18 Finite Element Methods

18.1 The Big Picture

There are 4 key components of the finite element method:

1. Weak form in an infinite-dimensional space V.
2. Weak form in a finite dimensional space $V_h \subset V$.
3. Construction of a basis for V_h.
4. Assembly and solution of the resulting linear system of equations.

We will sketch these components in the context of Poisson's equation.

18.2 Poisson's equation

Let Ω be our domain with boundary $\Gamma = \partial\Omega$. Given $f \in C(\Omega)$ and $g_1, g_2 \in C(\Gamma)$, find $u \in C^2(\Omega) \cap C(\overline{\Omega})$ such that

$$-\nabla^2 u = f \quad \text{in } \Omega, \tag{1}$$

$$u = g_1 \text{ on } \Gamma_D, \tag{2}$$

$$\frac{\partial u}{\partial \mathbf{n}} = g_2 \text{ on } \Gamma_N, \tag{3}$$

where $\mathbf{n}$ is the outward pointing normal to Γ and $\Gamma = \Gamma_D \cup \Gamma_N$. The boundary conditions on Γ_D, which specify the value of the solution, are called *Dirichlet* conditions and those on Γ_N, which specify the value of the normal derivative, are called *Neumann* conditions.

The solution $u \in C^2(\Omega) \cap C(\overline{\Omega})$ is the *Classical solution*. The conditions on both u and f mean that we can evaluate both sides of equation 1 at any point $x \in \Omega$. Equation 1 describes *pointwise equality*.

Think back to your courses in Partial Differential Equations (PDEs). Have you seen any example of valid solutions to PDEs that do not satisfy these continuity requirements? [search for 'shock' or 'hydraulic jump' on FYFD, for example].

18.3 An alternative interpretation of equality

As we have seen, pointwise equality is too restrictive. Instead we will use an alternative interpretation of what it means for two functions to be equal, based on *integration* against *test functions*.

An example using vectors

Imagine you have two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$. You are not able to look at the components of these vectors but you can compute their *inner product* with any *test vector* $\mathbf{v} \in \mathbb{R}^n$. If you find that

$$\mathbf{a} \cdot \mathbf{v} = \mathbf{b} \cdot \mathbf{v} \quad \forall \mathbf{v} \in \mathbb{R}^n \tag{4}$$

then you know that $\mathbf{a} = \mathbf{b}$. One way to see this is to choose $\mathbf{v}$ in the following way:

$$v_i = \delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

With this choice of $\mathbf{v}$, which is allowed because equation 18.3 holds *for all* $\mathbf{v} \in \mathbb{R}^n$, we can simplify the lhs of equation 18.3:

$$\mathbf{a} \cdot \mathbf{v} = \sum_i a_i v_i \tag{6}$$

$$= \sum_i a_i \delta_{ij} \tag{7}$$

$$= a_j \tag{8}$$

and similarly for the rhs, to get

$$a_j = b_j \quad \forall j, \tag{9}$$

which is component-wise equality.

Inner products

An *inner product* is a function $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ with the following properties:

- Non-negativity: $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0 \quad \forall \mathbf{x} \in \mathbb{R}^n$ with $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \,\text{iff}\, \mathbf{x} = 0$.
- Linearity: $\langle \alpha \mathbf{x} + \beta \mathbf{y}, \mathbf{z} \rangle = \alpha \langle \mathbf{x}, \mathbf{z} \rangle + \beta \langle \mathbf{y}, \mathbf{z} \rangle \quad \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n, \forall \alpha, \beta \in \mathbb{R}$.
- Symmetry: $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

An appropriate inner product for the functions we will be considering is the L_2 inner product, defined by

$$\langle f(x), g(x) \rangle := \int_\Omega f(x) g(x) dx, \tag{10}$$

and it is defined for all *square integrable* functions f and g, i.e. those that satisfy

$$\int_\Omega |f(x)|^2 dx < \infty, \quad \int_\Omega |g(x)|^2 dx < \infty. \tag{11}$$

18.3.1 A weak form

We will now write down the weak form of equation 1. For simplicity, we will take and $g_1 = g_2 = 0$, and we will assume that our domain Ω is *polytopic* (i.e. a polygon or polyhedron). Let v, our *test function*, be any *sufficiently regular* function with $v = 0$ on Γ_D. Multiplying equation 1 by v and integrating over our domain gives

$$-\int_\Omega v\nabla^2 u dx = \int_\Omega f v dx. \tag{12}$$

Remember that we wish to reduce the differentiability requirements on our solution u. To do this we will integrate equation 12 by parts, transferring one of the derivatives from u to v.

Divergence theorem

$$\int_\Omega \nabla \cdot \mathbf{F} dx = \int_\Gamma \mathbf{F} \cdot \mathbf{n} ds, \tag{13}$$

where $\mathbf{n}$ is the outward pointing normal to Ω on Γ.

Green's identity

Applying the divergence theorem to $\mathbf{F} = v\nabla u$ we have

$$\int_\Omega \nabla u \cdot \nabla v dx + \int_\Omega v\nabla^2 u dx = \int_\Gamma v\nabla u \cdot \mathbf{n} ds. \tag{14}$$

This enables us to rewrite equation 12 as

$$\int_\Omega \nabla u \cdot \nabla v dx - \int_\Gamma v\nabla u \cdot \mathbf{n} ds = \int_\Omega f v dx, \tag{15}$$

which, due to the boundary conditions on v and $\frac{\partial u}{\partial \mathbf{n}} = \nabla u \cdot \mathbf{n}$, simplifies to

$$\int_\Omega \nabla u \cdot \nabla v dx = \int_\Omega f v dx. \tag{16}$$

In more abstract notation we can write this as

$$a(u, v) = F(v) \tag{17}$$

where

$$a(u, v) = \int_\Omega \nabla u \cdot \nabla v dx, \tag{18}$$

$$F(v) = \int_\Omega f v dx. \tag{19}$$

Now we can informally write down a definition of the function space V for which equation 17 makes sense for $u, v \in V$:

$$V = \{v : \Omega \to \mathbb{R} \text{ such that } a(v,v) < \infty \text{ and } F(v) < \infty \text{ and } v|_{\Gamma_D} = 0\} \tag{20}$$

Our abstract weak form of the problem can now be written as

$$\textit{find } u \in V \textit{ such that } a(u,v) = F(v) \quad \forall v \in V.$$

18.4 Galerkin approximation

Having formulated our weak form on the infinite dimensional function space V, we now aim to find an approximation on a finite dimensional subspace $V_h \subset V$.

This is the *Galerkin approximation*:

$$\textit{find } u_h \in V_h \textit{ such that } a(u_h, v_h) = F(v_h) \quad \forall v_h \in V_h.$$

Note that the discretisation has been applied to the solution, rather than the operators in the PDE. This means that the discretisation inherits some properties from the PDE, such as symmetry. We shall see more on this later.

18.5 Construction of the function spaces

The Galerkin approximation is general - there are many possibilities for constructing the finite dimensional subspace V_h. For the finite element method we make a particular choice of V_h, which we construct by covering our domain with a mesh made up of non-overlapping, *conforming*, cells (the *elements*), and specify *basis functions* that are piecewise polynomial within each cell. The cells are chosen to be simple shapes for which we can construct *quadrature rules* that enable us to integrate over each cell.

As a concrete example, we will use a mesh made up of triangles (a *triangulation* $\mathcal{T}$ of our domain), and the space P1 of piecewise linear polynomials on $\mathcal{T}$.

Triangulation

A triangulation $\mathcal{T}$ of our polytopic domain $\Omega \in \mathbb{R}^2$ is a set of triangles $\{K_i\}_{i=1}^N$, such that

- No triangles overlap: $\text{int}K_i \cap K_j = \emptyset$ where int denotes the interior of a set.
- $\mathcal{T}$ covers the domain: $\bigcup_{i=1}^N K_i = \overline{\Omega}$.
- Triangle vertices only meet other triangle vertices.

The P1 finite element space

Given a triangulation $\mathcal{T}$ of Ω, the P1 finite element space is the space of functions v defined by

$$V_h = \{v : v \in C^0(\Omega) \text{ and } v|_{K_i} \text{ is a linear function } \forall K_i \in \mathcal{T}\} \tag{21}$$

We can also define a subspace that accounts for our Dirichlet boundary conditions:

$$\mathring{V}_h = \{v \in V_h : v|_{\Gamma_D} = 0\}. \tag{22}$$

The Neumann boundary conditions are dealt with differently and we shall see this later

For this choice of V_h, each $v \in V_h$ is uniquely determined by its values at the vertices, which motives our choice of basis: the *nodal basis*.

Nodal basis

Given the locations of the vertices $x_i, i = 0, 1, ..., N-1$, the associated nodal basis $\phi_i, i = 0, 1, ..., N-1$ is given by

$$\phi_i(x_j) = \delta_{ij}. \tag{23}$$

This choice of V_h confers several advantages:

- most basis functions decouple - their supports do not intersect, which means that the resulting linear system is *sparse*,
- arbitrary triangulation can approximate complex geometries
- we can increase the accuracy of our approximation by enlarging V_h in two ways:
  - refine the mesh
  - increase the order of the polynomial basis

18.6 Linear algebraic formulation

Now that we have specified V_h and its basis:

$$V_h = \text{span}\{\phi_0, ..., \phi_{N-1}\} \tag{24}$$

we can write our Galerkin approximation as a linear system of equations.

We expand u_h and v_h in terms of our basis:

$$u_h = \sum_{j=1}^{N-1} U_j \phi_j, \tag{25}$$

and

$$v_h = \sum_{i=1}^{N-1} V_i \phi_i. \tag{26}$$

Our aim is now to calculate the *coefficients* U_j to give us our solution. Expanding v_h in our Galerkin approximation gives:

$$a(u_h, v_h) = F(v_h) \tag{27}$$

$$a\left(u_h, \sum_i V_i \phi_i\right) = F(\sum_i V_i \phi_i) \tag{28}$$

$$\sum_i V_i a(u_h, \phi_i) = \sum_i V_i F(\phi_i). \tag{29}$$

Since this holds for all possible values of V_i we have

$$a(u_h, \phi_i) = F(\phi_i) \text{ for } i = 0, 1, ..., N-1 \tag{30}$$

Similarly, expanding u_h gives:

$$a\left(\sum_j U_j \phi_j, \phi_i\right) = F(\phi_i) \tag{31}$$

$$\sum_j U_j a(\phi_j, \phi_i) = F(\phi_i). \tag{32}$$

In matrix notation this becomes

$$AU = b, \tag{33}$$

where

$$A_{ij} = a(\phi_j, \phi_i) \tag{34}$$

$$b_i = F(\phi_i). \tag{35}$$

We now have a linear system of equations that we can solve.

19 Compatible finite element methods

So far we have considered finite element methods for equations with just one prognostic field. When we have more than one prognostic field we are not constrained to use the same finite element function space for each field. Methods that choose different function spaces for different prognostic fields are called *mixed* finite element methods. When that choice is made in order to create a discretisation that respects, or mimics, a vector calculus identity that is true for the continuous fields, these methods are called *mimetic* or *compatible* finite element methods. These methods are a generalisation of the staggered grid finite element methods we saw earlier in the course and can be used as the basis for energy conserving finite element methods.

19.1 Linear shallow water equations

The linear shallow water equations are:

$$\mathbf{u}_t + f\mathbf{u}^\perp = -c^2 \nabla \eta, \tag{36}$$

$$\eta_t + \nabla \cdot \mathbf{u} = 0, \tag{37}$$

where $\mathbf{u}$ is the velocity, $h = H(1+\eta)$ is the layer thickness, f is the Coriolis parameter, $\mathbf{u}^\perp = \mathbf{k} \times \mathbf{u}$ with $\mathbf{k}$ the vertical normal, $c^2 = gH$ with g is the acceleration due to gravity and H is the mean depth. To obtain the weak form we multiply by test functions $w \in V_u$ and $\phi \in V_\eta$, where V_u and V_η will be our finite element function spaces for $\mathbf{u}$ and h respectively. We choose finite element function spaces with the properties that:

$$\underbrace{\mathbb{V}_\phi}_{\text{Continuous}} \xrightarrow{\nabla^\perp} \underbrace{\mathbb{V}_u}_{\text{Continuous normals}} \xrightarrow{\nabla\cdot} \underbrace{\mathbb{V}_\eta}_{\text{Discontinuous}}$$

1. if $\psi \in V_\psi$ then $\nabla^\perp \psi \in V_u$,
2. if $\boldsymbol{u} \in V_u$ with $\nabla \cdot \boldsymbol{u} = 0$ and $\int_\Omega \boldsymbol{u}\,\mathrm{d}x = 0$ then $\exists \psi \in V_\psi$ with $\boldsymbol{u} = \nabla^\perp \psi$,
3. if $\boldsymbol{u} \in V_u$ then $\nabla \cdot \boldsymbol{u} \in V_\eta$,
4. if $\eta \in V_\eta$ with $\int_\Omega \eta\,\mathrm{d}x = 0$ then $\exists \boldsymbol{u} \in V_u$ with $\nabla \cdot \boldsymbol{u} = \eta$.

Mimetic property: $\nabla \cdot \nabla^\perp \boldsymbol{u} = 0$

1. $\nabla\cdot$ maps from V_u onto V_η.
2. $\nabla^\perp$ maps from V_ψ onto the kernal of $\nabla\cdot$ in V_u.

We will see later that these properties of the finite element function spaces allow us to prove some important properties of our discretisation.

Using these function spaces on a domain Ω with periodic boundary conditions and integrating by parts to reduce the differentiability requirements on η gives:

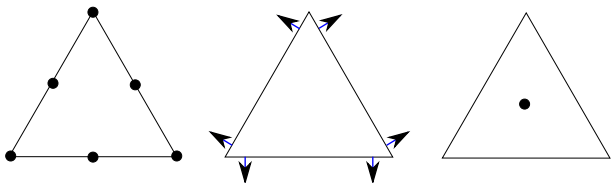
$$\int_\Omega \mathbf{w} \cdot \mathbf{u}_t\,dx + f\int_\Omega \mathbf{w} \cdot \mathbf{u}^\perp\,dx - c^2\int_\Omega \nabla \cdot \mathbf{w}\eta\,dx = 0, \tag{38}$$

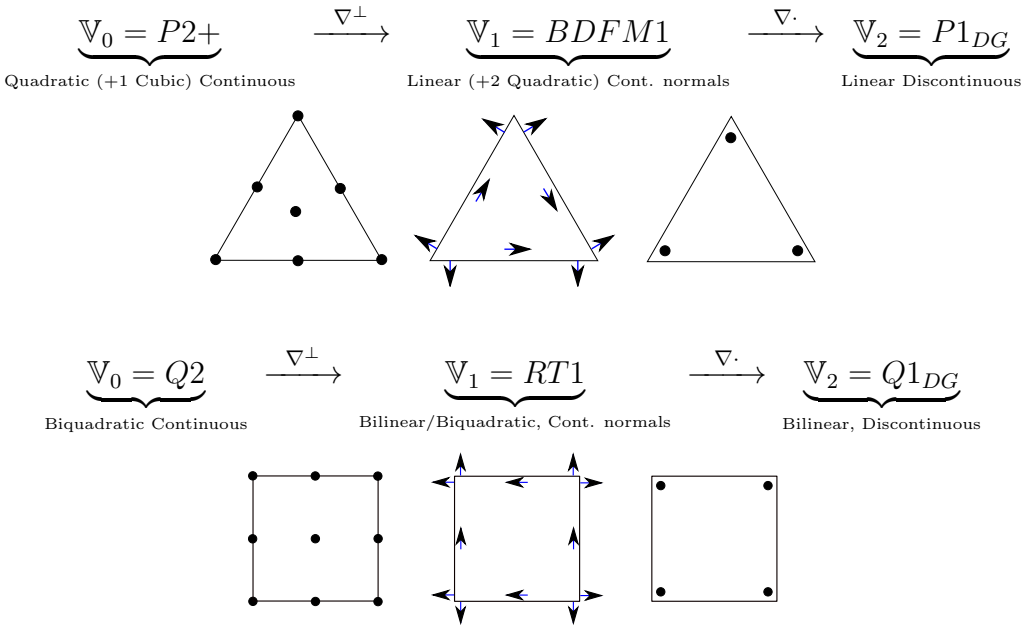
$$\int_\Omega \phi\eta_t\,dx + \int_\Omega \phi\nabla \cdot \mathbf{u}\,dx = 0. \tag{39}$$

19.1.1 Compatible finite element function spaces in 2D

There are several families of finite element function space that work for linear shallow water. Here are some examples:

$$\underbrace{\mathbb{V}_0 = P2}_{\text{Quadratic, Continuous}} \xrightarrow{\nabla^\perp} \underbrace{\mathbb{V}_1 = BDM1}_{\text{Linear, Continuous normals}} \xrightarrow{\nabla\cdot} \underbrace{\mathbb{V}_2 = P0}_{\text{Constant, Discontinuous}}$$





19.1.2 Energy conservation

Global energy conservation is required for the linear equations as it prevents numerical sources of fast waves. The conserved evergy for this system of equations is:

$$E = \frac{1}{2}\int_\Omega |\mathbf{u}|^2 + c^2\eta^2\, dx. \tag{40}$$

Looking at how this changes in time, we have

$$\frac{dH}{dt} = \int_\Omega \boldsymbol{u}\cdot\boldsymbol{u}_t + c^2\eta\eta_t\, dx \tag{41}$$

Since equations 38 hold for all $\boldsymbol{w} \in V_u$ and for all $\phi \in V_\phi$, we can take $\boldsymbol{w} = \boldsymbol{u}$ and $\phi = \eta$ in 38, and rearrange to get:

$$\begin{aligned}
\frac{dH}{dt} &= \int_\Omega \boldsymbol{u}\cdot\boldsymbol{u}_t + c^2\eta\eta_t\, dx, && (42)\\
&= \int_\Omega -f\boldsymbol{u}\cdot\boldsymbol{u}^\perp + c^2\nabla\cdot\boldsymbol{u}\eta - c^2\nabla\cdot\boldsymbol{u}\, dx, && (43)\\
&= 0, && (44)
\end{aligned}$$

since the first term is zero and the final two terms cancel out.

19.1.3 Local mass conservation

Numerical weather prediction models require local mass conservation as it prevents spurious sources and sinks of fluid. To prove that our discretisation has this property, we again use the fact that equations 38 hold for all test functions in the chosen function spaces. Noting that element indicator functions (i.e. those functions that are equal to 1 in one element and 0 in all others) are in V_η. Selecting the element indicator function for element e as the test function in the prognostic equation for η (the second equation in 38) we have

$$\frac{d}{dt}\int_e \eta\, dx + \int_e \nabla\cdot\boldsymbol{u}\, dx = 0, \tag{45}$$

which, on using the divergence theorem, becomes:

$$\frac{d}{dt}\int_e \eta\, dx + \int_{\partial e} \boldsymbol{u}\cdot\boldsymbol{n}\, dS = 0. \tag{46}$$

Since $\boldsymbol{u}$ has continuous normal components on element boundaries, this means that the flux of η is continuous and hence η is locally conserved.

19.1.4 Absence of spurious pressure modes

Spurious pressure modes are those that have a vanishing numerical gradient despite being oscillatory in space. These modes can pollute the numerical solution and are easily excited by initialisation from noisy data or by physics parameterisation schemes. Proof of this properties requires some more involved finite element theory (see Cotter CJ, Shipton J. Mixed finite elements for numerical weather prediction. Journal of Computational Physics. 2012 Aug 30;231(21):7076-91.)

19.1.5 Steady geostrophic modes

When the Coriolis parameter f is constant, geostrophically balanced states are steady. When f varies, as on the sphere, and nonlinear terms are introduced, these steady states become the slowly-evolving states that give rise to the large scale weather systems that forecasters want to predict. It is crucial that a numerical discretisation preserves these steady geostrophic states in the linear case otherwise the addition of nonlinearity and spherical geometry will lead to the emission of noise which generates fast inertia-gravity waves that pollute the numerical solution.

Starting from a geostrophically balanced state corresponding to a streamfunction $\psi \in V_\psi$, we set $\boldsymbol{u} = \nabla^\perp\psi$ and calculate η from the geostrophic balance relation:

$$c^2\int_\Omega \phi\eta\, dx = f\int_\Omega \phi\nabla\psi\, dx. \tag{47}$$

We can now check using equations 38 that for this choice of $\boldsymbol{u}$ and η, $\frac{d\boldsymbol{u}}{dt} = 0$ and $\frac{d\eta}{dt} = 0$ (Ex: try this and let me know if you have any questions.)