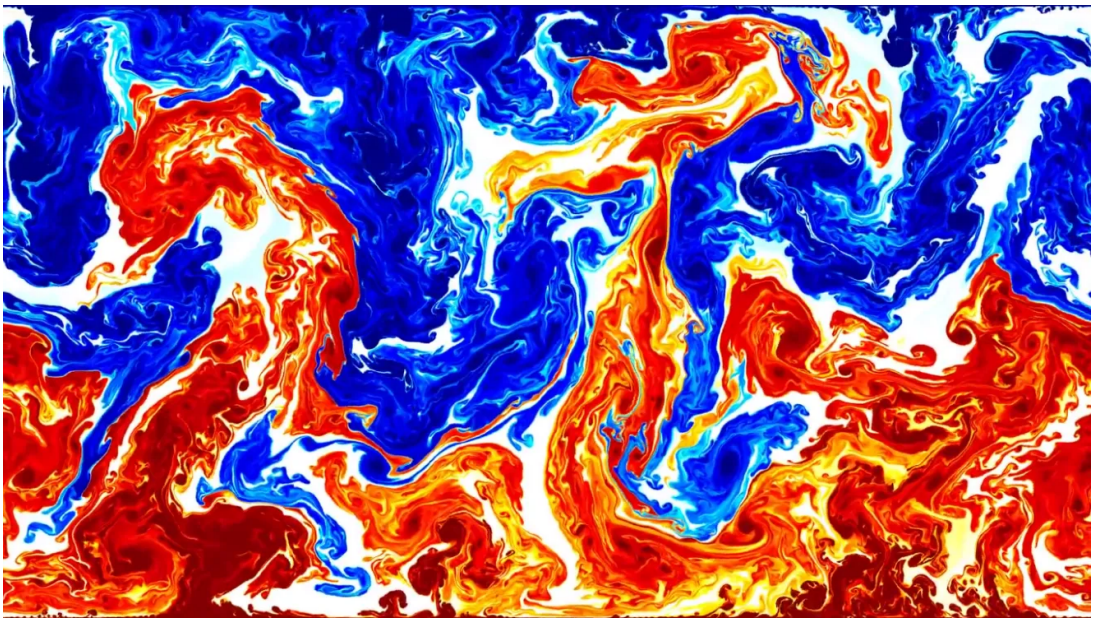
Numerical Analysis for Partial Differential Equations



Rayleigh–Bénard convection

Held by Prof. Antonietti Paola at Politecnico di Milano 2023/2024

Notes by Rayan Emara

Table of contents

1. Disclaimers and preface

These notes were taken during AY 2023/2024 using older material, your mileage may vary. They're meant to accompany the lectures and in no way aim to substitute lectures.

These notes are in part based on material by **Ravizza**

For any questions/mistakes you can reach me here.

Part 1

1. Boundary-value problems

In general, these types of problems are written as: "Some operator applied to some function u set to some value".

$$\begin{cases} \mathcal{L}u = f & \text{in } \Omega \\ \text{BC} & \text{in } \partial\Omega \end{cases}$$

Some specific examples include the diffusion problem, where:

$$\mathcal{L}u = -\text{div}(\mu(x)\nabla u)$$

Where $\mu(x)$ is the diffusion coefficient which is positive almost everywhere (from here on out a.e.).
There's also the ADR problem where:

$$\mathcal{L}u = -\text{div}(\mu(x)\nabla u) + \underline{b} \cdot \nabla u + \sigma u$$

Where $\mu(x)$ is as above, $\underline{b} \in \mathbb{R}^d$ and $\sigma = \sigma(x) \geq 0$.
In PDE courses you're usually taught to think of these as at least $L^2(\Omega)$ functions but for the purposes of this course
we'll relax the constraints and assume them to be $L^\infty(\Omega)$ to simplify the analysis.
Here, have some notation, this is what we're going to assume:

$$\begin{cases} \mu(x) \in L^\infty(\Omega) \\ \sigma(x) \in L^\infty(\Omega) \\ \underline{b} \in [L^\infty(\Omega)]^d \\ f \in L^2(\Omega) \end{cases}$$

1.1. Weak formulation

Consider

$$\begin{cases} \mathcal{L}u = f & \text{in}\Omega \\ +\text{B.C.} & \text{on}\partial\Omega \end{cases}$$

where Ω is an open bounded domain in $\mathbb{R}^d$ where d is the number of dimensions and $\mathcal{L}$ is a 2^{nd} order differential operator.

Examples of 2^{nd} order operators and a general boundary value problem

The following are respectively a non-conservative and a conservative form

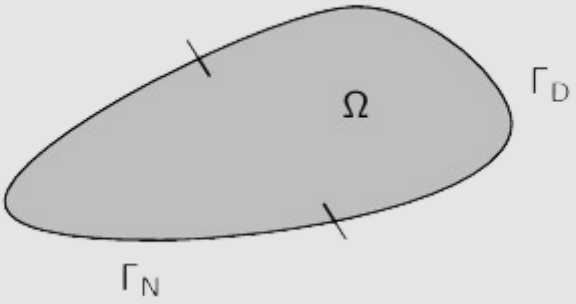
$$\mathcal{L}u = -\text{div}(\mu\nabla u) + \mathbf{b} \cdot \nabla u + \sigma u$$

$$\mathcal{L}u = -\text{div}(\mu\nabla u) + \text{div}(\mathbf{b}u) + \sigma u$$

The following is an example of an applied BVP

$$\begin{cases} \mathcal{L}u = -\operatorname{div}(\mu\nabla u) + \mathbf{b}\cdot\nabla u + \sigma u = f & \text{in } \Omega \\ u = 0 & \text{on } \Gamma_D \\ \mu\nabla u\cdot\mathbf{n} = g & \text{on } \Gamma_N \end{cases}$$

$$g \in L^2(\Gamma_N), \quad \partial\Omega = \Gamma_D \cup \Gamma_N, \quad \tilde{\Gamma}_D \cap \tilde{\Gamma}_N = \emptyset$$



This is a very general form which isn't very useful for numerical analysis, it's better to use some form of weak formulation, this will help us find an integral representation of the problem and derive numerical models from that.
Let's start quick and dirty, ignore the legality of steps, the regularity of v and just try to come up with an integral form, we'll worry about conditions later !

$$\int_\Omega [-\operatorname{div}(\mu(x)\nabla u) + \underline{b}\cdot\nabla u + \sigma u\cdot v] = \int_\Omega [f\cdot v]$$

Now, integrating by parts we get:

$$\int_\Omega \mu(x)\nabla u\cdot nv - \underbrace{\int_{\partial\Omega} \mu\nabla u\cdot nv}_{\text{notice the dominion}} + \int_\Omega \underbrace{b}\nabla uv + \int_\Omega \sigma uv = \int_\Omega fv \qquad \forall v$$

$$\underbrace{\int_\Omega \mu\nabla u\cdot\nabla v + \int_\Omega \mathbf{b}\cdot\nabla uv + \int_\Omega \sigma uv}_{=:z(u,v)}$$

$$= \int_\Omega fv + \underbrace{\int_{\Gamma_D} \mu\nabla u\cdot\mathbf{n}v}_{=0 \text{ if } v|_{\Gamma_D}=0} + \int_{\Gamma_N} \underbrace{\mu\nabla u\cdot\mathbf{n}}_{=g} v$$

whence

Abstract weak formulation

Find $u \in V$ such that

$$a(u,v) = F(v) \quad \forall v \in V$$

where $a : V \times V \to \mathbb{R}$ is a bilinear form and $F : V \to \mathbb{R}$ is a linear form $\langle F, v \rangle \equiv F(v) = \int_\Omega fv + \int_{\Gamma_N} gv$

Lax-Milgram Lemma

The following theorem provides sufficient conditions for the existence and uniqueness of a solution to a weakly formulated problem.
Let:

- V be a Hilbert space with norm $|| \cdot ||_V$ and inner product $(\cdot, \cdot)$
- $F \in V'$: $|F(v)| \leq \|F\|_{V'} \|v\|_V \forall v \in V$
- The bilinear form a is **continuous** meaning:

$$\exists M > 0 : |a(u, v)| \leq M \|u\|_V \|v\|_V \forall u, v \in V$$

- The bilinear form a is **coercive** meaning:

$$\exists \alpha > 0 : a(v, v) \geq \alpha \|v\|_V^2 \quad \forall v \in V$$

Then there exists a unique solution u for the abstract weak formulation.

More over

$$\alpha \|u\|_V^2 \leq a(u, u) = F(u) \leq \|F\|_{V'} \|u\|_V$$

therefore

$$\|u\|_V \leq \frac{\|F\|_{V'}}{\alpha}$$

which is a stability result that implies the continuous dependence of the solution from the data.

Note that I'm skipping the examples given by prof. Antonietti as they're already well described in the slides.

Poincarè inequality

Let Γ_D be a set of positive measure (in 1D it is sufficient that it contains a single point) then:

$$\exists C_{\mathbb{P}} > 0 : \|v\|_{L^2(\Omega)} \leq C_{\mathbb{P}} \|\nabla v\|_{L^2(\Omega)} \quad \forall v \in V = H^1_{\Gamma_{\mathbb{D}}}(\Omega)$$

which can be rewritten in terms of the $L^2(\Omega)$ norm starting from

$$\|v\|_V^2 = \|v\|_{L^2(\Omega)}^2 + \|\nabla v\|_{L^2(\Omega)}^2 \leq \left(1 + C_P^2\right) \|\nabla v\|_{L^2(\Omega)}^2$$

therefore

$$\|\nabla v\|^2_{L^2(\Omega)} \geq \left(1 + C_{\mathrm{P}}^2\right)^{-1} \|v\|^2_V$$

1.2. Galerkin approximation

We define V_h to be any finite dimensional subspace of V where $h > 0$.
We've seen how we can apply the Lax-Milgram lemma in order to find solutions for weakly formulated abstract problems. Given that $V_h \subset V$ we can conclude that if any problem of type

$$\text{Find } u \in V : a(u, v) = F(v) \quad \forall v \in V$$

has a solution then a problem of the following type will also have a solution

$$\text{Find } u_h \in V_h \colon a(u_h, v_h) = F(v_h) \quad \forall v_h \in V_h$$

where the $\dim(V_h) = N_h < +\infty$.
We'll denote the second problem as the (G) or Galerkin formulation problem.

Linear system equivalent for the galerkin problem

Problem (G) is equivalent to the following linear system of equations:

$$\text{Find } \boldsymbol{u} \in \mathbb{R}^{N_h} \text{ s. t. } \quad A\boldsymbol{u} = \boldsymbol{F}$$

where $A \in \mathbb{R}^{N_h \times N_h}$ and $F \in \mathbb{R}^{N_h}$

$$A = \begin{pmatrix} \dots & \dots & \dots \\ a_{i1} & \dots & a_{iN_h} \\ \dots & \dots & \dots \end{pmatrix}$$

Now. keep in mind that each element of A is an integral and as such, before we feed this linear system into our laptop, we need to approximate those integrals numerically.

Another important consideration is that u_h is a projection of $u \in V$ in V_h, the assumption that u_h converges to u, as V_h gets larger and larger, is called **space saturation**. We'll be making use of space saturation for our three step solution finding method.

- Try to write your problem as an abstract problem in order to apply *Lax-Milgram.*
- Use the Galerkin paradigm.
- Choose a subspace that satisfies *space saturation* to automatically get convergence for our method.

Finally we'll write elements of V_h as linear combinations of basis functions for V_h

$$\left\{\phi_i(\mathbf{z})\right\}_{j=1}^{N_h}$$

there any $v_h \in V_h$ can be expanded as a linear combination of the basis in the following fashion

$$v_h(\mathbf{x}) = \sum_{j=1}^{N_h} v_j \phi_j(\mathbf{x})$$

where v_j are the coefficients that identify v_h. This means that finding u_h in the galerkin paradigm actually means finding the **coefficients** that identify u_h *given* the basis functions (we'll find out later why the choice of basis functions is actually very important for numerical reasons).

Analysis of the Galerkin method

- **Existence and uniqueness** is a consequence of the Lax-Milgram lemma given that $V_h \subset V$
- **Stability** is also a consequence of the Lax-Milgram lemma from which we a uniform bound with respect to h

$$\|u_h\|_V \leq \frac{\|F\|_{V'}}{\alpha}$$

- **Consistency (or Galerkin orthogonality)** is basically measuring how well we're projecting the actual solution in V_h, for which we have the following result

$$a(u - u_h\ ,\ v_h) = 0 \quad \forall v_h \in V_h$$

which can be proven by computing difference between $a(\cdot, \cdot)$ both in *weak-formulation* and *galerkin form* and using $v = v_h$

- **Convergence (Cèa Lemma)**:

$$\begin{aligned}
\alpha\|u - u_h\|_V^2 &\leq a(u - u_h, u - u_h) \\
&= a(u - u_h, u - v_h) + \underbrace{a(u - u_h, v_h - u_h)}_{=0(\text{Galerkin orthogonality})} \\
&\leq M\|u - u_h\|_V \|u - v_h\|_V \quad \forall v_h \in V_h
\end{aligned}$$

$$\|u - u_h\|_V \leq \frac{M}{\alpha}\|u - v_h\|_V \quad \forall v_h \in V_h$$

$$\|u - u_h\|_V \leq \frac{M}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_V$$

which is the best approximation we could hope for.

As previously mentioned, one of the assumptions we'll make on V is that it satisfies **space saturation**, in other words that

$$V_h \longrightarrow V \quad \text{when } h \longrightarrow 0$$

which implies

$$\forall v \in V \quad \lim_{h \to 0} \inf_{v_h \in V_h} \|v - v_h\|_V = 0$$

1.3. The Finite Element Method

As in other numerical schemes, we'll try to find a **tessellation** to break up our domain space. Then we'll construct a finite dimensional space made by piece-wise polynomials in H_0^1.

Definition

We'll call $\mathcal{T}_h = \bigcup K$ a triangulation of our discrete space Ω_h, for any $r \geq 1$

$$V_h = \{v_h \in \mathcal{C}^0(\overline{\Omega}) \colon v_h|_K \in \mathbb{P}^r(K) \forall K \in \mathcal{T}_h, v_h|_{\Gamma_0} = 0\}$$

So essentially each element K is comprised of a continuous (up to and including the boundary) polynomial of degree r or less, that go to zero at some boundary Γ_0.
We can prove that for a suitable choice of an interpolant

$$\inf_{v_h \in V_h} \|u - v_h\|_V \leq \|u - \overline{u}_h\|_V$$

we'll later see that if we take $\overline{u}_h = \Pi_h^r u$ then we get space saturation

$$\|u - \overline{u}_h\| \leq C h^r |u|_{H^{r+1}(\Omega)}$$

Properties of basis/shape functions

We'll use some specific notation

- We'll use $v = (v_1, \dots, v_{N_h})^T$ to denote a real vector containing all the basis coefficients (also called **degrees of freedom**)
- A basis is called **Lagrangian** if it satisifes the following property

  $$\phi_i(\mathbf{x}_j) = \delta_{ij}$$

  for a *suitable* collection of points called **nodes**
- When the basis is Lagrangian the following property holds

  $$v_h(\mathbf{x}_i) = v_i \quad \forall 1 \leq j \leq N_h$$

Suppose I have a generic function $v \in V$ of which I want to compute the finite element interpolant $\prod_h^1 v \in V_h$, we'll restrict ourselves to $r = 1$ for now.
By definition this object has to be a piece-wise continuous linear polynomial over my mesh so basically

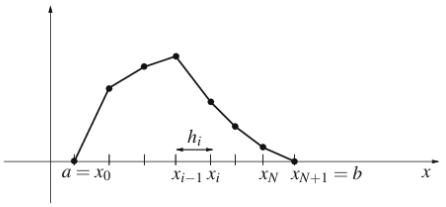
- $\phi_i(x)$ is a piece-wise continuous polynomial such that

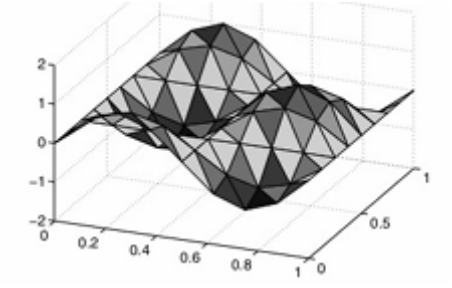
  $$\phi_i(\underline{x}_i) = 1 \text{ and } \phi_i(\underline{x}_{j \neq i}) = 1$$

We can then compute our **interpolant** (essentially the approximated version of our function) as the dot product between the coefficients and our basis functions

$$\prod_h^1 v(x) = \sum_{i=1}^{N_h} v(x_i)\phi_i(x)$$

Note that we set the basis function nodes to be zero on the boundary in order to comply with the boundary conditions, therefore we'd have something like $v_h(a) = v_h(b) = 0$ in the 1D case.





Note how overlapping between the basis functions is not allowed

A small (not in importance) note has to be made about the space V in which we're operating

$$V_h \subset V = H_0^1$$

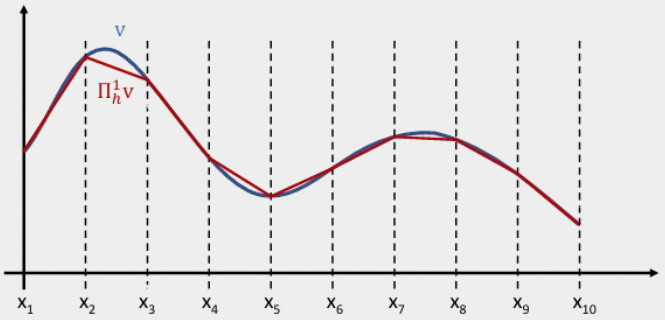
This is intuitively because V is comprised of functions that are (along with their gradients) L^2 while V_h is comprised of piece-wise polynomials that are trivially integrable on bounded intervals, the gradients of V_h are also integrable if we assume to *glue* them without jumps.

Error estimate and analysis

Bounds for the interpolation error

Let $r \geq 1$, $v \in H^{r+1}(\Omega)$, and $\prod_h^r$ to be the finite element interpolant of v at the finite element nodes, meaning

- $\Pi_h^r v \in X_h^r$
- $\Pi_h^r(\mathbf{x}_i) = v(\mathbf{x}_i) \quad \forall$ node $\mathbf{x}_i$ of T_h



Then, for $m = 0, 1$, $\exists C = C(r, m, \hat{k})$ such that:

$$|v - \Pi_h^r v|_{H^m(\Omega)} \leq C \left(\sum_{K \in \mathcal{T}_h} h_K^{2(r+1-m)} |v|^2_{H^{r+1}(K)} \right)^{1/2} \tag{8}$$

The constant then depends on r, the norm and the shape of the triangle. $h_K = \text{diam}(K)$ and since $h_K \leq h$ then $\forall K$ the following holds

$$|v - \Pi_h^r v|_{H^m(\Omega)} \leq C h^{r+1-m} |v|_{H^{r+1}(K)} \quad \forall v \in H^{r+1}(\Omega), m = 0, 1 \tag{9}$$

In essence (8) gives a localized element-wise interpolation error while equation (9) global bound

So know that the interpolation has an inherent error that thankfully goes to zero as h goes to zero. We can then try to get an estimate for $\|u - u_h\|_V$

$$\begin{aligned} \|u - u_h\|_V &= \|u - u_h\|_{H^1(\Omega)} \\ &\leq \frac{M}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_{H^1(\Omega)} \\ &\leq \frac{M}{\alpha} \|u - \Pi_h^r u\|_{H^1(\Omega)} \end{aligned}$$

where we can then use (8) and then (9) to get

$$\|u - u_h\|_V \leq C \frac{M}{\alpha} \left(\sum_{K \in \mathcal{T}_h} h_K^{2r} |u|^2_{H^{r+1}(\Omega)} \right)^{1/2} \quad \text{(using 8)}$$

$$\|u - u_h\|_V \leq C \frac{M}{\alpha} h^r |u|_{H^{r+1}(\Omega)} \quad \text{(using 9)}$$

Remember that we're assuming quite a bit of regularity, let s be the regularity for $u \in H$ then we have the following error estimates for u_H, remember that we're working on the ADR problem, we're therefore using $V = H^1$

$_r\backslash^s$	$s<2$	$s=2$	$s=3$	$s=4$
$r=1$	N/A	h^1	h^2	h^3
$r=2$	N/A	h^1	h^2	h^3
$r=3$	N/A	h^1	h^2	h^3

Convergence rates

Error estimates in the L^2 norm

We start by defining the **adjoint form** starting from a bilinear form

Adjoint form

Consider a bilinear form $a : V \times V \to \mathbb{R}$, the adjoint form a^* is defined as

$$\begin{aligned} &a^*: V \times V \to \mathbb{R} \\ &a^*(v,w) = a(w,v) \quad \forall v, w \in V \end{aligned}$$

It immediately follows that if a is symmetric then

$$a^* = a$$

An adjoint problem can be given in the following form

$$\begin{cases} \text{Find } \phi = \phi(g) \in V \\ a^*(\phi, v) = (g, v) = \int_\Omega gv \quad \forall v \in V \end{cases} \tag{12}$$

Example

Consider $\mathcal{L} = -\Delta$, then the solution of the Poisson problem

$$\begin{cases} -\Delta\phi = g & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

satisfies $\phi \in H^2(\Omega)$, moreover

$$\exists C_1 > 0 : \quad \|\phi(g)\|_{H^2(\Omega)} \leq C_1 \|g\|_{L^2(\Omega)} \tag{13}$$

This is a direct result of the fact that this a is symmetric in this case, therefore we can apply Lax-Milgram and conclude that there exists a unique $\phi \in V$ that continuously depends on the data.

We will now take g to be our error and do some magic ™ to get an error estimate in the L^2 norm.
Taking $g = e_h = u - u_h$ in (12)

$$\begin{aligned}\|e_h\|^2_{L^2(\Omega)} = (e_h, e_h) = a^*(\phi, e_h) &\underbrace{=}_{\text{by symmetry}} a(e_h, \phi) \\ &= a(e_h, \phi - \phi_h) \quad \text{(Galerkin orthogonality, for } \phi_h \in V_h) \\ &\leq M\|e_h\|_{H^1(\Omega)}\|\phi - \phi_h\|_{H^1(\Omega)}\end{aligned}$$

The second row is a consequence of the fact that the error $u - u_h$ is orthogonal to anything in the discrete space since it is a projection, we then use the continuity of the bilinear form and the fact that $V = H^1$ to get the inequality.
We assume $\phi \in H^2(\Omega) \cap V$, this is also referred to as **elliptic regularity**, we also take $\phi_h = \prod_h^1 \phi$, then

$$\begin{aligned}\|e_h\|^2_{L^2(\Omega)} &\leq M\|e_h\|_{H^1(\Omega)}\|\phi - \Pi_h^1\phi\|_{H^1(\Omega)} \\ &\leq M\|e_h\|_{H^1(\Omega)}C_2 h|\phi|_{H^2(\Omega)} \quad \text{(for (9) with } m = r = 1) \\ &\leq M\|e_h\|_{H^1(\Omega)}C_2 h C_1\|e_h\|_{L^2(\Omega)} \quad \text{(for (13))}\end{aligned}$$

then

$$\begin{aligned}\|e_h\|_{L^2(\Omega)} &\leq MC_1C_2h\|e_h\|_{H^1(\Omega)} \\ &\leq MC_1C_2hC_3h^r|u|_{H^{r+1}(\Omega)} \qquad \text{(for (11))}\end{aligned}$$

in conclusion

$$\|e_h\|_{L^2(\Omega)} \leq \overline{C}h^{r+1}|u|_{H^{r+1}(\Omega)}$$

In essence if the solution is regular enough (in $r = 1$) we get linear convergence in the energy norm but quadratic convergence in the L^2 norm. If the domain is convex the solution is in H^2 and the H^2 is bounded by the L^2 norm.
Finally it can be proven that if elliptic regularity isn't satisfied then the L^2 norm doesn't gain an order of convergence.

Stiffness matrix

Let's take the Poisson problem

$$\begin{cases}-\Delta\phi = f & \text{in } \Omega \\ u = 0 & \text{on } \Gamma_D\end{cases}$$

take $V = H_0^1(\Omega)$ the weak formulation then becomes

$$\int_\Omega \nabla u \cdot \nabla v = \int_\Omega fv \quad \forall\, v \in V$$

let

- Space be two dimensional $d = 2$
- Linear polynomials $r = 1$
- Basis functions be ϕ_i, also referred to as hat functions (associated to a vertex i)

recall that

- $\underline{u} \in \mathbb{R}^{N_h}$ is defined as

$$\underline{u} = [u_1, \ldots, u_{N_h}]^T$$

- A is a real valued $N_h \times N_h$ matrix

$$A_{i,j} = a(\phi_j, \phi_i) \quad \forall\, i, j = 1, \ldots, N_h$$

- $\underline{F}$ is a real valued N_h long vector defined as

$$F_i = F(\phi_i) \quad \forall\, i = 1, \ldots, N_h$$

We're going to define a special element called a **reference element** with a variable substitution on which we construct our shape functions.
The reference element is nothing but a equilateral triangle with coordinates $(0,0), (0,1), (1,0)$