely determined by its values at*

$$v(a^i)\,,\ i=1,2,3,\ \textit{and the three middle points}\ v(a^{12}),\ v(a^{23}),\ v(a^{31})\,.$$

As there are six parameters and six conditions, we expect to be able to determine the quadratic function uniquely. Highlights of the proof are as follows.

- We just need to prove the homogeneous case $v(a^i)=0$, $v(a^{ij})=0$, since the right-hand side does not affect the existence and uniqueness.

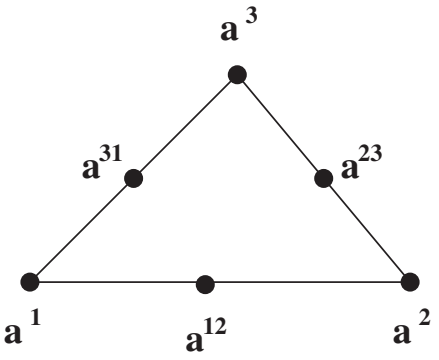


Figure 9.11. *A diagram of* 6 *points in a triangle to determine a quadratic function.*

- We can represent a quadratic function as a product of two linear functions, *i.e.*, $v(\mathbf{x}) = \psi_1(\mathbf{x})\omega(\mathbf{x}) = \psi_1(\mathbf{x})\psi_2(\mathbf{x})\omega_0$, with $\psi_i(\mathbf{x})$ denoting the local basis function such that $\psi_i(a^i) = 1$ and $\psi_i(a^j) = 0$ if $i \neq j$. Note that here we use $\mathbf{x} = (x, y)$ notation for convenience.
- It is easier to introduce a coordinate axis aligned with one of the three sides.

  Proof: We introduce the new coordinates (*cf.* Fig. 9.5)

$$\xi = (x - x_2)\cos\alpha + (y - y_2)\sin\alpha\,,$$
$$\eta = -(x - x_2)\sin\alpha + (y - y_2)\cos\alpha\,,$$

such that a^2 is the origin and a^2a^3 is the η- axis. Then $v(x, y)$ can be written as

$$v(x, y) = v(x(\xi, \eta), y(\xi, \eta)) = \bar{v}(\xi, \eta) = \bar{a}_{00} + \bar{a}_{10}\xi + \bar{a}_{01}\eta + \bar{a}_{20}\xi^2 + \bar{a}_{11}\xi\eta + \bar{a}_{02}\eta^2\,.$$

Furthermore, under the new coordinates, we have

$$\psi_1(\xi, \eta) = \sigma + \beta\xi + \gamma\eta = \beta\xi\,, \quad \beta \neq 0,$$

since $\psi_1(a^2) = \psi_1(a^3) = 0$. Along the η-axis ($\xi = 0$), $\bar{v}(\xi, \eta)$ has the following form

$$\bar{v}(0, \eta) = \bar{a}_{00} + \bar{a}_{01}\eta + \bar{a}_{02}\eta^2.$$

Since $\bar{v}(a^2) = \bar{v}(a^3) = \bar{v}(a^{23}) = 0$, we get $\bar{a}_{00} = 0$, $\bar{a}_{01} = 0$ and $\bar{a}_{02} = 0$, therefore,

$$\begin{aligned}\bar{v}(\xi, \eta) &= \bar{a}_{10}\xi + \bar{a}_{11}\xi\eta + \bar{a}_{20}\xi^2 = \xi\,(\bar{a}_{10} + \bar{a}_{11}\eta + \bar{a}_{20}\xi) \\ &= \beta\xi\left(\frac{\bar{a}_{10}}{\beta} + \frac{\bar{a}_{20}}{\beta}\xi + \frac{\bar{a}_{11}}{\beta}\eta\right) \\ &= \psi_1(\xi, \eta)\omega(\xi, \eta)\,.\end{aligned}$$

Similarly, along the edge a^1a^3

$$v(a^{13}) = \psi_1(a^{13})\,\omega(a^{13}) = \frac{1}{2}\,\omega(a^{13}) = 0$$
$$v(a^1) = \psi_1(a^1)\omega(a^1) = \omega(a^1) = 0,$$

i.e.,

$$\omega(a^{13}) = 0, \quad \omega(a^1) = 0\,.$$

By similar arguments, we conclude that

$$\omega(x,y) = \psi_2(x,y)\,\omega_0\,,$$

and hence

$$v(x,y) = \psi_1(x,y)\psi_2(x,y)\omega_0\,.$$

Using the zero value of v at a^{12}, we have

$$v(a^{12}) = \psi_1(a^{12})\,\psi_2(a^{12})\,\omega_0 = \frac{1}{2}\,\frac{1}{2}\,\omega_0 = 0\,,$$

so we must have $\omega_0 = 0$ and hence $v(x,y) \equiv 0$.

Continuity along the edges

Along each edge, a quadratic function $v(x,y)$ can be written as a quadratic function of one variable. For example, if the edge is represented as

$$y = ax + b \quad \text{or} \quad x = ay + b\,,$$

then

$$v(x,y) = v(x, ax+b) \quad \text{or} \quad v(x,y) = v(ay+b, y).$$

Thus the piecewise quadratic functions defined on two triangles with a common side are identical on the entire side if they have the same values at the two end points and at the mid-point of the side.

Representing quadratic basis functions using linear functions

To define quadratic basis functions with minimum compact support, we can determine the six nonzero functions using the values at three vertices and the mid-points $\mathbf{v} = (v(a^1), v(a^2), v(a^3), v(a^{12}), v(a^{23}), v(a^{31})) \in \mathbf{R}^6$. We can either take $\mathbf{v} = \mathbf{e}_i \in \mathbf{R}^6$, $i = 1, 2, \cdots, 6$ respectively, or determine a quadratic function on the triangle using the linear basis functions as stated in the following theorem.

Theorem 9.14. *A quadratic function on a triangle can be represented by*

$$\begin{aligned} v(x,y) \quad = \quad & \sum_{i=1}^{3} v(a^i)\phi_i(x,y)\left(2\phi_i(x,y) - 1\right) \\ & + \sum_{i,j=1, i<j}^{3} 4\,v(a^{ij})\,\phi_i(x,y)\,\phi_j(x,y)\,. \end{aligned} \tag{9.39}$$

Proof: It is easy to verify the vertices if we substitute a^j into the right-hand side of the expression above,

$$v(a^j)\phi_j(a^j)\left(2\phi_j(a^j) - 1\right) = v(a^j)\,,$$

since $\phi_i(a^j) = 0$ if $i \neq j$. We take one mid-point to verify the theorem. On substituting a^{12} into the left expression, we have

$$\begin{aligned} & v(a^1)\phi_1(a^{12})\left(2\phi_1(a^{12})-1\right) + v(a^2)\phi_2(a^{12})\left(2\phi_2(a^{12})-1\right) \\ & \quad + v(a^3)\phi_3(a^{12})\left(2\phi_3(a^{12})-1\right) + 4v(a^{12})\phi_1(a^{12})\phi_2(a^{12}) \\ & \quad + 4v(a^{13})\phi_1(a^{12})\phi_3(a^{12}) + 4v(a^{23})\phi_2(a^{12})\phi_3(a^{12}) \\ & \quad\quad = v(a^{12})\,, \end{aligned}$$

since $2\phi_1(a^{12}) - 1 = 2 \times \frac{1}{2} - 1 = 0$, $2\phi_2(a^{12}) - 1 = 2 \times \frac{1}{2} - 1 = 0$, $\phi_3(a^{12}) = 0$ and $4\phi_1(a^{12})\phi_2(a^{12}) = 4 \times \frac{1}{2} \times \frac{1}{2} = 1$. Note that the local stiffness matrix is 6×6, when using quadratic functions.

9.6.2 Cubic basis functions in $H^1 \cap C^0$

There are several ways to construct cubic basis functions in $H^1 \cap C^0$ over a triangulation, but a key consideration is to keep the continuity of the basis functions along the edges of neighboring triangles. We recall that the degree of freedom of a cubic function in 2D is 10, and one way is to add two auxiliary points along each side and one auxiliary point inside the triangle. thus together with the three vertices, we have ten points on a triangle to match the degree of the freedom (*cf.* Fig. 9.12). Existence and uniqueness conditions for such a cubic function are stated in the following theorem.

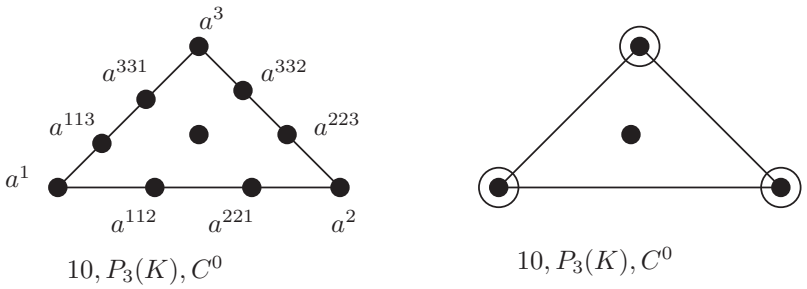


Figure 9.12. *A diagram of the freedom used to determine two different cubic basis function $H^1 \cap C^0$. We use the following notation: $\bullet$ for function values; $\circ$ for values of the first derivatives.*

Theorem 9.15. *A cubic function $v \in P_3(K)$ is uniquely determined by the values of*

$$v(a^i), \quad v(a^{iij}),\ i,j = 1,2,3,\ i \neq j \quad \textit{and} \quad v(a^{123})\,, \tag{9.40}$$

where

$$a^{123} = \frac{1}{3}\left(a^1 + a^2 + a^3\right), \quad a^{iij} = \frac{1}{3}\left(2a^i + a^j\right),\ i,j = 1,2,3,\ i \neq j\,. \tag{9.41}$$

Sketch of the proof: Similar to the quadratic case, we just need to prove that the cubic function is identically zero if $v(a^i) = v(a^{iij}) = v(a^{123}) = 0$. Again using local coordinates where one of the sides of the triangle T is on an axis,

$$v(\mathbf{x}) = C\phi_1(\mathbf{x})\phi_2(\mathbf{x})\phi_3(\mathbf{x})\,,$$

where C is a constant. Since $v(a^{123}) = C\phi_1(a^{123})\phi_2(a^{123})\phi_3(a^{123}) = C\frac{1}{3}\times\frac{1}{3}\times\frac{1}{3} = 0$, we conclude that $C = 0$ and hence $v(\mathbf{x}) \equiv 0$.

With reference to continuity along the common side of two adjacent triangles, we note that the polynomial of two variables again becomes a polynomial of one variable there, since we can substitute either x for y or y for x from the line equations $l_0 + l_{10}x + l_{01}y = 0$. Furthermore, a cubic function of one variable is uniquely determined by the values of four distinct points.

There is another choice of cubic basis functions, using the first order derivatives at the vertices, *cf.* the right diagram in Fig. 9.12. This alternative is stated in the following theorem.

Theorem 9.16. *A cubic function $v \in P_3(K)$ is uniquely determined by the values of*

$$v(a^i), \quad \frac{\partial v}{\partial x_j}(a^i),\ i = 1,2,3,\ j = 1,2 \text{ and } i \neq j, \quad v(a^{123})\,, \tag{9.42}$$

where $\partial v/\partial x_j(a^i)$ represents $\partial v/\partial x(a^i)$ when $j = 1$ and $\partial v/\partial y(a^i)$ when $j = 2$, at the nodal point a^i.

At each vertex of the triangle, there are three degrees of freedom, namely, the function value and two first order partial derivatives, so total there are nine degrees of freedom. An additional degree of freedom is the value at the centroid of the triangle. For the proof of the continuity, we note that on a common side of two adjacent triangles a cubic polynomial of one variable is uniquely determined by its function values at two distinct points plus the first order derivatives in Hermite interpolation theory. The first order derivative is the tangential derivative along the common side defined as $\partial v/\partial t = \partial v/\partial x\, t_1 + \partial v/\partial y\, t_2$, where $\mathbf{t} = (t_1, t_2)$ such that $t_1^2 + t_2^2 = 1$ is the unit direction of the common side.

9.6.3 Basis functions in $H^2 \cap C^1$

Basis functions are also needed for fourth order PDE such as the 2D biharmonic equation

$$\Delta\left(u_{xx} + u_{yy}\right) = u_{xxxx} + 2u_{xxyy} + u_{yyyy} = 0\,. \tag{9.43}$$

Since second order partial derivatives are involved in the weak form, we need to use polynomials with degree more than three. On a triangle, if the function values and partial derivatives up to second order are specified at the three vertices, the degree of freedom would be at least 18. The closest polynomial would be of degree five, as a polynomial $v(\mathbf{x}) \in P_5$ has degree of freedom 21, *cf.* the left diagram in Fig. 9.13.

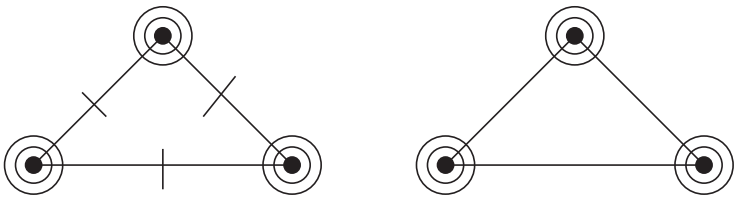


Figure 9.13. *A diagram of the freedom used to determine two different fifth order polynomial basis functions in $H^2 \cap C^1$. In addition to the previous notation, we introduce $\bigcirc$ for values of the second derivatives; $/$ for values of the first derivatives; and $\nearrow$ for values of the mixed derivatives .*

Theorem 9.17. *A quintic function $v(x,y) \in P_5(K)$ is uniquely determined by the values of*

$$D^\alpha v(a^i)\,,\ i=1,2,3,\ |\alpha| \le 2, \quad \frac{\partial v}{\partial n}(a^{ij})\,,\ i,j=1,2,3,\ i<j\,, \tag{9.44}$$

where $\partial v/\partial n(a^i) = n_1 \partial v/\partial x(a^i) + n_2 \partial v/\partial y(a^i)$ represents the normal derivative of $v(x)$ at a^i and $n=(n_1,n_2)$ such that $n_1^2+n_2^2=1$ is the outward unit normal at the boundary of the triangle.

Sketch of the proof: We just need to show that $v(\mathbf{x}) = 0$ if $D^\alpha v(a^i) = 0$, $i = 1,2,3$, $|\alpha| \le 2$ and $\partial v/\partial n(a^{ij}) = 0$, $i,j = 1,2,3$, $i<j$. A fifth order polynomial $v(s)$ of one variable s is uniquely determined by the values of v and its derivatives $v'(s)$ and $v''(s)$ at two distinct points, so along a^2a^3, $v(\mathbf{x})$ must be zero for the given homogeneous conditions. We note that $\frac{\partial v}{\partial n}(\mathbf{x})$ is a fourth order polynomial of one variable along a^2a^3. Since all of the first and second order partial derivatives are zero at a^2 and a^3,

$$\frac{\partial v}{\partial n}(a^i) = 0\,, \qquad \frac{\partial}{\partial t}\left(\frac{\partial v}{\partial n}\right)(a^i) = 0,\ i=2,3\,,$$

and $\frac{\partial v}{\partial n}(a^{23}) = 0$. Here again, $\frac{\partial}{\partial t}$ is the tangential directional derivative. From the five conditions, we have $\frac{\partial v}{\partial n}(\mathbf{x}) = 0$ along a^2a^3, so we can factor $\phi_1^2(\mathbf{x})$ out of $v(\mathbf{x})$ to get

$$v(\mathbf{x}) = \phi_1^2(\mathbf{x})\, p_3(\mathbf{x})\,, \tag{9.45}$$

where $p_3(\mathbf{x}) \in P_3$. Similarly, we can factor out $\phi_2^2(\mathbf{x})$ and $\phi_3^2(\mathbf{x})$ to get

$$v(\mathbf{x}) = \phi_1^2(\mathbf{x})\,\phi_2^2(\mathbf{x})\,\phi_3^2(\mathbf{x})\,C\,, \tag{9.46}$$

where C is a constant. Consequently $C=0$, otherwise $v(\mathbf{x})$ would be a polynomial of degree six, which contradicts that $v(\mathbf{x}) \in P_5$.

The continuity condition along a common side of two adjacent triangles in C^1 has two parts, namely, both the function and the normal derivative must be continuous. Along a common side of two adjacent triangles, a fifth order polynomial of $v(x,y)$ is actually a fifth order polynomial of one variable $v(s)$, which can be uniquely determined by the values $v(s)$, $v'(s)$ and $v''(s)$ at two distinct points. Thus the two fifth order polynomials

on two adjacent triangles are identical along the common side if they have the same values of $v(s)$, $v'(s)$ and $v''(s)$ at the two shared vertices. Similarly, for the normal derivative along a common side of two adjacent triangles, we have a fourth order polynomial of one variable $\partial v/\partial n(s)$. The polynomials can be uniquely determined by the values $\partial v/\partial n(s)$ and $(d/ds)\,(\partial v/\partial n)\,(s)$ at two distinct points plus the value of a $\partial v/\partial n(s)$ at the mid-point. Thus the continuity of the normal derivative is also guaranteed.

An alternative approach is to replace the values of $\frac{\partial v}{\partial n}(a^{ij})$ at the three mid-points of the three sides by imposing another three conditions. For example, assuming that along a^2a^3 the normal derivative of the fifth order polynomial has the form

$$\frac{\partial v}{\partial n} = \widetilde{a_{00}} + \widetilde{a_{10}}\eta + \widetilde{a_{20}}\eta^2 + \widetilde{a_{30}}\eta^3 + \widetilde{a_{40}}\eta^4\,,$$

we can impose $\widetilde{a_{40}} = 0$. In other words, along the side of a^2a^3 the normal derivative of $\partial v/\partial n$ becomes a cubic polynomial of one variable. The continuity can again be guaranteed by the Hermite interpolation theory. Using this approach, the degree of the freedom is reduced to 18 from the original 21, *cf.* the right diagram in Fig. 9.13 for an illustration.

9.6.4 Finite element spaces on quadrilaterals

While triangular meshes are intensively used, particularly for arbitrary domains, meshes using quadrilaterals are also popular for rectangular regions. Bilinear functions are often used as basis functions, and let us first consider a bilinear function space in $H^1 \cap C^0$. A bilinear function space over a quadrilateral K in 2D, as illustrated Fig. 9.14, is defined as

$$Q_1(K) = \Big\{\, v(x,y), \quad v(x,y) = a_{00} + a_{10}x + a_{01}y + a_{11}xy \Big\}\,, \tag{9.47}$$

where $v(x,y)$ is linear in both x and y. The degree of the freedom of a bilinear function in $Q_1(K)$ is 4.

Theorem 9.18. *A bilinear function $v(x,y) \in Q_1(K)$ is uniquely determined by its values at four corners.*

Proof: without loss of the generality, assume that the quadrilateral is determined by the four corners a^i: (0, 0), (x_1, 0), (x_1, y_1) and (0, y_1). The coefficient matrix of the linear system of algebraic equations that determines the coefficients a_{ij}, $i,j = 0,1$ is

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & x_1 & 0 & 0 \\ 1 & 0 & y_1 & 0 \\ 1 & x_1 & y_1 & x_1y_1 \end{pmatrix},$$

with determinant $det(A) = x_1^2y_1^2 \neq 0$ since $x_1y_1 \neq 0$. Indeed, we have analytic expressions

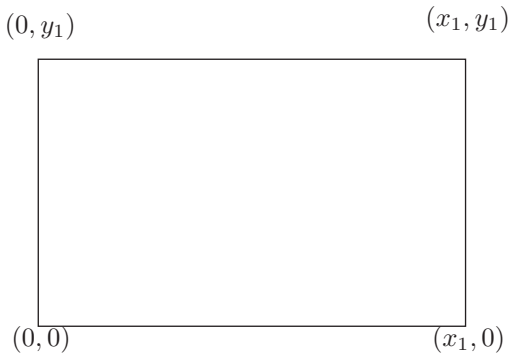


Figure 9.14. *A standard rectangle on which four bilinear basis functions can be defined.*

for the four nonzero basis functions over the quadrilateral, namely,

$$\phi_1(x,y) = 1 - \frac{x}{x_1} - \frac{y}{y_1} + \frac{xy}{x_1 y_1} \,, \tag{9.48}$$

$$\phi_2(x,y) = \frac{x}{x_1} - \frac{xy}{x_1 y_1} \,, \tag{9.49}$$

$$\phi_3(x,y) = \frac{xy}{x_1 y_1} \,, \tag{9.50}$$

$$\phi_4(x,y) = \frac{y}{y_1} - \frac{xy}{x_1 y_1} \,. \tag{9.51}$$

On each side of the rectangle, $v(x,y)$ is a linear function of one variable (either x or y), and uniquely determined by the values at the two corners. Thus any two basis functions along one common side of two adjacent rectangles are identical if they have the save values at the two corners, although it is hard to match the continuity condition if arbitrary quadrilaterals are used instead of rectangles or cubic boxes.

A bi-quadratic function space over a rectangle is defined by

$$\begin{aligned} Q_2(K) \quad = \quad & \Big\{ v(x,y), \quad v(x,y) = a_{00} + a_{10}x + a_{01}y + \\ & a_{20}x^2 + a_{11}xy + a_{20}y^2 + a_{21}x^2 y + a_{12}xy^2 + a_{22}x^2 y^2 \Big\} \,. \end{aligned} \tag{9.52}$$

The degree of the freedom is 9. To construct basis functions in $H^1 \cap C^0$, as for the quadratic functions over triangles we can add four auxiliary points at the mid-points of the four sides plus a point in the rectangle. In general, a bilinear function space of order k over a rectangle is defined by

$$Q_k(K) \quad = \quad \left\{ v(x,y), \quad v(x,y) = \sum_{i,j=0, i\le k, j\le k} a_{ij} x^i y^j \right\} . \tag{9.53}$$

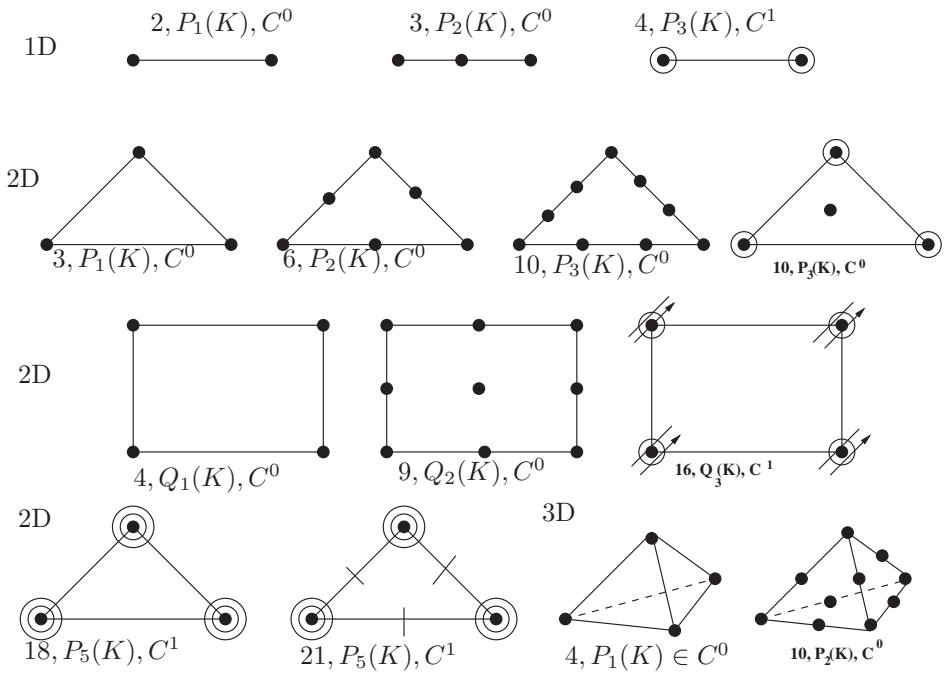


Figure 9.15. *Diagrams of different FE spaces in 1D, 2D and 3D. The degree of the freedom, polynomial type and the space is listed in each case. We use the following notation:* • *for function values;* ∘ *for values of the first derivatives;* ◯ *for values of the second derivatives;* / *for values of the first derivatives; and* ↗ *for values of the mixed derivatives.*

9.6.5 Other FE spaces

The diagrams in Fig. 9.15 illustrate different FE spaces over a line segment in 1D, a triangle or quadrilateral in 2D and a tetrahedron in 3D. From these diagrams, we can find the degree of the freedom, the polynomial basis functions and the Sobolev spaces.

9.7 The FE method for parabolic problems

We can use the FE method for time dependent problems, and there are two approaches. One approach is to discretize the space variables using the FE method while discretizing the time variable using some FD method. This is possible if the PDE is separable. Another way is to discretize both the space and time variables using the FE method. In this section, we briefly explain the first approach, since it is simply and easy to implement.

Let us consider the PDE

$$\frac{\partial u}{\partial t} = \nabla \cdot (p\nabla u) + qu + f(x,y,t)\,,\ (x,y) \in \Omega,\ 0 \le t \le T\,, \tag{9.54}$$

$$u(x,y,0) = 0\,,\ (x,y) \in \Omega\,, \quad \text{the initial condition,} \tag{9.55}$$

$$u(x,y,t)\Big|_{\partial\Omega} = g(x,y,t)\,, \quad \text{the boundary condition,} \tag{9.56}$$

where p, q, f and g are given functions with usual regularity assumptions. Multiplying the PDE by a test function $v(x,y) \in H^1(\Omega)$ on both sides, and then integrating over the domain, once again leads to the weak form:

$$\iint_\Omega u_t v\, dxdy = \iint_\Omega (qv - p\nabla u \cdot \nabla v)\, dxdy + \iint_\Omega fv\, dxdy\,, \tag{9.57}$$

where $u_t = \partial u/\partial t$. The weak form above can be simplified as

$$(u_t, v) = -a(u,v) + (f,v)\ \forall v \in H^1(\Omega)\,, \tag{9.58}$$

where $a(u,v) = \iint_\Omega (p\nabla u \cdot \nabla v - qv)\ dxdy$.

Given a triangulation T_h and FE space $V_h \in H^1(\Omega) \cap C^0(\Omega)$, with $\phi_i(x,y)$, $i = 1, 2, \cdots, M$ denoting a set of basis functions for V_h, we seek the FE solution of form

$$u_h(x,y,t) = \sum_{j=1}^{M} \alpha_j(t)\, \phi_j(x,y)\,. \tag{9.59}$$

Substituting this expression into (9.58), we obtain

$$\left(\sum_{j=1}^{M} \alpha_i'(t)\phi_i(x,y)\ ,\ v_h\right) = -a\left(\sum_{j=1}^{M} \alpha_i(t)\phi_i(x,y),\ v_h\right) + (f, v_h) \tag{9.60}$$

and then take $v_h(x,y) = \phi_i(x,y)$ for $i = 1, 2, \cdots, M$ to get the linear system of ordinary differential equations in the $\alpha_j(t)$:

$$\begin{bmatrix} (\phi_1,\phi_1) & (\phi_1,\phi_2) & \cdots & (\phi_1,\phi_M) \\ (\phi_2,\phi_1) & (\phi_2,\phi_2) & \cdots & (\phi_2,\phi_M) \\ \vdots & \vdots & \vdots & \vdots \\ (\phi_M,\phi_1) & (\phi_M,\phi_2) & \cdots & (\phi_M,\phi_M) \end{bmatrix} \begin{bmatrix} \alpha_1'(t) \\ \alpha_2'(t) \\ \vdots \\ \alpha_M'(t) \end{bmatrix} =$$

$$\begin{bmatrix} (f,\phi_1) \\ (f,\phi_2) \\ \vdots \\ (f,\phi_M) \end{bmatrix} - \begin{bmatrix} a(\phi_1,\phi_1) & a(\phi_1,\phi_2) & \cdots & a(\phi_1,\phi_M) \\ a(\phi_2,\phi_1) & a(\phi_2,\phi_2) & \cdots & a(\phi_2,\phi_M) \\ \vdots & \vdots & \vdots & \vdots \\ a(\phi_M,\phi_1) & a(\phi_M,\phi_2) & \cdots & a(\phi_M,\phi_M) \end{bmatrix} \begin{bmatrix} \alpha_1(t) \\ \alpha_2(t) \\ \vdots \\ \alpha_M(t) \end{bmatrix},$$

and the corresponding problem can therefore be expressed as

$$B\frac{d\vec{\alpha}}{dt} + A\vec{\alpha} = F, \qquad \alpha_i(0) = u(N_i,0),\ i = 1, 2, \cdots, M. \tag{9.61}$$

There are many methods to solve the above problem involving the system of first order ODE. We can use the ODE Suite in Matlab, but note that the ODE system is known to be very stiff. We can also use FD methods that march in time, since we know the initial condition on $\vec{\alpha}(0)$. Thus with the solution $\vec{\alpha}^k$ at time t^k, we compute the solution $\vec{\alpha}^{k+1}$ at the time $t^{k+1} = t^k + \Delta t$ for $k = 0, 1, 2, \cdots$.

Explicit Euler method

If the forward FD approximation is invoked, we have

$$B\frac{\vec{\alpha}^{k+1} - \vec{\alpha}^k}{\Delta t} + A\vec{\alpha}^k = F^k, \tag{9.62}$$

$$\text{or} \quad \vec{\alpha}^{k+1} = \vec{\alpha}^k + \Delta t B^{-1}\left(F^k - A\vec{\alpha}^k\right). \tag{9.63}$$

Since B is a non-singular tridiagonal matrix, its inverse and hence $B^{-1}\left(F^k - A\vec{\alpha}^k\right)$ can be computed. However, the CFL (Courant-Friedrichs-Lewy) condition

$$\Delta t \leq Ch^2. \tag{9.64}$$

must be satisfied to ensure numerical stability. Thus we need to use a rather small time step.

Implicit Euler method

If we invoke the backward FD approximation, we get

$$B\frac{\vec{\alpha}^{k+1} - \vec{\alpha}^k}{\Delta t} + A\vec{\alpha}^{k+1} = F^{k+1}, \tag{9.65}$$

$$\text{or} \quad (B + \Delta t A)\,\vec{\alpha}^{k+1} = \vec{\alpha}^k + \Delta t F^{k+1} \tag{9.66}$$

then there is no constraint on the time step and thus the method is called unconditionally stable. However, we need to solve a linear system of equations similar to that for an elliptic PDE at each time step.

The Crank-Nicolson method

Both of the above Euler methods are first order accurate in time and second order in space, *i.e.*, the error in computing $\vec{\alpha}$ is $O(\Delta t + h^2)$. We obtain a second order scheme in time as well on space if we use the central FD approximation:

$$B\frac{\vec{\alpha}^{k+1} - \vec{\alpha}^k}{\Delta t} + \frac{1}{2}A\left(\vec{\alpha}^{k+1} + \vec{\alpha}^k\right) = \frac{1}{2}\left(F^{k+1} + F^k\right), \tag{9.67}$$

$$\text{or} \quad \left(B + \frac{1}{2}\Delta t A\right)\vec{\alpha}^{k+1} = \left(B - \frac{1}{2}\Delta t A\right)\vec{\alpha}^k + \frac{1}{2}\Delta t\left(F^{k+1} + F^k\right). \tag{9.68}$$

This Crank-Nicolson method is second order accurate in both time and space, and it is unconditionally stable for linear PDE. The challenge is to solve the resulting linear system of equations efficiently.

9.8 Exercises

1. Derive the weak form for the following problem:

$$
\begin{aligned}
&-\nabla \cdot (p(x,y)\nabla u(x,y)) + q(x,y)u(x,y) = f(x,y)\ ,\ (x,y) \in \Omega\ , \\
&u(x,y) = 0\ ,\ (x,y) \in \partial\Omega_1\ , \qquad \frac{\partial u}{\partial \mathbf{n}} = g(x,y),\ (x,y) \in \partial\Omega_2\ , \\
&a(x,y)u(x,y) + \frac{\partial u}{\partial \mathbf{n}} = c(x,y)\ ,\ (x,y) \in \partial\Omega_3\ ,
\end{aligned}
$$

   where $q(x,y) \geq q_{min} > 0$, $\partial\Omega_1 \cup \partial\Omega_1 \cup \partial\Omega_3 = \partial\Omega$ and $\partial\Omega_i \cap \partial\Omega_j = \phi$. Provide necessary conditions so that the weak form has a unique solution. Show your proof using the Lax-Milgram Lemma but without using the Poincaré inequality.

2. Derive the weak form and appropriate space for the following problem involving the bi-harmonic equation:

$$
\begin{aligned}
\Delta\Delta u(x,y) &= f(x,y)\ ,\ (x,y) \in \Omega\ , \\
u(x,y)|_{\partial\Omega} = 0\ , &\quad u_n(x,y)|_{\partial\Omega} = 0\ .
\end{aligned}
$$

   What kind of basis function do you suggest, to solve this problem numerically?
   Hint: Use Green's theorem twice.

3. Consider the problem involving the Poisson equation:

$$
\begin{aligned}
-\Delta u(x,y) &= 1\ ,\ (x,y) \in \Omega\ , \\
u(x,y)|_{\partial\Omega} &= 0\ ,
\end{aligned}
$$

   where Ω is the unit square. Using a uniform triangulation, derive the stiffness matrix and the load vector for $N = 2$; in particular, take $h = 1/3$ and consider

   (a) the nodal points ordered as $(1/3, 1/3)$, $(2/3, 1/3)$; $(1/3, 2/3)$, and $(2/3, 2/3)$; and

   (b) the nodal points ordered as $((1/3, 2/3)$, $(2/3, 1/3)$; $(1/3, 1/3)$, and $(2/3, 2/3)$.

   Write down each basis function explicitly.

4. Use the Matlab PDE toolbox to solve the following problem involving a parabolic equation for $u(x, y, t)$, and make relevant plots:

$$
\begin{aligned}
u_t = u_{xx} + u_{yy}, &\quad (x,y) \in [-1\ \ 1] \times [-1\ \ 1]\ , \\
u(x,y,0) = 0\ .&
\end{aligned}
$$

   The geometry and the BC are defined in Fig. 9.16. Show some plots of the solution (mesh, contour, *etc.*).

5. Download the Matlab source code *f.m, my_assemb.m, uexact.m* from

   `http://www4.ncsu.edu/~zhilin/FD\_FEM\_Book.`

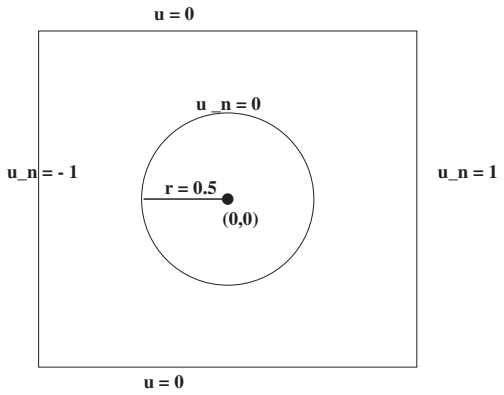


Figure 9.16. *Diagram for Exercise 2.*

Use the exported mesh of the geometry generated from Matlab, see Fig. 9.16 to solve the Poisson equation

$$-(u_{xx} + u_{yy}) = f(x, y)\,,$$

subject to the Dirichlet BC corresponding to the exact solution

$$u(x, y) = \frac{1}{4}\left(x^2 + y^4\right) \sin \pi x \cos 4\pi y\,.$$

Plot the solution and the error.

6. Modify the Matlab code to consider the generalized Helmholtz equation

$$-(u_{xx} + u_{yy}) + \lambda\, u = f(x, y)\,.$$

Test your code with $\lambda = 1$, with reference to the same solution, geometry and BC as in Problem 5. Adjust $f(x, y)$ to check the errors.

7. Modify the Matlab code to consider the Poisson equation

$$-\nabla\left(p(x, y) \cdot \nabla u(x, y)\right) = f(x, y)\,, \tag{9.69}$$

using a third order quadrature formula. Choose two examples with nonlinear $p(x, y)$ and $u(x, y)$ to show that your code is bug-free. Plot the solutions and the errors.

9.9 Matlab PDE-Toolbox Lab Exercises

Purpose: to learn the Matlab *Partial Differential Equation* toolbox.

Use the Matlab PDE toolbox to solve some typical second order PDE on some regions with various BC. Visualize the mesh triangulation and the solutions, and export the triangulation.

Reference: Partial Differential Equation Toolbox, MathWorks.

Test Problems

1. Poisson equation on a unit circle:

$$\begin{aligned} -\Delta u &= 1\,, \quad x^2 + y^2 < 1\,, \\ u|_{\partial\Omega} &= 0\,, \quad x \le 0\,, \\ u_n|_{\partial\Omega} &= 1\,, \quad x > 0\,. \end{aligned}$$

2. Wave equation on a unit square $x \in [-1,1] \times y \in [-1,1]$:

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} &= \Delta u\,, \\ u(x,y,0) &= \arctan\left(\cos\frac{\pi x}{2}\right), \\ u_t(x,y,0) &= 3\sin(\pi x) e^{\sin(\pi y/2)}\,, \\ u = 0 \text{ at } x = -1 \text{ and } x = 1\,, &\quad u_n = 0 \text{ at } y = -1 \text{ and } y = 1\,. \end{aligned}$$

3. Eigenvalue problem on an L-shape:

$$-\Delta u = \lambda u, \quad u = 0 \text{ on } \partial\Omega.$$

   The domain is the L-shape with corners (0,0) , $(-1,0)$, $(-1,-1)$, $(1,-1)$, (1,1) , and (0,1).

4. The heat equation:

$$\frac{\partial u}{\partial t} = \Delta u\,.$$

   The domain is the rectangle $[-0.5 \quad 0.5] \times [-0.8 \quad 0.8]$, with a rectangular cavity $[-0.05 \quad 0.05] \times [-0.4 \quad 0.4]$; and the BC are:

   - $u = 100$ on the left-hand side;
   - $u = -10$ on the right-hand side; and
   - $u_n = 0$ on all other boundaries.

5. Download the Matlab source code 2D.rar from

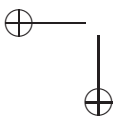
   `http://www4.ncsu.edu/~zhilin/FD\_FEM\_Book`

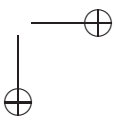
   Export the mesh of the last test problem from Matlab and run assemb.m to solve the example.

General Procedure

- Draw the geometry;

- define the BC;
- define the PDE;
- define the initial conditions if necessary;
- solve the PDE;
- plot the solution;
- refine the mesh if necessary; and
- save and quit.





Bibliography

[1] J. Adams, P. Swarztrauber, and R. Sweet. Fishpack: Efficient Fortran subprograms for the solution of separable elliptic partial differential equations. http://www.netlib.org/fishpack/.

[2] D. Braess. *Finite Elements*. Cambridge University Press, 1997.

[3] S.C. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer New York, 2002.

[4] R. L. Burden and J. D. Faires. *Numerical Analysis*. 2010, PWS-Kent Publ. Co. 2006, Brooks &. Cool, 9 ed, 2010.

[5] D. Calhoun. A Cartesian grid method for solving the streamfunction-vorticity equation in irregular regions. *J. Comput. Phys.*, 176:231–275, 2002.

[6] G. F. Carey and J. T. Oden. *Finite Element, I-V*. Prentice-Hall, Inc, Englewood Cliffs, 1983.

[7] A. J. Chorin. Numerical solution of the Navier-Stokes equations. *Math. Comp.*, 22:745–762, 1968.

[8] P. G. Ciarlet. *The finite element method for elliptic problems*. North Holland, 1978, and SIAM Classic in Applied Mathematics 40, 2002.

[9] D. De Zeeuw. Matrix-dependent prolongations and restrictions in a blackbox multigrid solver. *J. Comput. Appl. Math.*, 33:1–27, 1990.

[10] L. C. Evans. *Partial Differential Equations*. AMS, 1998.

[11] G. Golub and C. Van Loan. *Matrix computations*. The Johns Hopkins University Press, 2nd ed., 1989.

[12] A. Iserles. *A First Course in the Numerical Analysis of Differential Equations*. Cambridge, 2008.

[13] Jr. J. E. Dennis and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. SIAM, 1996.

[14] C. Johnson. *Numerical Solution of Partial Differential Equations by the Finite Element Method*. Cambridge University Press, 1987.

[15] R. J. LeVeque. Clawpack and Amrclaw – Software for high-resolution Godunov methods. 4-th Intl. Conf. on Wave Propagation, Golden, Colorado, 1998.

[16] R. J. LeVeque. *Finite Difference Methods for Ordinary and Partial Differential Equations, Steady State and Time Dependent Problems*. SIAM, 2007.

[17] Z. Li and M-C. Lai. The immersed interface method for the Navier-Stokes equations with singular forces. *J. Comput. Phys.*, 171:822–842, 2001.

[18] Z. Li and C. Wang. A fast finite difference method for solving Navier-Stokes equations on irregular domains. *J. of Commu. in Math. Sci.*, 1:180–196, 2003.

[19] M. Minion. A projection method for locally refined grids. *J. Comput. Phys.*, 127:158–178, 1996.

[20] K. W. Morton and D. F. Mayers. *Numerical Solution of Partial Differential Equations*. Cambridge press, 1995.

[21] Y. Saad. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 7:856–869, 1986.

[22] G. Strang and G. J. Fix. *An Analysis of the Finite Element Method*. Prentice-Hall, 1973.

[23] J. C. Strikwerda. *Finite Difference Scheme and Partial Differential Equations*. Wadsworth & Brooks, 1989.

[24] J. W. Thomas. *Numerical Partial Differential Equations: Finite Difference Methods*. Springer New York, 1995.

Index