Escuela de Verano 2015, EAFIT

Theoretical Aspects and Numerical Methods for Incompressible Fluids

Instructors: Charles Dapogny†, Pascal Frey‡

† Laboratoire J. Kuntzmann, Grenoble, France
‡ Laboratoire J.L. Lions, Sorbonne Universités, UPMC, Paris, France
Outline of the lectures

PART I - CLASSICAL METHODS

1. FLUID MECHANICS
   1.1 Notations, vectors, tensors
   1.2 Conservation laws,
   1.3 Flow models and simplifications.

2. THE STOKES MODEL
   2.1 Mathematical and numerical analysis, variational formulation,
   2.2 Finite element approximation, resolution,
   2.3 Unsteady Stokes problem
   2.4 The Finite Element Method

3. THE NAVIER-STOKES MODEL
   3.1 Steady state problem, analysis
   3.2 Discretization procedures
Outline of the lectures

**PART II - ADVANCED METHODS**

4. **TWO FLUID OR TWO-PHASE FLOW PROBLEMS**
   4.1 Problem statement, modelling
   4.2 Evolution of the interface: level set formalism, scheme
   4.3 Numerical resolution

5. **SHAPE OPTIMIZATION**
   5.1 Framework of shape optimization, examples
   5.2 Shape sensitivity analysis using shape derivatives
   5.3 Céa’s method to compute derivatives
   5.4 Numerical issues

6. **ERROR ESTIMATES AND MESH ADAPTATION**
   6.1 Residual and geometric estimates
   6.2 Mesh adaptation
   6.3 Mesh adaptation for level sets

A. **Appendix:** Variational approximation of linear problems
References

Functional and numerical analysis

References (2)

- **Fluid mechanics**
  

- **Computational Fluid Dynamics**
  
References (3)

- **Computational Fluid Dynamics (cont’d)**


- **Numerical programming**


References (4)

- Shape optimization


Introduction to Fluid Dynamics

• Fluid flows are governed by partial differential equations (PDE) which represents conservation laws (mass, momentum, and energy).

• Computational Fluid Dynamics (CFD): virtual experimental laboratory
  ○ consists in replacing suitably the PDE problem by algebraic equations that are then solved numerically using computers,
  ○ provides a prediction of fluid flows based on
    – mathematical models (continuous)
    – numerical schemes (discrete)
    – algorithmic techniques (meshers, solvers, visualization).

What is CFD?
Computational Fluid Dynamics (CFD) provides a qualitative (and sometimes even quantitative) prediction of fluid flows by means of
• mathematical modeling (partial differential equations)
• numerical methods (discretization and solution techniques)
• software tools (solvers, pre- and postprocessing utilities).

CFD enables scientists and engineers to perform 'numerical experiments' (i.e. computer simulations) in a 'virtual flow laboratory'

Why use CFD?
Numerical simulations of fluid flow (will) enable
• architects to design comfortable and safe living environment
• designers of vehicles to improve the aerodynamic characteristics
• chemical engineers to maximize the yield from their equipment
• petroleum engineers to devise optimal oil recovery strategies
• surgeons to cure arterial diseases (computational hemodynamics)
• meteorologists to forecast the weather and warn of natural disasters
• safety experts to reduce health risks from radiation and other hazards
• military organizations to develop weapons and estimate the damage
• CFD practitioners to make big bucks by selling colorful pictures :-)

real simulation

numerical simulation
The two pilars: experiments and simulations

Investigation of the flow patterns to understand the flow phenomena by

1. Experiments: description of the phenomena using measurements
   - main features: laboratory scale, one variable at a time, at a few locations, operating conditions
     \[ \Rightarrow \text{experiments are expensive, slow, sequential} \]
   - error sources: flow measurements, probes, instruments, interferences,
     \[ \Rightarrow \text{calibration, reproducibility} \]

2. Simulations: prediction of the phenomena using computers
   - main features: scale 1:1, high resolution, virtually any problem, run scenarios
     \[ \Rightarrow \text{software codes are versatile, portable, easy to modify, cheap, fast, parallel,} \]
   - error sources: hypothesis, modelling, discretization, implementation,
     \[ \Rightarrow \text{robustness, stability and convergence issues.} \]

Hence, CFD is a now a complement and tends to be a substitute to experiments.
Fluid characteristics

1. Flow properties (macroscopic)

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>density</td>
</tr>
<tr>
<td>$\mu$</td>
<td>viscosity</td>
</tr>
<tr>
<td>$p$</td>
<td>pressure</td>
</tr>
<tr>
<td>$v$</td>
<td>velocity</td>
</tr>
<tr>
<td>$T$</td>
<td>temperature</td>
</tr>
</tbody>
</table>

2. Flow classification

<table>
<thead>
<tr>
<th>Viscous</th>
<th>Inviscid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compressible</td>
<td>Incompressible</td>
</tr>
<tr>
<td>Laminar</td>
<td>Turbulent</td>
</tr>
<tr>
<td>Steady</td>
<td>Unsteady</td>
</tr>
<tr>
<td>Single-phase</td>
<td>Multi-phase</td>
</tr>
</tbody>
</table>

3. Reliability issues in simulations

- related to mathematical models: Stokes, Navier-Stokes, Euler, Saint-Venant, …
- input data may be inaccurate,
- accuracy related to computer features (memory, architecture),
- sensitivity of turbulence models,
- difficulty of tracking the interfaces between phase (or species).
1. Problem analysis:
   - physical phenomena, type of flow (laminar, steady, ...)
   - domain geometry, interfaces, free surfaces,
   - objectives: computation of integral values (lift, drag), shape optimization

2. Mathematical modelling:
   - select flow model, identify forces,
   - write conservation laws (mass, momentum, energy): partial differential equations,
   - define computational domain and specify boundary + initial conditions

3. Space and time discretization:
   - mesh generation, references and domains identification,
   - discrete weak formulation (finite elements),
   - approximation of temporal derivatives: explicit vs implicit scheme

4. Resolution and visualization:
   - solve sparse algebraic systems
   - compute derived quantities (streamlines, vorticity), visualization.
CFD analysis: problems at stake

1. Validation of models and certification of CFD codes:
   - check if model is adequate for solving problem
   - compare numerical solutions with experimental results
   - introduce sensitivity analysis
   - switch between models, domain geometry and boundary conditions
   *i.e.* the goal is to ensure that the codes produce *reasonable results* for a certain range of flow problems

2. Uncertainty: usually related to the lack of knowledge (e.g. turbulence models)

3. Errors: may have various causes:
   - physical modelling due to (over) simplifications or incorrect parameter values
   - approximation of PDEs (space and time discretizations)
   - convergence of iterative procedures
   - round-off (truncation) of computer arithmetic, computer programming
Several CFD software are considered robust, reliable and efficient for performing accurate simulations. But they all require a knowledge of the underlying numerical methods and physics.

Below is a (non exhaustive) list of general-purpose free CFD softwares:

<table>
<thead>
<tr>
<th>Software</th>
<th>License</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>OpenFOAM</td>
<td>open-source</td>
<td><a href="http://www.openfoam.com">http://www.openfoam.com</a></td>
</tr>
<tr>
<td>Geris</td>
<td>open-source</td>
<td><a href="http://gfs.sourceforge.net">http://gfs.sourceforge.net</a></td>
</tr>
<tr>
<td>FeatFlow</td>
<td>open-source</td>
<td><a href="http://www.featflow.de">http://www.featflow.de</a></td>
</tr>
<tr>
<td>OpenFlower</td>
<td>open-source</td>
<td><a href="http://openflower.sourceforge.net">http://openflower.sourceforge.net</a></td>
</tr>
</tbody>
</table>

A list of commercial codes can be found at:
[http://www.cfd-online.com/Wiki/Codes#Commercial_codes](http://www.cfd-online.com/Wiki/Codes#Commercial_codes)

Recent research approaches (stochastic models, error estimates, complex models) have not been yet fully integrated and require further developments.
Schedule

Mornings: lectures 21 hours

- **day 1**: Preliminaries, fluid models 3 hours
  introduction to fluid problems and derivation of the incompressible fluid models (conservation laws, physical considerations,..);

- **day 2**: Stokes I 3 hours
  theoretical and numerical analysis of steady Stokes problem, variational formulation, a saddle point approach, FE approximation of the problem, solution methods;

- **day 3**: Stokes II + Navier-Stokes I 1.5 + 1.5 hours
  unsteady Stokes analysis, discretization in time, numerical issues. Navier-Stokes model, variational formulation, theoretical results;

- **day 4**: Navier-Stokes II 3 hours
  Navier-Stokes problem, nonlinear iterative procedures, time discretization procedures, numerical issues;

- **day 5**: Two fluids or two-phase flows 3 hours
  problem statement, level set formulation, numerical resolution.

- **days 6 + 7**: Introduction to shape optimization 3 + 3 hours
  Framework, shape sensitivity - shape derivatives, Céa’s method, numerical issues.

Afternoons: numerical experiments using FreeFem++ 21 hours.
Chapter I

Fluid Mechanics
Section 1.1

Notations, vectors, tensors
# Differential operators

$(x_1, \ldots, x_d)$ Cartesian coordinates of point $x \in \mathbb{R}^d$

$\partial_i u$ partial derivative (in the distributional sense) of $u$ with respect to $x_i$

$\partial_i^m u$ partial derivative (in the distributional sense) of order $m$ of $u$ with respect to $x_i$, $m \in \mathbb{N}$

$\partial_{ij} u$ partial derivative (in the distributional sense) of $u$ with respect to $x_i$ and $x_j$

$\alpha$ multi-index: $\alpha = (\alpha_1, \ldots, \alpha_d)^t \in \mathbb{N}^d$ with $\alpha_i \in \mathbb{N}$ for every $i \in \{1, \ldots, d\}$

$|\alpha|$ length of the multi-index: $\alpha : |\alpha| = \alpha_1 + \cdots + \alpha_d$

$\partial^\alpha u = \partial_{x_1}^{\alpha_1} \ldots \partial_{x_d}^{\alpha_d}$

$\nabla u$ gradient of $u$: $\nabla u = (\partial_1 u, \ldots, \partial_d u)^t \in \mathbb{R}^d$ if $u$ is a scalar function, and $\nabla u = (\partial_j u_i)_{1 \leq i \leq m, 1 \leq j \leq d} \in \mathbb{R}^{m,d}$ if $u$ is a function with values in $\mathbb{R}^m$

$\nabla \cdot u$ divergence of $u$: $\nabla \cdot u = \text{div} u = \sum_{i=1}^d \partial_i u_i$ where $u$ is a function with values in $\mathbb{R}^d$

$\nabla \times u$ curl of $u$: $\nabla \times u = (\partial_2 u_3 - \partial_3 u_2, \partial_3 u_1 - \partial_1 u_3, \partial_1 u_2 - \partial_2 u_1)^t$ where $u$ is a function with values in $\mathbb{R}^3$

$\Delta u$ Laplacian of $u$: $\Delta u = \nabla \cdot (\nabla u) = \sum_{i=1}^d \partial_{ii} u$
Vectors in fluid dynamics

- **Velocity gradient**

\[
\nabla u = (\nabla u_x, \nabla u_y, \nabla u_z) = \begin{pmatrix}
\partial x u_x & \partial x u_y & \partial x u_z \\
\partial y u_x & \partial y u_y & \partial y u_z \\
\partial z u_x & \partial z u_y & \partial z u_z \\
\end{pmatrix}
\]

The trace of \( \nabla u \) is \( \text{div } u = \nabla \cdot u \)

- **Deformation rate tensor**

\[
D(u) = \frac{1}{2}(\nabla u + (\nabla u)^t) = \frac{1}{2}
\begin{pmatrix}
2\partial x u_x & (\partial x u_y + \partial y u_x) & (\partial x u_z + \partial z u_x) \\
(\partial y u_x + \partial x u_y) & 2\partial y u_y & (\partial y u_z + \partial z u_y) \\
(\partial z u_x + \partial x u_z) & (\partial z u_y + \partial y u_z) & 2\partial z u_z \\
\end{pmatrix}
\]

- **Convective derivative**

\[
u \cdot \nabla u = u_x \partial x u + u_y \partial y u + u_z \partial z u
\]
Let consider $T = \{t_{ij}\} \in \mathbb{R}^{3 \times 3}$, $u = (u_1, u_2, u_3) \in \mathbb{R}^3$ and $\alpha \in \mathbb{R}$. We recall:

1. $\alpha T = \{\alpha t_{ij}\}$,

2. $T^1 + T^2 = \{t^1_{ij} + t^2_{ij}\}$

3. $u \cdot T = (u_1, u_2, u_3) \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix} = \sum_{i=1}^{3} u_i \underbrace{(t_{i1}, t_{i2}, t_{i3})}_{i\text{-th row}}$

4. $T \cdot u = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \sum_{j=1}^{3} \begin{pmatrix} t_{1j} \\ t_{2j} \\ t_{3j} \end{pmatrix} u_j \ (j\text{-th column})$

5. $T^1 \cdot T^2 = \begin{pmatrix} t^1_{11} & t^1_{12} & t^1_{13} \\ t^1_{21} & t^1_{22} & t^1_{23} \\ t^1_{31} & t^1_{32} & t^1_{33} \end{pmatrix} \begin{pmatrix} t^2_{11} & t^2_{12} & t^2_{13} \\ t^2_{21} & t^2_{22} & t^2_{23} \\ t^2_{31} & t^2_{32} & t^2_{33} \end{pmatrix} = \left\{ \sum_{k=1}^{3} t^1_{ik} t^2_{kj} \right\}$

6. $T^1 : T^2 = \text{tr}(T^1 \cdot (T^2)^t) = \sum_{i=1}^{3} \sum_{k=1}^{3} t^1_{ik} t^2_{ik}$
Gauss-Green theorem

- Let $\Omega$ be a bounded, open subset of $\mathbb{R}^n$ and $\partial \Omega$ is $C^1$. Then, along $\partial \Omega$ is defined the outward unit normal vector field $n$.

- Let $u \in C^1(\bar{\Omega})$, the normal derivative of $u$ is defined by: $\frac{\partial u}{\partial n} = \nabla u \cdot n = \frac{\partial u}{\partial x} \cdot n$.

**Theorem 1 (Gauss-Green Theorem)** Suppose $u \in C^1(\bar{\Omega})$; then

$$\int_{\Omega} \frac{\partial u}{\partial x_i} \, dx = \int_{\partial \Omega} u n_i \, ds \quad (i = 1, \ldots, n).$$

**Theorem 2 (Integration by Parts Formula)** Let $u, v \in C^1(\bar{\Omega})$; then

$$\int_{\Omega} \frac{\partial u}{\partial x_i} v \, dx = -\int_{\Omega} u \frac{\partial v}{\partial x_i} \, dx + \int_{\partial \Omega} uv n_i \, dx \quad (i = 1, \ldots, n).$$

**Theorem 3 (Green’s Formulas)** Let $u, v \in C^2(\bar{\Omega})$; then

1. $\int_{\Omega} \Delta u \, dx = \int_{\partial \Omega} \frac{\partial u}{\partial n} \, ds$;

2. $\int_{\Omega} \nabla v \cdot \nabla u \, dx = -\int_{\Omega} u \Delta v \, dx + \int_{\partial \Omega} \frac{\partial v}{\partial n} u \, ds$;

3. $\int_{\Omega} u \Delta v - v \Delta u \, dx = \int_{\partial \Omega} u \frac{\partial v}{\partial n} - v \frac{\partial u}{\partial n} \, ds$. 
Section 1.2
Conservation laws
Mathematical models

Partial Differential Equations (PDE) form the kernel of very many mathematical models and are used to describe physical, chemical, biological, economical phenomena, . . .

- **Mathematical Modelling:**
  - **role:** description of the behavior of a system using a set of PDEs endowed with initial and boundary conditions;
  - **challenge:** theoretical framework for mathematical analysis;
  - **well-posedness:** *existence, uniqueness* (or multiplicity) of solutions and *stability* w/r to the data.

- **Numerical Simulation:**
  - **target:** imitating / reproducing the behavior of the complex system to analyze and possibly improve it;
  - **methods:** relies on mathematical analysis, numerical analysis, scientific computing.
Numerical simulation

• has become a consequent part of the mathematical modelling process and is possibly the sole that can give access to specific solutions that cannot be solved analytically;

• the choice of the approximation method is strongly related to the model and to the nature and the behavior of the phenomenon represented as well as to the desired properties of the solutions (regularity, stability, etc.);

• the quality of the approximation method is related to mathematical properties (stability, convergence and order of convergence, error estimate) and to more practical issues (accuracy, reliability, efficiency). Both criteria are often antagonists and incompatible with one another.

• numerical analysts shall then try to obtain a good compromise between these criteria, preserving the accuracy of the solution at a reasonable computational effort.
From modelling to simulation

• the mathematical modelling and numerical simulation of a PDE problem is a highly multi-disciplinary topic to obtain the numerical solution;

• process requires advanced knowledge of theoretical and applied mathematics and of computer science.

• the numerical simulation bridges the gap between theoretical analysis and experiments.

![Diagram](image)

Figure 1: *A continuum from modelling to simulation.*
**Objective:** describe the motion of a fluid filling a domain $\Omega_t \subset \mathbb{R}^d, t > 0$, given $\Omega_0$.

**Representations:** motion of fluid flow can be modeled in two ways:

1. In the **Lagrangian** description, the trajectory $t \mapsto \chi(x_0, t, t_0)$ of any individual fluid particle, occupying position $x_0$ at time $t_0$, is followed during the time period. This approach is widely used in solid and particle mechanics, but yields more tedious analysis in the case of fluid flow.

2. In the **Eulerian** representation, the study takes place in a fixed referential, and no particular fluid particle is followed. Each point in the domain $\Omega_t \subset \mathbb{R}^d$ -whose Cartesian coordinates are denoted by $x = (x_i)_{i=1,\ldots,d}$ - is considered over the time period, independently of which specific particle is occupying the place.
Notations

Let \( u(x, t) \) denote the velocity of the fluid particle occupying the position \( x \) in \( \Omega_t \) at time \( t \). At each instant \( t \), \( u \) is a vector field on \( \Omega_t \), called the velocity field of the fluid.

**Definition 1**

1. The flow field is said to be conservative if there exists a scalar function \( \phi(x, t) \) (called a scalar potential)

\[
u = \nabla \phi.
\]

such that

2. The vorticity \( \omega \) of the flow is defined as:

\[
\omega = \text{curl}(u) = \nabla \times u.
\]

The flow is said irrotational (or curl-free) if \( \omega = 0 \).

3. The flow field is said incompressible (or divergence-free) if the material density \( \rho \) is constant in time and space; this is equivalent to:

\[
\text{div } u = \nabla \cdot u = 0.
\]

4. The deformation tensor \( D = D(u) \) of the flow is defined as:

\[
D(u) = \frac{1}{2}(\nabla u + \nabla u^t).
\]
Physical meaning

**Proposition 1** Let \( u : U \rightarrow \mathbb{R}^d \) any vector field, defined over a subset \( U \subset \mathbb{R}^d \). Then the following asymptotic expansions hold:

\[
(d = 2), \quad u(x + h) = u(x) + \nabla u(x).h + \frac{1}{2} \text{curl}(u)(x) h^\perp + O(h^2).
\]

\[
(d = 3), \quad u(x + h) = u(x) + \nabla u(x).h + \frac{1}{2} \text{curl}(u)(x) \times h + O(h^2).
\]

Hence, locally in the vicinity of any point \( x \), \( u \) is submitted to:

1. an **infinitesimal deformation** \( \nabla u \), which admits \( d \) principal deformation directions: the principal directions of the symmetric matrix \( \nabla u \).

2. an **infinitesimal rotation** \( h \mapsto \frac{1}{2} \text{curl}(u)(x) \times h \).

   If \( d = 3 \), the axis of the rotation is the unit vector in the direction of \( \text{curl}(u)(x) \), and the speed of the rotation is proportional to the modulus of \( \text{curl}(u)(x) \).
Vector calculus

- **Exercises**: prove the following identities (\( u \) vector field, \( f \) scalar function)
  - \( \text{div}(fu) = u \cdot \nabla f + f \text{div} u; \)
  - \( \text{div(curl} \ u) = \text{div}(\nabla \times u) = 0 \) and \( \text{curl}(\nabla f) = 0; \)
  - \( \text{curl(curl} \ u) = \nabla (\text{div} u) - \Delta u. \)

- The **divergence theorem** relates the flow of a vector field through a surface to the behavior of the vector field in the domain.

**Theorem 4 (Divergence Theorem)** Suppose \( \Omega \) is a compact subset of \( \mathbb{R}^3 \) with a Lipschitz-continuous boundary \( \partial \Omega \). Then, if \( v \in C^1(\Omega) \) is a continuously differentiable vector field defined on a neighborhood of \( \Omega \), then we have

\[
\int_{\Omega} \nabla \cdot v \, dx = \int_{\partial \Omega} v \, n \, ds,
\]

where \( n \) is the outward pointing unit normal field of the boundary \( \partial \Omega \).
Motion of a particle

- The motion of a fluid, filling a domain $\Omega_t$, is considered during the time $I = [t_0, t]$.
- Usually, $\Omega_{t_0}$ (resp. $\Omega_t$) is called the reference configuration (resp. current configuration).
- In a Lagrangian perspective, the motion is described by a family of mappings and we denote by $\chi(\cdot, t, t_0)$ the diffeomorphism:

$$x_0 \in \Omega_{t_0} \mapsto x = \chi(x_0, t, t_0) \in \Omega_t,$$

which maps the position $x_0$ of a considered particle at time $t_0$ to its position $x$ at $t$.

**Definition 2** The trajectory of a particle of fluid is the curve $\{\chi(x_0, t, t_0)\}_{t \in I}$.

The velocity of this fluid particle at position $x$ at time $t$ is then the vector $u$, independent of the reference time $t = 0$, i.e.

$$u = u(x, t) = u(\chi(x_0, t, t_0), t) = \frac{\partial \chi}{\partial t}(x_0, t, t_0). \quad (2)$$
**Definition 3** The acceleration $\gamma(x, t)$ of the particle at time $t$ and at $x = \chi(x_0, t, t_0)$ is:

$$\gamma(x, t) = \gamma(\chi(x_0, t, t_0), t) = \frac{d}{dt} u(x, t) = \frac{d}{dt} u(\chi(x_0, t, t_0), t)$$

$$= \frac{\partial u}{\partial t}(\chi(x_0, t, t_0), t) + \sum_i \frac{\partial u}{\partial x_i}(\chi(x_0, t, t_0), t) \frac{\partial \chi_i}{\partial t}(x_0, t, t_0), \quad (3)$$

which also reads, using (2)

$$\gamma(\chi(x_0, t, t_0), t) = \frac{\partial u}{\partial t}(\chi(x_0, t, t_0), t) + \sum_i \frac{\partial u}{\partial x_i}(\chi(x_0, t, t_0), t) u_i(\chi(x_0, t, t_0), t)$$

$$= \frac{\partial u}{\partial t}(\chi(x_0, t, t_0), t) + (u \cdot \nabla) u(\chi(x_0, t, t_0), t). \quad (4)$$
Motion of a particle (3)

Considering the inverse mapping $\chi^{-1}$, for $x \in \Omega_t$ we deduce from the previous formula

$$\gamma(x, t) = \frac{\partial u}{\partial t} + (u \cdot \nabla) u(x, t)$$

which is known as the fundamental Eulerian formula for acceleration of fluids.

**Definition 4** The operator

$$\frac{D}{Dt} \overset{\text{def}}{=} \frac{\partial}{\partial t} + (u \cdot \nabla)$$

is called the total derivative or the material derivative. It accounts for the derivative of a quantity along the trajectory of a particle with velocity $u$.

For an arbitrary function $f(x, t)$, the chain rule allows to write

$$\dot{f}(x, t) = \frac{d}{dt} f(x, t) \overset{\text{def}}{=} \frac{\partial f}{\partial t}(x, t) + (u(x, t) \cdot \nabla) f(x, t) = \frac{Df}{Dt}(x, t).$$

With this notation, we have $\gamma(x, t) = \dot{u}(x, t)$. 

Notion of mass

- For each time $t > 0$, we assume that there exists a positive measure $\mu_t$ carried by $\Omega_t$ called the **mass distribution**, which is regular with respect to the Lebesgue measure $\text{d}x$.

- Hence, there exists a regular function $\rho = \rho(x, t)$ (sufficiently smooth) such that:

  \[
  d\mu_t(x) = \rho(x, t) \text{d}x,
  \]

  where $\rho$ is called the **mass density** of the system at point $x$ at time $t$.

- For any subset $W_t \subset \Omega_t$, the mass of fluid contained in $W_t$ at time $t$ is defined by:

  \[
  m(W_t) = \int_{W_t} d\mu_t(x) = \int_{W_t} \rho(x, t) \text{d}x.
  \]

  Notice that the assumption of existence for $\rho$ is a continuum assumption that no longer holds at molecular level.

*Non mathematicians can simply consider $d\mu_t(x)$ as a notation for the measure $\rho(x, t)\text{d}x$. 
Forces

The motion of a material system is entailed by the action of external forces of two types, represented by a regular vector measure $d\varphi_t(x)$ carried by a volume or a surface.

1. **Contact (or stress) forces:**
   expressed as forces across the surface $\partial W_t$ of any subsystem $W_t \subset \Omega_t$. Their description relies on Cauchy’s theory of stress tensor.

**Definition 5** In a $d$-dimensional evolving continuum $\Omega_t$, **Cauchy’s stress tensor** $\sigma(x, t)$ is a $d$-dimensional tensor, defined by the property that, for any subsystem $W_t \subset \Omega_t$ with smooth enough boundary $\partial W_t$, the surface density of forces applied on $\partial W_t$ by the rest of the continuum reads:

$$s(x, t) = \sigma(x, t) \cdot n(x, t),$$

The fact that there is no momentum of the stress efforts inside a fluid in static equilibrium has the following fundamental mathematical translation.

**Theorem 5** **Cauchy’s stress tensor** $\sigma$ is symmetric.
Forces (2)

The total force exerted on the fluid inside $W_t$ by means of stress on its boundary is

$$s_{\partial W_t} = \int_{\partial W_t} \sigma(x, t) \cdot n(x, t) \, ds(x).$$

where $ds$ is a regular surface measure.

**Definition 6** A fluid is said to be an ideal fluid when there exists a function $p(x, t)$, called the pressure, such that the stress tensor $\sigma$ is of the form:

$$\sigma(x, t) \cdot n = -p(x, t) \, n,$$

for any unit vector $n$.

This means that there is no tangential component in the contact forces. This definition also implies that there cannot be a rotation initiating in such fluid.
**Definition 7** A viscous Newtonian fluid is defined by the form of its stress tensor $\sigma$:

$$\sigma = -p\operatorname{Id} + \mathcal{L}(D(u)), \quad (6)$$

where $p$ is the pressure of the fluid, $D(u)$ is the deformation tensor and $\mathcal{L}$ is a linear mapping.

Further hypothesis (isotropic medium, invariance under a change of observer), lead to the following form of the stress tensor

$$\sigma = -p\operatorname{Id} + \lambda \text{div} u \operatorname{Id} + 2\mu D(u). \quad (6)$$

where $\lambda$, the volume viscosity, and $\mu > 0$, the dynamic viscosity, are the viscosity coefficients of the fluid (also called the Lamé coefficients).

Viscosity measures the resistance of a fluid under a shear stress deformation.
Forces (4)

2. **External (or body) forces:**

- exert a force per unit volume on the fluid; e.g. gravity
- described by a density

\[
d\varphi_t(x) = f(x, t) \, dx , \quad \text{for } x \in \Omega_t .
\]

For instance, in the classical case of *gravity*, we have

\[
f(x, t) = -\rho g e_3 ,
\]

where \( g \in \mathbb{R}^3 \) is the gravitational acceleration and \( e_3 \) is the vertical unit vector pointing upward.
Energy and mechanical work

Two kinds of mechanical works (energy exchange with exterior) can be distinguished, due to external or interior forces.

1. **External work:**

   **Definition 8** Suppose a fluid in motion, with velocity field $u(x, t)$, experiences an external volume force $f(x, t)$; the instantaneous mechanical work at time $t$ done by $f$ on a portion $W_t$ of $\Omega_t$ is defined as:

   $$\int_{W_t} f(x, t) \cdot u(x, t) \, dx.$$ 

   Similarly, if external loads $g(x, t)$ are applied on the boundary of $\Omega_t$, the mechanical work done by $g$ reads:

   $$\int_{\Omega_t} g(x, t) \cdot u(x, t) \, ds.$$ 

   When $f$ and $u$ have the same orientation, then $f \cdot u \leq 0$, and the mechanical work done by $f$ is positive, in the sense that the action of the force "helps" the motion.
2. Internal work:

**Definition 9** The instantaneous power received by a portion of fluid in $W_t \subset \Omega_t$ at time $t$ reads:

$$-\int_{W_t} \sigma : D(u) \, dx$$

To get an intuition, assume that the considered fluid is incompressible and viscous. As seen before, one possibility for modelling the stress-strain relation reads:

$$\sigma = -p \text{Id} + 2\mu D(u),$$

where $p$ stands for the scalar pressure field. Then, the instantaneous power received by a portion $W_t$ of fluid is:

$$-\int_{W_t} \sigma : D(u) \, dx = -\int_{W_t} p \, \text{div}(u) + 2\mu |D(u)|^2 \, dx = -2\int_{W_t} \mu |D(u)|^2 \, dx.$$ 

Hence viscous forces only entail energy dissipation.
Conservation principles

Establishing the partial differential equations that describe the evolution of $\Omega_t$ according to Newtonian mechanics relies on three conservation laws:

1. mass conservation: mass is neither created nor destroyed;

2. balance of momentum: according to Newton’s second law, the rate of change of the linear momentum equals the total applied force on the considered system;

3. conservation of energy: energy is neither created nor destroyed.

The derivation of the three forthcoming equations of conservation follows a common general sketch: a small portion of fluid, evolving in time, is chosen, and the quantities attached to it (mass, momentum, and energy) should satisfy the aforementioned conservation principles.

Giving a precise mathematical meaning to this idea involves differentiation of integrals on moving domains.
Differentiation of a volume integral

It is often required to compute the time derivative of integrals of the form:

\[ g(t) = \int_{\Omega_t} f(x, t) \, dx, \]

where \( f = f(x, t) \) is a given scalar function and \( \Omega_t \) is a bounded domain in \( \mathbb{R}^d \), evolving with respect to a velocity field \( u(x, t) \).

**Theorem 6 (Transport theorem (Liouville’s theorem))** Assume \( f = f(x, t) \) is a function of class \( C^1 \) for \( x \in \Omega_t \) and \( t \in I \), and that \( u \) is of class \( C^1 \) with respect to \( x \) and \( t \). Then,

\[
\frac{dg}{dt}(t) = \int_{\Omega_t} \frac{\partial f}{\partial t}(x, t) \, dx + \int_{\Omega_t} \text{div}(f \cdot u)(x, t) \, dx, \tag{7}
\]

\[
= \int_{\Omega_t} \frac{\partial f}{\partial t}(x, t) \, dx + \int_{\partial \Omega_t} f(x, t) u(x, t) \cdot n(x) \, ds(x), \tag{8}
\]

where \( \partial \Omega_t \) denotes the boundary of \( \Omega_t \), \( n \) is the unit outward normal vector on \( \partial \Omega_t \) and \( ds \) is the surface measure on \( \partial \Omega_t \).
Conservation of mass

The rate of change of mass in a subset $W_t$ is

$$\frac{d}{dt} m(W_t) = \frac{d}{dt} \int_{W_t} \rho(x, t) \, dx.$$ 

$= 0 \ (\text{conservation})$

Using the Transport theorem 6 with $f = \rho$ yields, for all $W_t$ in $\Omega_t$:

$$\int_{W_t} \frac{\partial \rho}{\partial t} + \text{div}(\rho u) \, dx = 0,$$

We obtain the so-called continuity equation:

$$\frac{\partial \rho}{\partial t} + \text{div}(\rho u) = 0 \ \text{in} \ \Omega_t, \quad (9)$$

In case of an incompressible fluid, $\rho$ is constant and equation (9) becomes

$$\text{div} \ u = 0 \ \text{in} \ \Omega_t.$$
Conservation of mass (2)

Based on this continuity equation, the following convenient form of theorem 6 is suited to deal with mass quantities:

**Corollary 1** Let \( f = f(x, t) \) is a function of class \( C^1 \) for \( x \in \Omega_t \) and \( t \in I \), and assume that \( u \) is of class \( C^1 \) with respect to \( x \) and \( t \). Define the integral quantity \( g(t) \) as:

\[
g(t) = \int_{\Omega_t} \rho(x, t) f(x, t) \, dx.
\]

Then,

\[
\frac{dg}{dt}(t) = \int_{\Omega_t} \rho(x, t) \frac{Df}{Dt}(x, t) \, dx.
\]  

(10)

**Proof:**

\[
\frac{dg}{dt}(t) = \int_{\Omega_t} \frac{\partial (\rho f)}{\partial t}(x, t) + \text{div}((\rho f)u)(x, t) \, dx,
\]  

(11)

\[
= \int_{\Omega_t} \rho \left( \frac{\partial f}{\partial t} + u \cdot \nabla f \right)(x, t) + \left( \frac{\partial \rho}{\partial t} + \text{div}(\rho u) \right)(x, t) \, dx,
\]  

(12)

+ definition 4 of the material derivative \( \frac{Df}{dt} \).
Balance of momentum

The total force $F$ acting on $W_t \subset \Omega_t$ filled by the fluid reads:

$$F = \int_{W_t} f(x, t) \, dx + \int_{\partial W_t} \sigma(x, t).n(x, t)) \, ds(x) .$$  \hfill (13)

Balance of momentum:

- related to Newton’s second law $F = m\gamma$;
- change of momentum related to the force acting on $\Omega_t$.

Given $W_t = \chi(W, t, t_0)$, volume of fluid moving, the linear momentum in $W_t$ writes:

$$M_L(W_t) = \int_{W_t} \rho(x, t) \, u(x, t) \, dx ,$$  \hfill (14)

and the balance of momentum reads, thanks to (13):

$$\frac{d}{dt} M_L(W_t) = \int_{W_t} f(x, t) \, dx + \int_{\partial W_t} \sigma(x, t).n(x, t)) \, ds(x) ,$$

where $n = n(x, t)$ is the unit outward normal to $\partial W_t$. 
Balance of momentum (2)

The left-hand side of previous equation rewrites (cf. corollary 1):

\[
\frac{d}{dt} M_L(W_t) = \int_{W_t} \rho(x, t) \frac{Du}{Dt}(x, t) \, dx = \int_{W_t} \rho(x, t) \dot{u}(x, t) \, dx.
\]

Now, using the divergence theorem 4, one has:

\[
\int_{\partial W_t} \sigma(x, t) n(x, t) \, ds(x) = \int_{W_t} \text{div} \, \sigma(x, t) \, dx.
\]

Since \( W_t \) is an arbitrary subset, we can gather all previous results to write the linear momentum formula:

\[
\rho \dot{u} = \text{div} \, \sigma + f. \tag{15}
\]

Interestingly, the linear momentum law can be used in the context of energy considerations.
Balance of energy

Let $W_t \subset \Omega_t$ any evolving portion of fluid, $u$ the velocity field in the frame of reference.

**Definition 10** The *kinetic energy* of $W_t$ at time $t$ is given by:

$$E_c = \frac{1}{2} \int_{\Omega_t} |u(x, t)|^2 \, d\mu_t(x) = \frac{1}{2} \int_{\Omega_t} \rho(x, t)|u(x, t)|^2 \, dx,$$

The rate of change of kinetic energy of a moving subset $W_t$ of fluid is calculated as follows

$$\frac{d}{dt}E_c = \frac{d}{dt} \left( \frac{1}{2} \int_{W_t} \rho(x, t)|u(x, t)|^2 \, dx \right)$$

$$= \frac{1}{2} \int_{W_t} \rho(x, t) \frac{D|u(x, t)|^2}{Dt} \, dx \quad \text{(corollary 1)}$$

$$= \frac{1}{2} \int_{W_t} \rho(x, t) \left( u \cdot \left( \frac{\partial u}{\partial t} + (u \cdot \nabla)u \right) \right) (x, t) \, dx$$

in other terms,

$$\frac{d}{dt}E_c = \int_{W_t} \rho u \cdot \dot{u} \, dx. \quad (16)$$
**THEOREM 7** The instantaneous rate of change in kinetic energy $\frac{d}{dt}E_c$ in any portion of fluid $W_t \subset \Omega_t$ equals the sum of:

1. the power of internal forces within $W_t$:

$$- \int_{W_t} \sigma : D(u) \, dx,$$

2. the mechanical work done by surface loads $\sigma \cdot n$ (owing to the continuity of the stress tensor in $\Omega_t$) applied on $\partial W_t$:

$$\int_{\partial W_t} (\sigma \cdot n) \cdot u \, dx,$$

3. the mechanical work done by external forces on $W_t$:

$$\int_{W_t} f \cdot u \, dx.$$

Hence, we write:

$$\int_{W_t} \rho u \cdot \dot{u} \, dx = - \int_{W_t} \sigma : D(u) \, dx + \int_{\partial W_t} (\sigma \cdot n) \cdot u \, dx + \int_{W_t} f \cdot u \, dx. \quad (17)$$
Section 1.3

Flows models
Summary of the equations

1. **Continuity equation**
   Conservation of mass
   \[
   \frac{\partial \rho}{\partial t} + \text{div}(\rho u) = 0 \tag{18}
   \]

2. **Momentum equation**
   Newton’s second law
   \[
   \frac{\partial (\rho u)}{\partial t} + \text{div}(\rho u \times u) = \text{div} \sigma + f
   \]

3. **Energy equation**
   First law of thermodynamics
   \[
   \frac{\partial (\rho E_c)}{\partial t} + \text{div}(\rho E_c u) = -D(u) : \sigma + f \cdot u
   \]
The Navier-Stokes equations

- **Physical considerations:**
  we consider the motion of a *Newtonian incompressible viscous fluid*, with the following assumptions on the Cauchy stress tensor $\sigma$:
  
  1. $\sigma$ depends linearly on the velocity gradients $\nabla u$
     - normal stress: stretching
     - shear stress: deformation
  2. $\sigma$ is invariant under translations and rigid rotations (Galilean invariance);
  3. $\sigma$ is symmetric (consequence of the balance of angular momentum).

We introduced $\sigma$ as (viscous fluid $\mu > 0$):

$$\sigma = -p \text{Id} + \lambda \text{div} u \text{Id} + 2\mu \mathcal{D}(u),$$

which for an incompressible fluid ($\text{div} u = 0$) can be simplified:

$$\sigma = -p \text{Id} + 2\mu \mathcal{D}(u).$$
The Navier-Stokes equations (2)

Substituting \( \sigma \) in the expression \( \rho \dot{u} = \text{div} \sigma + f \) yields the Navier-Stokes equations for a viscous incompressible fluid filling a domain \( \Omega_t \), for any \( t \in [0, T] \):

\[
\begin{cases}
\rho \left( \frac{\partial u}{\partial t} + (u \cdot \nabla)u \right) = -\nabla p + 2 \text{div}(\mu \nabla u) + f \\
\text{div} u = 0
\end{cases}
\]  

(19)

where \( \rho = \rho(x, t) \) denotes the density, \( u = u(x, t) \) the velocity of the fluid particle at position \( x \) at time \( t \), \( p = p(x, t) \) the pressure, and \( f = f(x, t) \) the volume forces.

Given \( \mu > 0 \) and \( \text{div} u = 0 \), then \( \text{div} \nabla u^t = \nabla \text{div} u \) and thus \( \text{div}(\mu \nabla u) = \mu \Delta u \), the "classical" form of Navier-Stokes equations is then:

\[
\begin{cases}
\text{Given } f; \text{ find } (u, p) \text{ such that } \\
\rho \left( \frac{\partial u}{\partial t} + (u \cdot \nabla)u \right) = -\nabla p + \mu \Delta u + f \\
\text{div} u = 0
\end{cases}
\]  

(20)
The Navier-Stokes equations (3)

\[ \rho \left( \frac{\partial u}{\partial t} + (u \cdot \nabla)u \right) = -\nabla p + \mu \Delta u + f \]

1. The inertial or non-linear term \((u \cdot \nabla)u\) is related to convective acceleration: it contributes to the transfer of kinetic energy.

2. The term \(-\nabla p\) is the pressure gradient:
   A fluid reacts to attempt to change its volume according to a pressure. This term arises from the isotropic part of the Cauchy-stress tensor (6) \(p = 1/3 \text{tr}(\sigma)\). The pressure is only involved through its gradient in Navier-Stokes equations: accelerating the fluid from high to low pressure.
   Pressure can be interpreted as a Lagrange multiplier corresponding to \(\text{div}\ u = 0\).
The Navier-Stokes equations (4)

\[ \rho \left( \frac{\partial u}{\partial t} + (u \cdot \nabla)u \right) = -\nabla p + \mu \Delta u + f \]

3. The term \( \mu \Delta u \) is the **dissipative viscous term**;
   \( \mu \) is the **dynamic viscosity** of the fluid.
   Viscosity operates in a diffusion of a momentum, much alike the diffusion of heat in the heat transfer equation (a Laplacian in space). Diffusion of fluid momentum is the result of friction between fluid particles moving at uneven speed that tend gradually to become the same.

4. The vector field \( f \) represents the **body forces** (forces per unit volume) such as gravity.
Initial and boundary conditions

1. Initial condition:
   - initial velocity field must be divergence-free in $\Omega_t$ and at $\partial \Omega_t$:
     $$ u(x, 0) = u_0(x), \quad \text{div} \, u_0 = 0. $$

2. Boundary conditions for viscous fluids
   - Need to consider molecular interactions between fluid and surface of a solid body;
   - **No-slip** condition: normal and tangential components of the velocity vanish:
     $$ u(x, t) = 0, \quad \forall x \in \partial \Omega_t, \quad \forall t \in I; $$
   - General situation: $\partial \Omega_t$ moving at speed $v(x, t)$, then solid and fluid share the same velocity at the boundary, i.e.
     $$ u(x, t) = v(x, t), \quad \forall x \in \partial \Omega_t, \quad \forall t \in I. $$
The Reynolds number

- Write the equations in a non-dimensional form, independent of the system of units.

- Define the dimensionless variables as follows:
  \[ u' = \frac{u}{U}, \quad x' = \frac{x}{L}, \quad t' = \frac{tU}{L}, \quad p' = \frac{p}{\rho U^2}. \]

- Substitute the new variables in the Navier-Stokes equations:
  \[
  \begin{cases}
    \frac{\partial u'}{\partial t'} + (u' \cdot \nabla)u' = -\nabla p' + \frac{1}{Re} \Delta u' + f', \\
    \text{div } u' = 0
  \end{cases}
  \] (21)

- New dimensionless parameter: the Reynolds number of the flow:
  \[ Re = \frac{UL}{\nu} = \frac{\rho U^2}{\mu L} = \frac{\text{acceleration forces}}{\text{viscosity forces}}. \]

  provides a measure of the viscosity of the flow.
Chapter II

The Stokes Model

2.4. Resolution of Stokes Examples

Figure 2.15: Example of a Couette flow using the pPCG method: convergence of the velocity field for $\tau = 10^n$.

Comprehensible since we have only taken into account the viscosity ratio, and since the condition number of the stiffness matrix is related to the minimal size of the triangulation $T_h$.

Figure 2.16: Couette flow: streamlines of the velocity field in the vicinity of the interface.
Section 2.1
Mathematical and numerical analysis
Viscous incompressible flows

Take conservation laws and make simplifications: $\rho$ and $\mu$ constant

- continuity equation: $\frac{\partial \rho}{\partial t} + \text{div}(\rho u) = 0 \rightarrow \text{div} u = 0$

- inertial terms: $\frac{\partial (\rho u)}{\partial t} + \text{div}(\rho u \times u) = \rho \left( \frac{\partial u}{\partial t} + u \cdot \nabla u \right) = \rho \frac{du}{dt}$

- stress tensor: $\text{div} \sigma = \mu \text{div}(\nabla u + \nabla u^t) = \mu (\text{div} \nabla u + \nabla \nabla \cdot u) = \mu \Delta u$

- Stokes flow, $Re \rightarrow 0$, then $\frac{du}{dt} \approx \frac{\partial u}{\partial t} \approx 0$

  - momentum equation:
    
    $$\rho \frac{\partial u}{\partial t} = -\nabla p + \mu \Delta u$$

The Stokes model

Model: linear Stokes problem

- posed in a bounded regular domain $\Omega$, endowed with
- homogeneous Dirichlet boundary conditions

\[
\begin{aligned}
\begin{cases}
\text{Given } f; \text{ find } (u, p) \text{ such that:} \\
- \mu \Delta u + \nabla p = f, & \text{in } \Omega \\
\text{div } u = 0, & \text{in } \Omega \\
u = 0, & \text{on } \partial \Omega
\end{cases}
\end{aligned}
\]

(22)

(and we consider hereafter the viscosity parameter $\mu$ equal to 1, for simplicity).

- Remark: not physically realistic because of homogeneous Dirichlet boundary conditions.
Classical formulation

1. **Classical solution:**
   
a pair \((u, p) \in (C^2(\Omega) \cap C^0(\bar{\Omega})) \times (C^1(\Omega))\) is a classical solution of the Stokes problem if it satisfies eq. (22).

2. **Non-uniqueness:**
   
   - if \((u, p)\) is solution then \((u, p + c)\) is a solution, too.
   - to avoid the undeterminacy, 2 possibilities:
     
     (a) either impose the value of \(p\) at a given point \(x_0\) in the domain, \(p(x_0) = p_0\)
     
     but, not consistent with the functional setting of \(p \in L^2(\Omega)\).
     
     (b) or require the pressure to have null average:

     Here, we impose the zero mean condition of \(p\) over the domain \(\Omega\):

     \[
     \int_{\Omega} p \, dx = 0
     \]
Variational formulation

- **Concept:**
  - equation is not required to hold absolutely but has weak solutions with respect to test functions;
  - problem reformulated in the sense of distributions.

- **Approach:**
  by multiplying the first and second equations by a test function \( v \in V \) and \( q \in Q \) and by integrating over \( \Omega \), we obtain, thanks to the divergence theorem, the set of equations:

\[
\begin{align*}
\int_{\Omega} \nabla u : \nabla v \, dx - \int_{\Omega} \text{div} \, v \, p \, dx &= \int_{\Omega} f \, v \, dx + \int_{\partial\Omega} \left[ \frac{\partial u}{\partial n} \right] \, v \, ds \\
\int_{\Omega} \text{div} \, u \, q \, dx &= 0.
\end{align*}
\]

where \( V \) and \( Q \) are suitable functional spaces.
Variational formulation

- **Functional spaces**: Hilbert spaces
  - velocity space $V$ (to account for homogeneous BC):
    \[ V = H^1_0(\Omega) \] (23)
  - pressure space $Q$:
    \[ Q = L^2_0(\Omega) = \left\{ q \in L^2(\Omega) ; \int_{\Omega} q \, dx = 0 \right\}. \] (24)

- **Abstract formulation**: we define the bilinear forms $a : V \times V \to \mathbb{R}$ and $b : V \times Q \to \mathbb{R}$
  \[ a(u, v) = \int_{\Omega} \nabla u : \nabla v \, dx, \quad \text{and} \quad b(u, q) = -\int_{\Omega} \text{div} \ u \, q \, dx. \] (25)

and we introduce the following abstract **weak formulation**:

\[ \left\{ \begin{aligned}
\text{Given } f; \text{ find } (u,p) \in V \times Q, \text{ such that for all } (v,q) \in V \times Q \\
a(u, v) + b(v, p) = (f, v), \\
b(u, q) = 0,
\end{aligned} \right. \] (26)

where $(\cdot, \cdot)$ is the duality pairing between functions $f \in H^{-1}(\Omega)$ and $v \in H^1_0(\Omega)$. 
Variational formulation

- **Non-homogeneous Dirichlet boundary conditions**: \( u = g \) on \( \partial \Omega \):
  
  yields the following weak formulation:

  \[
  \begin{aligned}
  \text{Given } f ; \text{ find } (w, p) \in V \times Q, \text{ such that for all } (v, q) \in V \times Q
  
  &a(w, v) + b(v, p) = (f, v) - a(u_g, v) \\
  &- b(w, q) = b(u_g, q),
  \end{aligned}
  \]  

  (27)

  where \( V \) and \( Q \) are the spaces introduced above and \( u_g \in H^1(\Omega) \) is a lifting of the boundary condition \( g \):

  \[
  w = u - u_g.
  \]

- Problems (26)-(27) are called **saddle-point problems**.

- Pressure \( p \) is the **Lagrange multiplier** associated with the incompressibility constraint.
Linear operators

- **Linear operator** and **bilinear form** are related by a one-to-one mapping:
  
  - given $V = H_0^1(\Omega)$ and $Q = L_0^2(\Omega)$,
  
  - for the bilinear form $a : V \times V \rightarrow \mathbb{R}$, consider operator $A : V \rightarrow V'$:
    \[
    \langle Au, v \rangle = a(u, v) , \quad \forall (u, v) \in V \times V .
    \]  

  - for the bilinear form $b : V \times Q \rightarrow \mathbb{R}$, consider operator $B : V \rightarrow Q'$:
    \[
    \langle Bu, q \rangle = b(u, q) , \quad \forall (u, v) \in V \times Q .
    \]

    and the dual operator $B' : Q \rightarrow V'$:
    \[
    \langle B' q, u \rangle = \langle Bu, q \rangle = b(u, q) , \quad \forall (u, v) \in V \times Q .
    \]

  - Many properties of linear operators can be inherited from corresponding properties of bilinear forms: **continuity**, **positivity** and **ellipticity**, or **symmetry**.
Operator equations for Stokes

• Using these operators, the variational problem (26) is equivalent to the problem:

\[
\begin{aligned}
\text{Given } f; \text{ find } (u, p) \in V \times Q, \text{ such that for all } (v, q) \in V \times Q \\
Au + B'p &= f, \quad \text{in } V' \\
Bu &= 0, \quad \text{in } Q'
\end{aligned}
\]  

(29)

• Consider the spaces: \(X = \ker B\) in \(V\); \(X = \{v \in V; b(v, q) = 0, \forall q \in Q\}\), the polar set of \(X\); \(X^0 = \{g \in V'; \langle g, v \rangle = 0, \forall v \in X\}\).

• **Lemma 1** The following statements are equivalent

1. there exists a constant \(\beta > 0\) such that

\[
\inf_{q \in Q} \sup_{v \in V} \frac{b(v, q)}{\|v\|_V \|q\|_X} \geq \beta, \tag{30}
\]

2. the operator \(B'\) is an isomorphism from \(Q\) onto \(X^0\), and \(\|B'q\|_{V'} \geq \beta \|v\|_Q\),

3. the operator \(B\) is an isomorphism from \(X^\perp\) to \(Q'\) and \(\|Bv\|_{Q'} \geq \beta \|v\|_V\).
The Stokes model

Operator equations for Stokes

- **Existence** and **uniqueness** of a solution to the Stokes problem can be established:

**Theorem 8 (Brezzi splitting theorem)** The solution operator

\[ L : V \times Q \rightarrow V' \times Q' \]

\[ (u, p) \mapsto (f, 0) \]

is an **isomorphism** if and only if the following conditions are satisfied:

1. the bilinear form \( a(\cdot, \cdot) \) is **coercive** on \( V \), i.e., there exists \( \alpha > 0 \) such that
   \[ a(v, v) \geq \alpha \| v \|_V^2, \quad \forall v \in V. \]

2. the bilinear form \( b(\cdot, \cdot) \) satisfies the **inf-sup condition** (30).

Then, the Stokes problem (26) is well-posed.

- **The coercivity** of the bilinear form \( a(\cdot, \cdot) \) implies that the functional associated to Stokes’ problem is strictly convex on the feasible set.
A saddle-point approach

Stokes problem can be formulated as a minimization problem:

- consider the energy functional $J : V \to \mathbb{R}$ defined by
  \[
  J(v) = \frac{1}{2} a(v, v) - (f, v) = \frac{1}{2} a(v, v) - f(v). \tag{31}
  \]

- the solution $(u, p)$ is a minimizer of the objective function $J(\cdot)$ subject to the constraint $b(v, q) = 0$.

- we introduce an associated quadratic functional $\mathcal{L} : V \times Q \to \mathbb{R}$ defined by:
  \[
  \mathcal{L}(v, q) = J(v) + b(v, q) = \frac{1}{2} a(v, v) + b(v, q) - f(v). \tag{32}
  \]
  The functional $\mathcal{L}(\cdot, \cdot)$ is the Lagrangian functional associated with the Stokes problem.

- find a saddle-point $(u, p)$ of the Lagrangian functional over $V \times Q$, i.e. such that
  \[
  \mathcal{L}(u, q) \leq \mathcal{L}(u, p) \leq \mathcal{L}(v, p), \quad \forall (v, q) \in V \times Q. \tag{33}
  \]
A saddle-point approach

- Existence of a solution:

**Lemma 2** The pair \((u, p) \in V \times Q\) solves the Stokes problem (26) if and only if \((u, p)\) is a saddle-point of the Lagrangian functional \(L(v, q)\)

\[
L(v, q) = J(v) + b(v, q) = \frac{1}{2}a(v, v) + b(v, q) - f(v),
\]

that is, if and only if

\[
L(u, p) = \inf_{v \in V} \sup_{q \in Q} L(v, q).
\]  

- The pressure \(p \in Q\) plays the role of a Lagrange multiplier associated to the divergence-free constraint \(\text{div } u = 0\) (here represented by the constraint \(b(v, q) = 0\)).
Generalized Stokes problem

Model: generalized Stokes problem

- posed in a bounded domain $\Omega \subset \mathbb{R}^d$
- endowed with specific boundary conditions

\[
\begin{cases}
\text{Given } f; \text{ find } (u, p) \text{ such that:} \\
\alpha u - \mu \Delta u + \nabla p = f, & \text{in } \Omega \\
\text{div } u = 0, & \text{in } \Omega \\
u u = g_0 & \text{on } \Gamma_D, \quad \nu \frac{\partial u}{\partial n} = g_1 & \text{on } \Gamma_N
\end{cases}
\]

(35)

where $\alpha, \nu$ are given positive constants; $\Gamma_0$ and $\Gamma_1$ are two subsets of $\partial \Omega$ such that $\Gamma_D \cap \Gamma_N = \emptyset$, $\Gamma_D \cup \Gamma_N = \partial \Omega$ and $g_0$ and $g_1$ are two given functions defined on $\partial \Omega$;

- describe the motion of an incompressible viscous flow at low Reynolds number;

- derive from an implicit time discretization of the Navier-Stokes equations, in which case $\alpha \approx 1/\Delta t$.
Generalized Stokes problem

- Consider the functional spaces

\[ V = \{ v \in H^1(\Omega), \ v = 0 \text{ on } \Gamma_D \}, \text{ and } V_{g_0} = \{ v \in H^1(\Omega), \ v = g_0 \text{ on } \Gamma_D \}. \]

- Suppose that \( f \in L^2(\Omega), \tilde{g}_0 = g_0|_{\Gamma_D} \in H^1(\Omega) \) and \( g_1 \in L^2(\Gamma_N). \)

- If problem (35) admits a solution \((u, p) \in H^1(\Omega) \times L^2(\Omega)\), then this solution is such that \( u \in V_{g_0}, p \in L^2(\Omega) \) and, for all \( v \in V \):

\[
\begin{align*}
\int_\Omega (\alpha uv + \nu \nabla u \cdot \nabla v) \, dx - \int_\Omega p \, \text{div} \, v &= \int_\Omega f \, v \, dx + \int_{\Gamma_N} g_1 \, v \, ds,
\end{align*}
\]

\( \text{Theorem 9 (Uniqueness and Existence)} \) \text{ Suppose the previous hypothesis on } \nu, \alpha, f, g_0 \text{ and } g_1 \text{ hold. Then, there exists a unique solution } (u, v) \in V_{g_0} \times Q, \text{ if } p \in Q = L^2_0(\Omega).\)
Section 2.2

Finite element approximation
Discrete variational problem

- Weak formulation:

\[
\begin{cases}
\text{find } (u, p) \in V \times Q, \text{ such that for all } (v, q) \in V \times Q \\
a(u, v) + b(v, p) = (f, v), \\
b(u, q) = 0,
\end{cases}
\]

- Discrete weak formulation:

\[
\begin{cases}
\text{find } (u_h, p_h) \in V_h \times Q_h, \text{ such that for all } (v_h, q_h) \in V_h \times Q_h \\
a(u_h, v_h) + b(v_h, p_h) = (f, v_h), \\
b(u_h, q_h) = 0,
\end{cases}
\] (37)

- Existence and uniqueness of a solution: need to choose suitable \( V_h \) and \( Q_h \), such that the discrete inf-sup condition holds:

\[
\exists \beta > 0 ; \quad \inf_{q_h \in Q_h} \sup_{v_h \in V_h, v_h \neq 0} \frac{b(v_h, q_h)}{\|v_h\| \|q_h\|} \geq \beta .
\] (38)
Discrete variational problem

- The compatibility condition (38) is essential to ensure the uniqueness of the solution;
- Practical criteria for finding stable pairs of finite element spaces have been proposed by various authors; finite element spaces $V_h$ and $Q_h$ cannot be chosen independently;
- In some cases, the inf-sup condition may not hold: there exists a function $\tilde{q}_h \in Q_h$, such that $b(v_h, \tilde{q}_h) = 0$, for all $v_h \in V_h$ and we have
  \[
  a(u_h, v_h) + b(v_h, p_h + \tilde{q}_h) = a(u_h, v_h) + b(v_h, p_h) + b(v_h, \tilde{q}_h) = (f, v_h).
  \]
  if $(u_h, p_h)$ is a solution, so is $(u_h, p_h + \tilde{q}_h)$. Such functions are called spurious pressure modes and they usually yield numerical instabilities. The subspaces of these functions are called unstable or incompatible.
- The space $S_h$ of spurious pressure modes is defined as
  \[
  S_h = \ker B^t = \left\{ a_h \in Q_h, \int_{\Omega} q_h \text{div} v_h \, dx = 0, \forall v_h \in V_h \right\},
  \]
  and a necessary condition for the inf-sup to hold is to have
  \[
  S_h = \{0\}.
  \]
Algebraic formulation

The discrete finite element problem is equivalent to a system of linear algebraic equations.

- Let \( \{ \varphi_j \in V_h \} \) and \( \{ \phi_k \in Q_h \} \) be the bases of the spaces \( V_h \) and \( Q_h \). Then, we have

\[
\begin{align*}
  u_h(x) &= \sum_{j=1}^{N} u_j \varphi_j(x), & p_h(x) &= \sum_{k=1}^{M} p_k \phi_k(x), & N &= \dim V_h, & M &= \dim Q_h.
\end{align*}
\]

- Test functions chosen as functions \( \varphi_j \) and \( \phi_k \) in (37), leads to the algebraic system:

\[
\begin{cases}
  AU + B^t P = F \\
  BU = 0
\end{cases}
\]

or, in matrix form

\[
\begin{pmatrix}
  A & B^t \\
  B & 0
\end{pmatrix}
\begin{pmatrix}
  U \\
  P
\end{pmatrix}
= 
\begin{pmatrix}
  F \\
  0
\end{pmatrix}
\]

where \( U = [u_j] \) and \( P = [p_k] \) and \( F = [f_l] = [(f, \varphi_l)] = [\int_{\Omega} f \varphi_l] \) are vectors; and

\[
\begin{align*}
  A & \in \mathcal{M}_{N,N}(\mathbb{R}) = [a_{ij}] = [a(\varphi_j, \varphi_i)] = \left[ \int_{\Omega} \nabla \varphi_j \cdot \nabla \varphi_i \, dx \right] \\
  B & \in \mathcal{M}_{M,N}(\mathbb{R}) = [b_{km}] = [b(\varphi_m, \phi_k)] = \left[ \int_{\Omega} \phi_k \text{div} \varphi_m \, dx \right]
\end{align*}
\]
**Algebraic formulation**

Properties of the linear system:

- The linear system is **indefinite** since the matrix $S$ is block symmetric and indefinite (has positive and negative real eigenvalues).

- The matrix $A$ is **non-singular** since it is related to the **coercive** bilinear form $a(\cdot, \cdot)$, hence $A^{-1}$ is defined.

- We can derive the **dual problem**, by formally eliminating the velocity variable:

  $$(B \, A^{-1} \, B^t) \, P = (B \, A^{-1}) \, F$$  \hspace{1cm} (40)

  Matrix $BA^{-1}B^t$ (positive definite) is the **Schur complement** with respect to $P$.

- Then we compute the velocity as

  $$AU = F - B^t \, P.$$  \hspace{1cm} (41)

  This operation corresponds to Gaussian elimination on the initial system.
**Algebraic formulation**

Existence of a solution:

- System (40) has a unique solution $P$ if $R = B A^{-1} B^t$ is non-singular ($B^t$ non-singular).

- The kernel of matrix $B^t$ coincide with the space $S_h$ of spurious pressure modes:
  \[
  \ker B^t = \{0\}. \tag{42}
  \]

- The existence and uniqueness of $P$ allows to conclude that under condition (42), there exists a unique vector $U$ solution of system (41). Condition equivalent to the inf-sup condition.

**Proposition 2** The matrix $S$ is non-singular iif there is no zero eigenvalue.

- $A^{-1}$ is usually a dense matrix and can be expensive to compute.
  System (40) has to be resolved using iterative method (Uzawa, Schur complement).
Finite element spaces

Choice of suitable spaces

- Stokes equations are PDE of second order with respect to $u$ and first order with respect to $p$.

- The first obvious choice is to introduce a polynomial approximation space of degree $k \geq 1$ for the space $V_h$ and of degree $k - 1$ for the space $Q_h$.

- In this respect, the simplest choice of finite element spaces one can imagine is to use a standard piecewise linear $P_1$ approximation for velocity and a constant approximation $P_0$ for pressure defined on a triangulation $T_h$.

  Since the velocity is piecewise affine on the elements, its divergence is constant. Hence, by testing the divergence against constants, the divergence-free condition is easily strongly imposed on the discrete velocity field.

- Better choices exist.. maybe not so obvious.
Stable discretizations

1. $\mathbb{P}_2/\mathbb{P}_0$ and Crouzeix-Raviart mixed elements:

\[ V_h = \{v_h \in C^0(\bar{\Omega}); \quad v_h|_T \in \mathbb{P}_2, \ \forall T \in T_h, \ v_h|_{\partial \Omega} = 0\} \]

\[ Q_h = \{q_h \in L^2_0(\Omega); \quad q_h|_T \in \mathbb{P}_0, \ \forall T \in T_h\} , \]

- The discrete divergence-free condition writes on any triangle $T \in T_h$:

\[ \int_T \text{div} \ u_h \, dx = \int_{\partial T} u_h \cdot n \, ds = 0 , \]

and can be interpreted as a mass conservation law on every element.

- Extension: Crouzeix-Raviart mixed element;
  introduce normal component of the velocity at the mid-edge nodes. On each side of the triangle, the normal (resp. tangential) component of $u_h$ is quadratic (resp. affine).
Stable discretizations

2. $\mathbb{P}_2/\mathbb{P}_1$ Hood-Taylor element: stable

\[ V_h = \{ v_h \in C^0(\Omega) ; v_h|_T \in \mathbb{P}_2 , \forall T \in T_h , v_h|_{\partial\Omega} = 0 \} \]
\[ Q_h = \{ q_h \in L_0^2(\Omega) \cap C^0(\bar{\Omega}) ; q_h|_T \in \mathbb{P}_1 , \forall T \in T_h \} \]

- Suppose that every triangle in $T_h$ has no more than one edge on $\partial\Omega$. Then, there exists a constant $c > 0$ such that

\[
\sup_{v_h \in V_h} \frac{\int_{\Omega} q_h \text{div} v_h}{\| v_h \|_{H^1(\Omega)}} \geq c \left( \sum_{T \in T_h} h^2(T) |q_h|_{H^1(T)}^2 \right)^{1/2}.
\]
The Stokes model

\( P_2/P_1 \) Hood-Taylor element

- In element \( E \), the basis functions \( \varphi_i \) of \( V_h \) are defined by:

\[
\begin{align*}
\varphi_1(x, y) &= \left(1 - \frac{x}{h} - \frac{y}{2h}\right) \left(1 - \frac{x}{2h} - \frac{y}{2h}\right) \\
\varphi_2(x, y) &= \frac{x}{2h} \left(\frac{x}{h} - 1\right) \quad \varphi_3(x, y) = \frac{y}{2h} \left(\frac{y}{h} - 1\right) \\
\varphi_4(x, y) &= \frac{x}{h} \left(2 - \frac{x}{h} - \frac{y}{h}\right) \quad \varphi_6(x, y) = \frac{y}{h} \left(2 - \frac{x}{h} - \frac{y}{h}\right) \\
\varphi_5(x, y) &= \frac{xy}{h^2}
\end{align*}
\]

and are such that \( \varphi_i(x_j, y_j) = \delta_{ij} \)

- The basic functions \( \phi_i \) of \( Q_h \) are defined by:

\[
\begin{align*}
\phi_1(x, y) &= 1 - \frac{x}{2h} - \frac{y}{2h} \\
\phi_2(x, y) &= \frac{x}{2h} \quad \phi_3(x, y) = \frac{y}{2h}
\end{align*}
\]
Stable discretizations

3. $\mathbb{P}_1 b / \mathbb{P}_1$ element

- In each triangle $T$, a bubble function is defined at the barycenter

$$b \in H^1_0(T), \quad 0 \leq b \leq 1, \quad b(G) = 1,$$

$$V_h = \{v_h \in C^0(\bar{\Omega}) ; \; v_h|_T \in \mathcal{P}, \; \forall T \in T_h, \; v_h|_{\partial \Omega} = 0\}$$

$$Q_h = \{q_h \in L^2_0(\Omega) \cap C^0(\bar{\Omega}) ; \; q_h|_T \in \mathbb{P}_1, \; \forall T \in T_h\},$$

where, using the barycentric coordinates $\lambda_i$ of $T$,

$$\mathcal{P}(T) = \mathbb{P}_1(T) \oplus \text{span}(\lambda_1 \lambda_2 \lambda_3),$$

- **Lemma 3** If the triangulation $T_h$ is regular, there exists a constant $\beta$ such that the following estimate holds independently of $h$ (the discretization parameter):

$$\inf_{q_h \in Q_h} \sup_{v_h \in V_h, \; q_h \neq 0, \; v_h \neq 0} \frac{b(v_h, q_h)}{\|v_h\|_{H^1(\Omega)} \|q_h\|_{L^2(\Omega)}} \geq \beta.$$
Unstable discretizations

4. $\mathbb{P}_1/\mathbb{P}_0$ element: 

$$V_h = \{ v_h \in C^0(\overline{\Omega}); \ v_h|_T \in \mathbb{P}_1, \ \forall T \in T_h, \ v_h|_{\partial\Omega} = 0 \}$$
$$Q_h = \{ q_h \in L^2_0(\Omega); \ q_h|_T \in \mathbb{P}_0, \ \forall T \in T_h \}$$

- simplicity (obvious choice) but numerical locking
- this obvious choice does not satisfy the inf-sup condition.
Unstable discretizations

- Consider $\Omega = [0, 1] \times (0, 1]$ covered by a triangulation pictured below ($N \times N$ squares splitted into two triangles each): $\dim(V_h) = 2 (N - 1)^2 < \dim(Q_h) = 2 N^2$; the divergence matrix $B$ has more rows than columns and $(BU = 0) \Rightarrow (U = 0)$.

- More generally, Euler’s relations give the identity: $2 = 2N_v + N_e - N_t$ and we know that $\dim(V_h) = 2N_v$ and $\dim(Q_h) = N_t - 1$. The rank-nullity theorem writes

$$\dim(\text{im}(B_h^t)) + \dim(\ker(B_h^t)) = \dim(Q_h)$$

$$\Rightarrow \dim(\ker(B_h^t)) = \frac{N_t - 1 - 2N_v}{N_e - 3} \geq N_t - 1 - \dim(V_h)$$

where $B_h \in \mathcal{L}(V_h, Q_h); \langle B_h v_h, q_h \rangle = b(v_h, q_h)$ for all $v_h \in V_h$ and $q_h \in Q_h$. Hence, there is a least $N_e - 3$ spurious pressure modes.

The dimension of space $Q_h$ is too large: locking phenomenon.
Stabilization procedures

Elimination of spurious modes

- Equal degree interpolation for the velocity and the pressure do not usually satisfy the discrete inf-sup condition and are considered unstable.

- The coercivity property may not be inherited on the approximation subspaces and the conformity of these spaces is not sufficient to ensure the stability of the discrete problem.

- A stabilization technique can be used to remove the spurious modes in the finite dimensional problem.

These techniques are usually appreciated by engineers since they allow to use unstable yet simple finite elements to resolve the Stokes equations. Furthermore, stabilization is also convenient in a parallel (or multigrid) context.

- We have already seen a stabilization technique with the $P_1/P_1$ element, by adding bubbles to the velocity space leading to define the $P_1b/P_1$ mini-element.
Stabilization procedures

Least-squares formulation of the Stokes problem

- Stabilization methods involve an alteration of the discrete incompressibility condition in such a way that the discrete inf-sup condition is automatically satisfied for any choice of the discrete spaces $V_h \subset V$ and $Q_h \subset Q$.

- We consider only piecewise linear finite elements.

- Let introduce the space $W = H^1_0(\Omega) \times L^2_0(\Omega)$. The mixed formulation writes:

  \[
  \text{Find } (u, p) \in W; \text{ such that for all } (v, q) \in W, \quad A(u, p; v, q) = F(v, q),
  \]

  where we have set the functionals

  \[
  A(u, p; v, q) = a(u, v) + b(v, p) - b(u, q)
  = (\nabla u, \nabla v) - (\text{div } v, p) - (\text{div } u, q) = F(v, q)
  \]

  and $F(v, q) = (f, v)$.

  For $f \in L^2(\Omega)$ there exists a unique solution $(u, p)$. 

Least-squares formulation

Discrete Galerkin formulation

- Let consider the finite element spaces $V_h \subset H^1_0(\Omega)$ and $Q_h \subset L^2_0(\Omega)$.

  We set $W_h = V_h \times Q_h$;

- The Galerkin approximation of the problem is to find $(u_h, p_h) \in W_h$ such that

  $$A(u_h, p_h; v_h, q_h) = (f_h, v_h), \quad \text{for all } (v_h, q_h) \in W_h,$$

  where $A(u_h, p_h; v_h, q_h) = (\nabla u_h, v_h) - (p_h, \nabla \cdot v_h)$.

  And we know that if the finite element spaces satisfy the discrete inf-sup condition

  $$\inf_{q \in Q_h} \sup_{v \neq 0, q \neq 0} \frac{(\nabla v, q)}{\|v\|_{H^1(\Omega)} \|q\|_{L^2(\Omega)}} \geq \beta > 0,$$

  then $(u_h, p_h)$ is the unique solution of the weak formulation.

- Suppose the domain $\Omega$ is covered by an affine mesh $T_h$. 
Least-squares formulation

A stabilized variational formulation

- Define a modified bilinear form $A_h : W_h \times W_h \to \mathbb{R}$ and a modified functional $F_h$:

\[
A_h(u_h, p_h; v_h, q_h) = a(u_h, v_h) + b(v_h, q_h) - b(u_h, q_h) \\
+ \delta \sum_{T \in T_h} h_T^2 (-\Delta u_h + \nabla p_h, -\Delta u_h + \nabla p_h)_T,
\]  

(44)

\[
F_h(v_h, q_h) = (f, v_h) - \delta \sum_{T \in T_h} h_T^2 (f, -\Delta v_h + \nabla p_h)_T,
\]  

(45)

for a positive parameter $\delta$ suitably chosen.

- The Galerkin least-squares (GaLS) method is to find $(u_h, p_h) \in W_h$ such that

\[
A_h(u_h, p_h; v_h, q_h) = F_h(v_h, q_h), \quad \text{for all } (v_h, q_h) \in W_h.
\]  

(46)

- This is a strongly consistent approximation of Stokes the problem (22).

- If $\delta = 0$, the GaLS method reduces to the Galerkin approximation. The additional term depends on the residual of the (discrete) momentum equation and vanishes for the exact solution $(u, p)$. 
The Stokes model

Least-squares formulation

• find \((u_h, p_h) \in W_h\) such that

\[
A_h(u_h, p_h; v_h, q_h) = F_h(v_h, q_h), \quad \text{for all } (v_h, q_h) \in W_h. \tag{47}
\]

• Notice that this formulation leads to the identity

\[
A_h(u_h, p_h; u_h, p_h) = \| \nabla u_h \|^2_{L^2(\Omega)} + \delta \sum_{T \in T_h} \frac{h_T^2}{|T|} \| \nabla p_h \|^2_{L^2(T)},
\]

and observe that the kernel of the bilinear form \(A_h\) is reduced to the null function and thus there exists a unique solution to this problem.

**Lemma 4** The solution \((u_h, p_h) \in W_h\) to the stabilized Galerkin problem satisfies

\[
\| u - u_h \|_{H^1(\Omega)} + \left( \delta \sum_{T \in T_h} \frac{h_T^2}{|T|} \| \nabla p - \nabla p_h \|^2_{L^2(T)} \right)^{1/2} \leq C h,
\]

where \(C\) is a constant independent of \(h\).
Least-squares formulation

- Let now \{\varphi_j \in V_h\} and \{\phi \in Q_h\} be the bases for the spaces \(V_h\) and \(Q_h\), respectively. The matrix form of the stabilized Galerkin formulation becomes

\[
\begin{pmatrix}
    A & B^t \\
    B & -C
\end{pmatrix}
\begin{pmatrix}
    U \\
    P
\end{pmatrix} =
\begin{pmatrix}
    F \\
    G
\end{pmatrix},
\]

where

\[C = [c_{km} \text{ with } c_{km} = \delta \sum_{K} h_T^2 \int_{T_h} \nabla \phi_m \cdot \nabla \phi_k \, dx,\]

and the components of the right-hand side \(G\)

\[G = [g_l \text{ with } g_l = -\delta \sum_{K} h_T^2 \int_{T_h} f \text{ div } \phi_l \, dx].\]

As with the "classical" variational formulation, we can formally eliminate the velocity variable (since \(A^{-1}\) is well defined) and write

\[(B \ A^{-1} \ A^t + C)P = B \ A^{-1} \ F - G.\]

The matrix \(R = B \ A^{-1} \ B^t + C\) is non-singular and \(C\) is a positive definite matrix.
Solution methods for the Stokes system

Algebraic system

- The weak formulation equation is equivalent to a system of linear algebraic equations (39) for the components of vectors $U = [u_j]$ and $P = [p_k]$ with

$$u_h(x) = \sum_{j=1}^{N} u_j \varphi_j(x), \quad \text{and} \quad p_h(x) = \sum_{k=1}^{M} p_k \phi_k(x),$$

where $N = \text{dim}(V_h)$ and $M = \text{dim}(Q_h)$;

- The classical matrix formulation of the Stokes problem is:

$$K = \begin{pmatrix} A & B^t \\ B & 0 \end{pmatrix} \begin{pmatrix} U \\ P \end{pmatrix} = \begin{pmatrix} F \\ 0 \end{pmatrix}, \quad (48)$$

where $A$ is a regular matrix corresponding to the discretization of the Laplace operator for Dirichlet boundary conditions, often called the stiffness matrix; $B$ is the gradient matrix associated with the divergence matrix $B^t$. 
Resolution of the saddle-point problem

- The matrix of the linear system (48) is indefinite (neither positive nor negative). However, the Galerkin matrix is symmetric, because the continuous Stokes operator is self-adjoint.

- The Silvester law of inertia applied to the decomposition of the matrix

\[ K = \begin{pmatrix} A & B^t \\ B & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -BA^{-1}B^t \end{pmatrix} \begin{pmatrix} I & A^{-1}B^t \\ 0 & I \end{pmatrix} \]

shows that the number of positive and negative eigenvalues of the matrix \( K \) coincide with the one of the second factor. Hence, \( K \) has \( N \) positive eigenvalues and at least \( M - 1 \) negative eigenvalues, where \( N \) (resp. \( M \)) is the total number of degrees of freedom for the velocity (resp. the pressure).

- This system can be replaced by an equivalent system:

\[ \tilde{K} = \begin{pmatrix} A & B^t \\ -B & 0 \end{pmatrix} \]

where the matrix \( \tilde{K} \) is now positive definite (all eigenvalues have positive real part), but no longer symmetric.
Resolution of the saddle-point problem

Numerical considerations

- Solving this system using a direct method (i.e., elimination) may be acceptable for small systems \( n < 10^4 \), where \( n \) is the size of the system.

- In this case, the pivoting must be organized with row exchanges to avoid small pivots (since the matrix is indefinite).

- For large systems however, a direct method requires too much computational resources and memory \( O(n^3) \) and \( O(n^4) \). But, iterative methods cannot be chosen arbitrarily since many of them require a positive definite matrix. This feature precludes Gauss-Seidel, Jacobi or gradient-based techniques.

- Several other techniques can be used to solve this system, that take advantage of the structure of the system.
1. The penalty method

- Consists in introducing a perturbation into the original system (26) to penalize the divergence-free constraint:

\[
\begin{align*}
    a(u_\epsilon, v) + b(v, p_\epsilon) &= (f, v), & \text{for all } v &\in V \\
    b(u_\epsilon, q) - \epsilon \int_\Omega p_\epsilon q \, dx &= 0, & \text{for all } q &\in Q
\end{align*}
\]

where \( \epsilon > 0 \) is a small perturbation coefficient.

- The error introduced by this technique is \( O(\epsilon) \).

**Proposition 3** Let \((u, p)\) (resp. \((u_\epsilon, p_\epsilon)\)) denote the solution of the initial Stokes system (26) (resp. perturbed system (49)). We have the following error estimate

\[
\alpha \beta \| u - u_\epsilon \|_V + \alpha \beta^2 \| p - p_\epsilon \|_Q \leq c \epsilon \| p \|_Q .
\]

where \( \alpha > 0 \) is the coercivity constant associated with the bilinear form \( a(\cdot, \cdot) \) and \( \beta > 0 \) is the inf-sup constant.
Resolution of the saddle-point problem

- The matrix form of the penalized problem becomes

\[
\begin{pmatrix}
A & B^t \\
B & -\epsilon M_Q
\end{pmatrix}
\begin{pmatrix}
U_\epsilon \\
P_\epsilon
\end{pmatrix}
= \begin{pmatrix}
F \\
0
\end{pmatrix}
\]

or equivalently

\[
\begin{align*}
AU_\epsilon + B^t P_\epsilon &= F \\
BU_\epsilon - \epsilon M_Q P_\epsilon &= 0
\end{align*}
\]  \quad (50)

- Notice that the mass matrix \( M_Q \in \mathcal{M}_{M \times M}(\mathbb{R}) \) is invertible (symmetric and positive definite).

- By eliminating the pressure \( P_\epsilon \) from these equations, we obtain

\[
\left( A + \frac{1}{\epsilon} B^t (M_Q)^{-1} B \right) U_\epsilon = F.
\]

And once the velocity \( U_\epsilon \) has been computed by solving this equation, the pressure \( P_\epsilon \) can be easily obtained by solving the equation

\[
P_\epsilon = \frac{1}{\epsilon} (M_Q)^{-1} B U_\epsilon.
\]
Resolution of the saddle-point problem

- Notice that this penalization method requires computing the inverse of the matrix $M_Q$. Efficiency is directly related to the computational effort required to compute $(M_Q)^{-1}$.
  (For discontinuous pressure approximations ($P_0$ or $Q_0$ Lagrange FE), $M_Q$ can be inverted element-wise.)

- For small values of $\epsilon$, the system (50) is ill-conditioned $\Rightarrow$ convergence issue for iterative methods.

- A solution consists in considering the system

$$
\begin{pmatrix}
A & B^t \\
B & -\epsilon I
\end{pmatrix}
\begin{pmatrix}
U_\epsilon \\
P_\epsilon
\end{pmatrix}
=
\begin{pmatrix}
F \\
0
\end{pmatrix}
$$

or equivalently

$$
\begin{cases}
AU_\epsilon + B^t P_\epsilon = F \\
BU_\epsilon - \epsilon IP_\epsilon = 0
\end{cases}
$$

(51)

and likewise to eliminate the pressure from the equations

$$
\left( A + \frac{1}{\epsilon} B^t B \right) U_\epsilon = F.
$$

(52)

It can be shown that equation (52) is equivalent to the following equation

$$
(\nabla u_h, \nabla v_h) + \frac{1}{\epsilon} \sum_{i=1}^{M} (\text{div } u_h, q_i)(\text{div } v_h, q_i) = (f, v_h), \quad \forall u_h, v_h \in V_h.
$$
Resolution of the saddle-point problem

2. The augmented Lagrangian method:
\[ \rightsquigarrow \text{introduced for the solution of minimization problems with constraints.} \]

- Combines penalty and Lagrange multipliers techniques.

- Consider the Lagrangian functional \( L \in C^\infty(H_0^1(\Omega) \times L^2(\Omega)) \) associated with the Stokes problem

\[
L(v, q) = \frac{1}{2} a(v, v) + b(v, q) - f(v) = \frac{1}{2} \int_\Omega |\nabla v|^2 \, dx - \int_\Omega q \, \text{div} \, v \, dx - f(v),
\]

With \( r > 0 \), we associate with \( L \) the augmented Lagrangian functional \( L_r \) defined by

\[
L_r(v, q) = L + \frac{r}{2} \int_\Omega |\text{div} \, v|^2 \, dx, \quad \text{for all } (v, q) \in H^1(\Omega) \times L^2(\Omega).
\]

And we observe that

\[
L(v, q) \leq L_r(v, q), \quad \forall (v, q) \in H^1(\Omega) \times L^2(\Omega)
\]

\[
L(v, q) = L_r(v, q) = \frac{1}{2} \int_\Omega |\nabla v|^2 \, dx - f(v)
\]

\[= J(v), \quad \forall v \in H^2(\Omega); \text{div} \, v = 0, \quad \forall q \in L^2(\Omega).\]
Resolution of the saddle-point problem

**Theorem 10** The functionals $\mathcal{L}$ and $\mathcal{L}_r$ have the same saddle-points over $V \times L^2(\Omega)$.

- A simple iterative procedure can be employed to remove penalty errors:
  1. Assume $P_{\epsilon,0}$ is an arbitrary initial guess for the pressure.
  2. Knowing $P_{\epsilon,k}$, we compute $P_{\epsilon,k+1}$ from:

    \[
    \begin{pmatrix}
    A & B^t \\
    B & -\epsilon MQ \\
    \end{pmatrix}
    \begin{pmatrix}
    U_{\epsilon,k+1} \\
    P_{\epsilon,k+1} \\
    \end{pmatrix}
    =
    \begin{pmatrix}
    F \\
    -\epsilon MQ P_{\epsilon,k} \\
    \end{pmatrix},
    \]

  3. If $MQ$ can be inverted efficiently, we can write the previous equations as:

  \[
  \begin{cases}
  \left( A + \frac{1}{\epsilon} B^t (MQ)^{-1} B \right) U_{\epsilon,k} = F - B^t P_{\epsilon,k}, \\
  MQ P_{\epsilon,k+1} = MQ P_{\epsilon,k} + \frac{1}{\epsilon} BU_{\epsilon,k}.
  \end{cases}
  \]

- Convergence can be proved for any $\epsilon > 0$ and practically, setting $\epsilon$ to a small value ($\epsilon \leq 10^{-6}$) allows the previous sequence to converge in a few iterations.

- This iterative procedure is a special case of the Uzawa algorithm.
Resolution of the saddle-point problem

3. The Uzawa method - Schur complement

- Consider the saddle-point problem:

\[
\begin{cases}
\text{find } (u, p) \in X \times M \text{ such that } \\
\mathcal{L}(u, q) \leq \mathcal{L}(u, p) \leq \mathcal{L}(v, p), \quad \forall (v, q) \in X \times M,
\end{cases}
\]

with

\[
\mathcal{L}(v, q) = \frac{1}{2} a(u, v) + b(v, q) - f(v).
\]

- The algorithm for solving this problem is a gradient method applied to the minimization of the dual functional.

- Assume that \( \mathcal{L} \) is differentiable with respect to \( q \).
Resolution of the saddle-point problem

- The Uzawa algorithm consists in 3 steps:
  
  (a) given $p_0 \in M$
  
  (b) for $k \geq 0$, $p_k$ known, compute $u_k$ and $p_{k+1}$ in two steps:
    
    (b.1) compute velocity $u_k$ associated with $p_k$:
    
    $$u_k \in X, \text{ solving } \mathcal{L}(u_k, p_k) \leq \mathcal{L}(v, p_k), \quad \forall v \in X,$$
    
    (b.2) compute new pressure field $p_{k+1}$:
    
    $$p_{k+1} = \Pi_M(p_k + \rho \nabla_p \mathcal{L}(u_k, p_k)),$$
    
    where $\rho > 0$ and $\Pi_M$ denotes the orthogonal projection operator from $Q$ onto $M$.

- Convergence is proved for any $\rho$ sufficiently small.
Resolution of the saddle-point problem

- In matrix form, Uzawa’s method consists in eliminating the velocity $U$, thus yielding
  \[ B A^{-1} B^t P = B A^{-1} F. \]

- The Uzawa matrix $U_z = B A^{-1} B^t$ is given in implicit form since $A^{-1}$ is a dense matrix and the computational cost of inverting $A$ is too large in practice.

- Suitable iterative methods shall be used to compute it. Uzawa algorithm becomes:
  (a) given $P_0$
  (b) for $k = 0, 1, 2, \ldots$
    (b.1) solve $A U_k = F - B^t P_k$
    (b.2) set $P_{k+1} = P_k + \rho B U_k$

- Let $R_k = -B U_k$ denote the residual, then we have
  \[ R_k = -B A^{-1} (F - B^t P_k) = B A^{-1} B^t (P_k - P), \]

  and thus it is easy to show that
  \[ P_{k+1} - P_k = -\rho R_{k+1} = \rho B A^{-1} B^t (P - P_k) = \rho U_z (P - P_k). \]
Resolution of the saddle-point problem

• Remarks on Uzawa algorithm
  
  o Uzawa algorithm is equivalent to a gradient method with fixed time step $\rho$.

  o A speed-up can be obtained if a conjugate gradient method is used instead of the gradient method.

  o Preconditioning:
    
    – The iterative approaches may be very expensive, in particular because they are searching for the exact solution of the Laplace equation.
    
    – To speed up the process we can introduce a symmetric positive definite operator $C$ close to $U_z$:
      \[
      P_{k+1} = P_k - \rho C^{-1} B U_k .
      \]
    
    – $C^{-1}$ must not be expensive to compute.
    
    – Since only step $(b)$ is modified, the convergence properties of the method are preserved. The rate of convergence is still very dependent on the choice of $C$.  

Theoretical Aspects and Numerical Methods for Incompressible Fluids

EAFIT, Medellin, June 16-25, 2015

100
Section 2.3

Unsteady Stokes problem
The unsteady Stokes problem

Model: unsteady Stokes problem

- we study now the unsteady motion of a viscous incompressible fluid of constant density confined in a bounded domain $\Omega_t \subset \mathbb{R}^d$, subjected to a density of volume force per mass unit $f$, for $t \in [0, T]$;

- described by the following set of equations:

$$
\begin{aligned}
\begin{cases}
\text{Given } f; \text{ find } (u, p) \text{ such that:} \\
\frac{\partial u}{\partial t} - \nu \Delta u + \frac{1}{\rho} \nabla p = f, & \text{in } \Omega_t \times (0, T), \\
\text{div } u = 0, & \text{in } \Omega_t \times (0, T).
\end{cases}
\end{aligned}
$$

(54)

- endowed with the homogeneous Dirichlet boundary condition and the initial condition (i.e. a given velocity at time $t = 0$):

$$
\begin{aligned}
&u = 0, \quad \text{on } \partial \Omega_t \times (0, T), \\
u(x, 0) = u_0(x), & \text{with } \text{div } u_0 = 0 \text{ in } \Omega_t.
\end{aligned}
$$

(55)
Variational formulation

Equations (54) and (55) can be recast in a weak form using a mixed formulation.

- Consider the functional spaces \( V = H^1_0(\Omega_t) \) and \( Q = L^2_0(\Omega_t) \).

- By multiplying the first equation by a test function \( v \in V \) and by integrating over \( \Omega_t \), we get for almost every \( t \in (0, T) \):

\[
\int_{\Omega_t} \frac{\partial u}{\partial t} \cdot v \, dx - \int_{\Omega_t} \nu \Delta u \cdot v \, dx + \frac{1}{\rho} \int_{\Omega_t} \nabla p \cdot v \, dx = \int_{\Omega_t} f \cdot v \, dx.
\]

Using Green’s formula we can rewrite the partial terms and reformulate the previous equation as follows, for a.e. \( t \in (0, T) \):

\[
\int_{\Omega_t} \frac{\partial u}{\partial t} \cdot v \, dx + \int_{\Omega_t} \nu \nabla u \cdot \nabla v \, dx - \frac{1}{\rho} \int_{\Omega_t} p \, \text{div} \, v \, dx = \int_{\Omega_t} f \cdot v \, dx, \quad \forall v \in V \quad (56)
\]

- Likewise, we introduce the bilinear forms \( a(\cdot, \cdot) : (u, v) \in V \times V \to a(u, v) \in \mathbb{R} \) and \( b(\cdot, \cdot) : (u, q) \in V \times Q \to b(u, q) \in \mathbb{R} \), respectively defined as:

\[
a(u, v) = \int_{\Omega_t} \nu \nabla u \cdot \nabla v \, dx = \int_{\Omega} \nu \sum_{i=1}^{d} \nabla u_i \cdot \nabla v_i \, dx, \quad (57)
\]

\[
b(u, q) = -\int_{\Omega_t} \text{div} \, u \, q \, dx. \quad (58)
\]
Variational formulation

• To simplify, let assume the constant density $\rho = 1$.

• The weak form reads as follows:
  for given $f \in C^0([0, T]; L^2(\Omega_t))$ and $u_0 \in V$,

  $\begin{cases}
  \text{find a pair } (u, p) \in V \times Q \text{ such that for a.e. } t \in (0, T) \\
  \langle \partial_t u, v \rangle + a(u, v) + b(v, p) = (f, v), \\
  b(u, q) = 0, \\
  u(\cdot, t = 0) = u_0,
  \end{cases}$

  for all $v \in V$, 
  for all $q \in Q$, 

  (59)

• Here, we introduced the expression for almost every $t \in (0, T)$ to avoid any occurrence of a temporal discrepancy like a step function in the velocity field.

• We have the result:

  \textbf{Proposition 4} Given $f \in C^0([0, T]; L^2(\Omega_t))$ and $u_0 \in V$, there exists a unique solution to problem (59) for every $T > 0$.
The unsteady Stokes problem

- In order to avoid having an arbitrary value of the mean value of the pressure, we consider the functional spaces

\[
H = \{ v \in L^2(\Omega_t) ; \ \text{div} v = 0, \ v \cdot n = 0 \text{ on } \partial \Omega_t \}
\]

\[
H^1_{\int = 0}(\Omega_t) = \{ q \in H^1(\Omega_t) ; \int_{\Omega_t} q = 0 \},
\]

\[
V = \{ v \in H^1_0(\Omega_t) ; \ \text{div} v = 0 \},
\]

(and observe that \( V \) is dense in \( H \) and there is a continuous injection of \( V \) into \( H \).)

- For any \( \tilde{u} \in L^2(\Omega_t) \), there exists a unique pair \( u \in H \) and \( \phi \in H^1_{\int = 0}(\Omega_t) \) such that

\[
\begin{cases}
    u + \nabla \phi = \tilde{u}, & \text{in } \Omega_t, \\
    \text{div} u = 0, \\
    u \cdot n = 0, & \text{on } \partial \Omega_t.
\end{cases}
\] (60)

**Proposition 5** The functional space \( L^2(\Omega_t) \) can be decomposed as follows:

\[
L^2(\Omega_t) = H \oplus \nabla(H^1_{\int = 0}(\Omega_t)).
\] (61)
Finite element approximation

- In the Stokes problem (59), the space and time variables have different meanings and contribute differently to the phenomenon. This allows to proceed in different ways for the time and the space approximations.

- We introduce a Galerkin approximation for the space discretization the same way we did for the Stokes problem. This means that we have to pay attention to the choice of the approximation spaces to make sure that they lead to stable discretizations of the problem, i.e. that the discrete LBB inf-sup condition is satisfied.

- Then, for the resulting system of coupled ordinary differential equations, where the time is the independent variable, we consider a suitable approximation in time.
Space approximation

- Let \( \{ V_h \subset V \} \) and \( \{ Q_h \subset Q \} \) be two families of finite-dimensional vector spaces depending on a scalar parameter \( h > 0 \), and let \( u_{0h} \in V_h \) be an approximation of \( u_0 \).

- \( V_h \) and \( Q_h \) are compatible subspaces that approach \( V \) and \( Q \) as \( h \to 0 \).

- Problem (59) can then be replaced by the following Galerkin approximation:

\[
\begin{cases}
\text{find } u_h \in C^1([0, T]; V_h), p_h \in C^0([0, T]; Q_h) \text{ s.t. } & \forall t \in [0, T] \\
\langle \partial_t u_h, v_h \rangle + a(u_h, v_h) + b(v_h, p_h) = (f, v_h), & \text{for all } v_h \in V_h, \\
b(u_h, q_h) = 0, & \text{for all } q_h \in Q_h, \\
u_h(\cdot, t = 0) = u_{0h},
\end{cases}
\]

which is called a semi-discrete scheme.
Analysis of the semi-discrete scheme

• The spaces $V_h \subset H^1_0(\Omega_t)$ and $Q_h \subset L^2_0(\Omega_t)$ must satisfy a compatibility condition: there exists a constant $\beta > 0$ (independent of the parameter $h$) such that:

$$
\inf_{q_h \in Q_h} \sup_{v_h \in V_h} \frac{b(v_h, q_h)}{\|v_h\|_{H^1(\Omega_t)} \|q_h\|_{L^2(\Omega_t)}} \geq \beta \tag{63}
$$

• Furthermore, let assume that for all $v \in V$ and $q \in Q$ we have

$$
\inf_{v_h \in V_h} \|v - v_h\|_{H^1(\Omega_t)} + \inf_{q_h \in Q_h} \|q - q_h\|_{L^2(\Omega_t)} = O(h) \tag{64}
$$

• We have the following error estimate:

**Proposition 6** Under the previous hypothesis, the solution $(u_h, p_h)$ of the problem (62) is such that:

$$
\|u(t) - u_h(t)\|_{L^2(\Omega_t)} \leq \|u - u_{0h}\|_{L^2(\Omega_t)} e^{-c_1 t/2} + c_2(u, p) h^2,
$$

$$
\left(\frac{1}{T} \int_0^T \|u(t) - u_h(t)\|_{H^1(\Omega_t)} + \|p(t) - p_h(t)\|_{L^2(\Omega_t)} \right)^{1/2} \leq \frac{c_3}{\sqrt{T}} \|u - u_{0h}\|_{L^2(\Omega_t)} + c_4(u, p) h, \tag{65}
$$
The Stokes model

Time discretization

Actually, the approximate problem (62) can be discretized in time using techniques designed for parabolic problems.

- Let $\Omega_t$ be a domain of $\mathbb{R}^d$. We consider the following parabolic problem:

\[
\begin{cases}
\frac{\partial u}{\partial t} + Lu = f, & \forall x \in \Omega_t, \ t \geq 0, \\
u(x, t) = 0, & \forall x \in \partial \Omega_t, \ t > 0, \\
u(x, 0) = u_0(x), & \forall x \in \Omega_t,
\end{cases}
\]  

(66)

where $f = f(x, t)$ is a given function, $L = L(x)$ is an abstract elliptic operator acting on $u = u(x, t)$, the unknown.

- In order to find a weak formulation of problem (66), we multiply for each $t > 0$ the equation by a test function $v = v(x)$ taken in the functional space $V = H^1_0(\Omega)$, and integrate over $\Omega_t$:

\[
\int_{\Omega_t} \frac{\partial t}{\partial t} uv \, dx + a(t; u(t), v) = \int_{\Omega_t} f(t)v \, ds, \quad \forall v \in V
\]  

(67)

with $u(0) = u_0$ and $a(t; \cdot, \cdot)$ is the bilinear form associated with the elliptic operator $L$. 

Theoretical Aspects and Numerical Methods for Incompressible Fluids EAFIT, Medellin, June 16-25, 2015 109
Parabolic problems

- We introduce the following hypothesis:

  (i) the function $t \mapsto a(t; u, v)$ is measurable, $\forall u, v \in V$;

  (ii) the bilinear form $a(t; \cdot, \cdot)$ is continuous, that is:
  
  there exists $M$ such that $|a(t; u, v)| \leq M \|u\|_V \|v\|_V$, for a.e. $t > 0$;

  (iii) the form $a(t; \cdot, \cdot)$ is weakly coercive, that is for a.e $t > 0$, for all $u \in V$, there exists $\gamma \geq 0$, $\alpha > 0$ such that $a(t; u, u) \geq \alpha \|u\|_V^2 - \gamma \|u\|_{L^2(\Omega_t)}^2$.

- Furthermore, we require here $u_0 \in L^2(\Omega_t)$ and $f \in L^2(\Omega_t)$.

Then, we have the following result.

**Theorem 11 (J.L. Lions)** Under the hypothesis (i)-(iii), the problem (67) admits a unique solution.
Galerkin approximation of parabolic problems

- Consider an approximation space $V_h \subset V$ and $u_{0h}$ an approximation of $u_0$ in $V_h$.

- The Galerkin approximation of problem (67) reads:

\[
\begin{aligned}
\text{for each } t > 0, & \text{ find } u_h(t) \in V_h \text{ such that:} \\
\int_{\Omega_t} \partial_t u_h(t) \, v_h \, dx + a(t; u_h, v_h) &= \int_{\Omega_t} f(t) \, v_h \, dx, \quad \forall v_h \in V_h.
\end{aligned}
\]  

(68)

- This problem is called a semi-discretization of (67) since the time variable $t$ has not been discretized.

- The new problem is a system of coupled ordinary differential equations.

- The Cauchy-Lipschitz theorem guarantees the existence and the uniqueness of the solution $u_h(t)$ in $C^1([0, T]; V_h)$. 
Galerkin approximation of parabolic problems

• We introduce a finite element basis $(\phi_j)_{1 \leq j \leq N}$ for the space $V_h$.

• The solution to (68) can be written as follows, for each $t > 0$:

$$u_h(x, t) = \sum_{j=1}^{N} u_j(t) \phi_j(x)$$

where the coefficients $u_j(t)$ represent the unknowns of problem (68).

• We denote $d_t u_j(t)$ the derivative of the function $u_j(t)$ with respect to time. The previous system can be reformulated as:

$$\sum_{j=1}^{N} d_t u_j(t) \int_{\Omega_t} \phi_j \varphi_i \, dx + \sum_{j=1}^{N} a(\phi_j, \varphi_i) = \int_{\Omega_t} f(t) \varphi_i \, dx, \quad i = 1, 2, \ldots, N.$$  

• We define the vector of unknowns $u(t) = (u_i(t))_{1 \leq i \leq N}$ in $\mathbb{R}^N$, the right-hand side vector $f(t) = (f_i(t))_{1 \leq i \leq N}$, the mass matrix $M = (m_{ij})_{1 \leq i, j \leq N}$ and the stiffness matrix $A(t) = (a_{ij})_{1 \leq i, j \leq N}$ and the problem reads:

$$M d_t u(t) + A(t) u(t) = f(t), \quad t > 0.$$
Time-integration schemes

- Numerous time-integration methods have been proposed. Let $\delta t > 0$ be a scalar value representing the time step, $\delta t = t^{n+1} - t^n$ and denote $u^n_h$ the value of $u_h$ at time $t^n$.

- A classical method, the $\theta$-scheme, approximates the time derivative $d_t u(t^{n+1})$ by the differential quotient $\frac{1}{\delta t}(u^{n+1}_h - u^n_h)$ and compute the sequence $(u_0^h, \ldots, u^n_h), u^j_h \in V_h$, depending on the real parameter $\theta \in [0, 1]$ as:

\[
\frac{1}{\delta t}(u^{n+1}_h - u^n_h, v_h) + a(t^{n+1}; u^{n+1}_{\theta h}, v_h) = (f^\theta(t^{n+1}), v_h), \quad \forall v_h \in V_h
\]

where $u^{n+1}_{\theta h} = \theta u^{n+1}_h + (1 - \theta)u^n$ and $f^\theta(t^{n+1}) = \theta f(t^{n+1}) + (1 - \theta)f(t^n)$.

- In algebraic form, the $\theta$-scheme reads:

\[
M \frac{1}{\delta t}(u^{n+1} - u^n) + A(\theta u^{n+1} + (1 - \theta)u^n) = \theta f^{n+1} + (1 - \theta)f^n.
\]
The Stokes model

Time-integration schemes

- The $\theta$-scheme is known as
  - the **implicit Euler** scheme for $\theta = 1$ and writes:
    \[
    \frac{1}{\delta t}(u_h^{n+1} - u_h^n, v_h) + a(t^{n+1}; u_h^{n+1}, v_h) = (f(t^{n+1}), v_h), \quad \forall v_h \in V_h
    \]
  - the **implicit Crank-Nicolson** scheme for $\theta = 1/2$ and as
  - the **explicit Euler** scheme for $\theta = 0$.

- The **BDF2** scheme (second-order backward finite difference) is a two-step method:
  \[
  \frac{1}{\delta t}(3u_h^{n+1} - 4u_h^n + u_h^{n-1}, v_h) + a(t^{n+1}; u_h^{n+1}, v_h) = (f(t^{n+1}, v_h), \quad \forall v_h \in V_h.
  \]

- The **fractional-step scheme** consists in subdividing each time interval $[t^n, t^{n+1}]$ into three subintervals and then solve three systems.

- **Stability** and **consistency** are important numerical issues when dealing with stiff initial value problems. The implicit Euler method is consistent at order one, while the Crank-Nicolson scheme is consistent at order two in space and time. For $\theta \geq 1/2$, the $\theta$-scheme is unconditionally stable.
Backward Euler scheme

- Unsteady Stokes problem (62) is solved by constructing a sequence \((u_h, p_h)_{n \geq 0}\):
  1. solve \((u_h^0, v_h) = (u_{0h}, v_h)_{L^2(\Omega_t)},\) for all \(v_h \in V_h,\)
  2. find \(u_h^{n+1} \in V_h\) and \(p_h^{n+1} \in Q_H\) such that, for all \(v_h \in V_h, q_h \in Q_h:\)

\[
\begin{cases}
\frac{1}{\delta t} (u_h^{n+1} - u_h^n, v_h) + a(u_h^{n+1}, v_h) + b(v_h, p_h^{n+1}) = (f(t^{n+1}), v_h) \\
b(u_h^{n+1}, q_h) = 0.
\end{cases}
\] (69)

- At each time step, one has to resolve the following problem, for all \(v_h \in V_h, q_h \in Q_h:\)

\[
\begin{cases}
(u_h^{n+1}, v_h) + \delta t a(u_h^{n+1}, v_h) + b(v_h, \delta t p_h^{n+1}) = (u_h^n + \delta t f(t^{n+1}), v_h) \\
\tilde{a}(u_h^{n+1}, v_h) = b(u_h^{n+1}, q_h) = 0
\end{cases}
\]

a steady Stokes problem (no initial condition for the pressure is required).
The Stokes model

Backward Euler scheme

- This discretization method is unconditionally stable and is first-order accurate in time.

- It corresponds to the finite element space approximation of the strong steady formulation:

\[
\begin{aligned}
\frac{1}{\delta t} (u^{n+1} - u^n) - \nu \Delta u^{n+1} + \nabla p^{n+1} &= f^{n+1}, \\
\text{div } u^{n+1} &= 0,
\end{aligned}
\]

(70)

in which \((u^n, p^n)\) represents an approximation of \((u(x, n\delta t), p(x, n\delta t))\).

- Equation (69) can be also obtained from the time discretization of the weak form (59).

- We have the commutative diagram:

\[
\begin{array}{ccc}
\text{Eqns.}(54) & \xrightarrow{\text{time discretization}} & \text{Eqns.}(70) \\
\text{weak form} & \downarrow & \text{weak form} \\
\text{Eqn.(59)} & \xrightarrow{\text{time discretization}} & \text{Eqn.(69)}
\end{array}
\]
Projection methods

To solve the unsteady Stokes problem, we consider a Chorin-Temam projection scheme:

- At each time step, this approach involves 2 steps:
  1. a prediction step to determine a velocity field that does not consider the pressure, and
  2. a correction step to project this velocity field onto the space of divergence-free functions taking into account a pressure variable, thanks to the decomposition (61).

- Let $u^n \in V \approx u(\cdot, t^n)$. Set $u^0 = u_0$; then at every time step $t^n$ predict $\tilde{u}^{n+1}$:

  \begin{equation}
  \begin{cases}
    \frac{1}{\delta t} (\tilde{u}^{n+1} - u^n) - \Delta \tilde{u}^{n+1} = f^{n+1}, & \text{in } \Omega_t \\
    \tilde{u}^{n+1} = 0, & \text{on } \partial \Omega_t
  \end{cases}
  \end{equation}

  Then, $\tilde{u}^{n+1}$ is projected onto the space of divergence-free functions by solving:

  \begin{equation}
  \begin{cases}
    u^{n+1} + \nabla \phi^{n+1} = \tilde{u}^{n+1}, & \text{on } \Omega_t \\
    \text{div } u^{n+1} = 0, & \text{on } \partial \Omega \\
    u^{n+1} \cdot n = 0, & \text{on } \partial \Omega.
  \end{cases}
  \end{equation}

  to find $u^{n+1} \in H$ and $\phi^{n+1} \in H^1_f = 0(\Omega_t)$. 
Projection methods

- Actually, this projection step is equivalent to solving the following system:

\[
\begin{align*}
(ii) \quad & \begin{cases} 
\frac{1}{\delta t} (u^{n+1} - \tilde{u}^{n+1}) + \nabla p^{n+1} = 0, & \text{on } \Omega_t \\
\text{div } u^{n+1} = 0, & \\
u^{n+1} \cdot n = 0, & \text{on } \partial \Omega.
\end{cases}
\end{align*}
\]

(72)

Indeed, by multiplying the first equation by $1/\delta t$ and adding it to (71), we obtain the following expression:

\[
\frac{1}{\delta t} (u^{n+1} - u^n) - \Delta \tilde{u}^{n+1} + \nabla \left( \frac{1}{\delta t} \phi^{n+1} \right) = f^{n+1},
\]

that can be interpreted as an approximation of the momentum equation, where $1/\delta t \phi^{n+1}$ is considered as an approximation of the pressure term.

Now, setting $\phi^{n+1} = \delta t p^{n+1}$ leads to the previous system.
Projection methods

Two theoretical results follow:

- **Lemma 5 (Rannacher)** If the solution \((u, p)\) of problem (59) is sufficiently smooth in time and space, the solution of the fractional-step method (71)-(72) is such that:

  \[
  \sup_{t^n \in [0, T]} (\|u(t^n) - u^n\|_{L^2(\Omega_t)}, \|u(t^n) - \tilde{u}^n\|_{L^2(\Omega_t)}) \leq c \delta t, \tag{73}
  \]

  \[
  \sup_{t^n \in [0, T]} (\|p(t^n) - p^n\|_{L^2(\Omega_t)}, \|u(t^n) - \tilde{u}^n\|_{H^1(\Omega_t)}) \leq c (\delta t)^{1/2}. \tag{74}
  \]

- **Corollary 2** If the solution \((u, p)\) of problem (59) is sufficiently smooth in time and space, the solution of the fractional-step method (76)-(77) is such that:

  \[
  \sup_{t^n \in [0, T]} (\|u(t^n) - u^n\|_{L^2(\Omega_t)}, \|u(t^n) - \tilde{u}^n\|_{L^2(\Omega_t)},
  \|u(t^n) - \tilde{u}^n\|_{H^1(\Omega_t)}, \|p(t^n) - p^n\|_{L^2(\Omega_t)}) \leq c \delta t. \tag{75}
  \]
The numerical algorithm has been improved by modifying the computation of the intermediate velocity.

The modified version reads:

Set $u^0 = u_0$ and $p^0 = p|_{t=0}$ and solve at every time step $t^n$ the system:

\[(i) \begin{cases} \frac{1}{\delta t} (\tilde{u}^{n+1} - u^n) - \Delta \tilde{u}^{n+1} = f^{n+1} - \nabla p^n, & \text{in } \Omega_t \\ \tilde{u}^{n+1} = 0, & \text{on } \partial \Omega_t \end{cases}\]

\[(76)\]

and the projection stage is modified accordingly as follows:

\[(ii) \begin{cases} \frac{1}{\delta t} (u^{n+1} - \tilde{u}^{n+1}) + \nabla (p^{n+1} - p^n) = 0, & \text{on } \Omega_t \\ \text{div } u^{n+1} = 0, \\ u^{n+1} \cdot n = 0, & \text{on } \partial \Omega. \end{cases}\]

\[(77)\]
Finite element approximation

The FE approximation of the Chorin-Temam projection schemes is straightforward.

- Given the approximation spaces $V_h$ and $Q_h$, the problems to solve is:

$$
\begin{aligned}
\text{Find } \tilde{u}^{n+1} \text{ in } V_h \text{ such that for every } v_h \in V_h \\
\frac{1}{\delta t}(\tilde{u}_h^{n+1} - \tilde{u}_h^n, v_h) + a(\tilde{u}_h^{n+1}, v_h) + b(v_h, p_h^{n+1} - p_h^n) = (f(t^{n+1}), v_h),
\end{aligned}
$$

(78)

- and then the uncoupled pressure problem:

$$
\begin{aligned}
\text{Find } p^{n+1} \text{ in } Q_h \text{ such that for every } q_h \in Q_h \\
(\nabla (p_h^{n+1} - p_h^n), \nabla q_h) = -\frac{1}{\delta t} (\text{div } \tilde{u}_h^{n+1}, q_h).
\end{aligned}
$$

(79)

- These two problems are two coercive problems that are not required to satisfy the inf-sup condition when solved one after another. However, the inf-sup condition must be satisfied for the projection scheme to be convergent.
Finite element approximation

- We can easily recast the sub-problems (76)-(77) as follows:

\[
(i) \quad \frac{1}{\delta t} (\tilde{u}^{n+1} - u^n) - \Delta \tilde{u}^{n+1} + \nabla p^n = f^{n+1}, \quad \tilde{u}^{n+1}|_{\partial \Omega_t} = 0
\]

\[
(ii) \quad \begin{cases}
\frac{1}{\delta t} (u^{n+1} - \tilde{u}^{n+1}) + \nabla \phi = 0 \\
- \Delta \phi = -\frac{1}{\delta t} \text{div} \tilde{u}^{n+1} \\
\text{div} u^{n+1} = 0, \quad u^{n+1} \cdot n|_{\partial \Omega} = 0
\end{cases}
\]

\[
(iii) \quad p^{n+1} = p^n + \phi
\]

- For the prediction step, the boundary condition $\tilde{u}^{n+1} = 0$ is considered and it is updated to $u^{n+1} \cdot n = \nabla \phi \cdot n = 0$ on $\partial \Omega_t$ in the second step.
Finite element approximation

- When discretized using a finite element method, we obtain an algebraic system of the form:

\[
\begin{align*}
(i) & \quad \frac{1}{\delta t} (M\tilde{U}^{n+1} - MU^n) + A\tilde{U}^{n+1} + B^t P^n = F^{n+1} \\
(ii) & \quad L\Phi = \frac{1}{\delta t} B\tilde{U}^{n+1}, \quad MU^{n+1} = M\tilde{U}^{n+1} - \delta t B^t \Phi \\
(iii) & \quad P^{n+1} = P^n + \Phi
\end{align*}
\]

(81)

where \(M\) represents the velocity mass matrix, and \(A, B\) and \(L\) are the matrices associated with the velocity Laplacian, the pressure gradient and the pressure Laplacian operators, respectively.

- A BDF2 scheme has been proposed to approximate the time-derivative:

\[
\frac{1}{2\delta t} (3\tilde{u}^{n+1} - 4u^n + u^{n-1}) - \Delta \tilde{u}^{n+1} = f^{n+1} - \nabla p^n, \quad \tilde{u}^{n+1}|_{\partial\Omega_t} = 0
\]

(82)

\[
\begin{align*}
\frac{3}{2\delta t} (u^{n+1} - \tilde{u}^{n+1}) + \nabla (p^{n+1} - p^n) &= 0, \quad \text{in } \Omega_t \\
\text{div } u^{n+1} &= 0, \\
u^{n+1} \cdot n &= 0, \quad \text{on } \partial\Omega_t
\end{align*}
\]

(83)
Section 2.4
the Finite Element Method
Fundamentals of finite element method

The Finite Element Method (FEM) is a generic approach based on the approximation of continuous functions using piecewise polynomial basis functions.

- Domain $\Omega$ is covered by a triangulation $T_h$ composed of finite elements

\[
\Omega = \bigcup_{K \in T_h} K
\]

**Definition 11** A triangulation $T_h$ of the domain $\Omega$ is a decomposition of the domain $\Omega$ in a finite number of simplices $K$, obtained from a so-called reference element $\hat{K}$ and from one to one (bijective) transformations $T_K$ such that

\[
K = T_K(\hat{K}).
\]

- The triangulation is said to be **conforming**, when for every couple $(K_i, K_j)_{i \neq j}$ of simplices in $T_h$, the intersection $K_i \cap K_j$ is
  - either empty, or a vertex in dimension 1;
  - either empty, or a common vertex or an edge in dimension 2;
  - either empty, or a common vertex or an edge or a face in dimension 3.
Definitions and terminology

- The parameter $h$ represents the *grain* of the triangulation $T_h$

$$h = \max_{K \in T_h} h_K \quad \text{with} \quad h_K = \text{diam}(K).$$

**Definition 12** In an isotropic triangulation, all triangles shall satisfy a uniform estimate, for a scalar constant $C$, of the form

$$\sigma_K = \frac{h_K}{\rho_K} \leq C.$$

- The finite element subspace $V_h$ is composed of piecewise-polynomial functions;

$$V_h = \{ v_h \in C^k(\Omega) ; \ v_h|_K \in \mathbb{P}_k, \ \forall K \in T_h \}.$$

- Any function of $V_h$ is uniquely determined by a finite number of degrees of freedom (values of the function at mesh nodes).

- Each basis function $\varphi_i$ with compact support is associated with one degree of freedom.
Lagrange finite element

- A finite element is locally defined by a set \((K, P_K, \Sigma_K)\), where:
  1. \(K\) is a non-empty compact and connected set,
  2. \(P_K\) is a vector space of polynomials of dimension \(M\),
  3. \(\Sigma_k\) is a set of \(M\) independent linear forms \(\varphi_i\) defined on \(P_k\).

- The set \(\Sigma_K\) is said to be unisolvent on \(P_k\) if for each set of \(M\) scalar values \((\alpha_i)_{1 \leq i \leq M}\), there exists a unique \(p \in P_K\) such that
  \[
  \forall 1 \leq i \leq M, \quad \varphi_i(p) = \alpha_i.
  \]

- If \(\Sigma_K\) is unisolvent on \(P_K\), the triple \((K, P_K, \Sigma_K)\) is a Lagrange finite element.

- We consider the basis \(\{p_i\}_{i=1,\ldots,M}\) of \(P_K\) defined by \(\forall 1 \leq j \leq M, \varphi_j(p_i) = \delta_{i,j}\), and every \(p \in P_K\) is written as:
  \[
  p = \sum_{i=1}^{M} \varphi_i(p)p_i.
  \]
Finite elements

- $\Sigma_k$ is called the set of the degrees of freedom of the finite element $(K, P_K, \Sigma_K)$.

- We consider the interpolation operator $\Pi_K$: if $v$ belongs to the definition set of the linear forms of $\Sigma_K$,

  $$\Pi_K(v) = \sum_{i=1}^{M} \varphi_i(v)p_i.$$ 

- Similarly, $\Pi_K(v)$ is the unique element of $P_K$ such that

  $$\forall 1 \leq i \leq M, \quad \varphi_i(\Pi_K(v)) = \varphi_i(v).$$ 

- $\Pi_K(v)$ interpolates $v$ on $P_K$ and we have

  $$\forall p \in P_K, \quad \Pi_K(p) = p.$$
Polynomial spaces

- We introduce the set $P_k$ of the polynomials $p$ with scalar coefficients of $\mathbb{R}^d$ in $\mathbb{R}$ of degree less than or equal to $k$:

$$P_k = \left\{ p(x) = \sum_{i_1, \ldots, i_d \geq 0}^{i_1 + \ldots + i_d \leq k} \alpha_{i_1, \ldots, i_d} x_1^{i_1} \ldots x_d^{i_d}, \quad \alpha_{i,j} \in \mathbb{R}, \quad x = (x_1, \ldots, x_d) \right\}.$$

- It is easy to verify that $P_k$ is a vector space of dimension:

$$\dim(P_k) = \sum_{l=0}^{k} \binom{d+l-1}{l} = \binom{d+k}{k} = \begin{cases} k + 1 & d = 1 \\ \frac{1}{2}(k+1)(k+2) & d = 2 \\ \frac{1}{6}(k+1)(k+2)(k+3) & d = 3 \end{cases}$$

- The notion of lattice $\Sigma_k$ of a simplex $K$ allows to define a bijective mapping between a space of polynomials $P_k$ and a set of points $(\sigma_j)_{1 \leq j \leq N_k}$.
Triangular finite elements

- The finite element method for triangular Lagrange $\mathbb{P}_k$ elements involves the discrete finite dimensional functional space:

$$V_h^k = \left\{ v \in C^0(\Omega), \ v_h|_{K_j} \in \mathbb{P}_k, \ K_j \in T_h \right\},$$

and its subspace:

$$V_{0,h}^k = \left\{ v_h \in V_h^k, \ v_h = 0 \text{ on } \partial\Omega \right\}.$$ 

- **Lemma 6** The space $V_h^k$ is a subspace of the space $H^1(\Omega)$ of finite dimension corresponding to the number of degrees of freedom. Furthermore, there exists a basis $(\varphi_j)_{1 \leq j \leq N_{dof}}$ of $V_h^k$ defined by:

$$\varphi_i(a_j) = \delta_{ij}, \quad 1 \leq i, j \leq N_{dof},$$

such that every function $v_h \in V_h^k$ can be uniquely written as

$$v_h(x) = \sum_{i=1}^{N_{dof}} v_h(a_i) \varphi_i(x).$$
The reference finite element

• By convention, the vertices of the reference simplex $\hat{K}$ are given by $\hat{a}_0 = (0, \ldots, 0)$ (origin) and the points $\hat{a}_i = (0, \ldots, 1, \ldots, 0)$.

• $F_K : \mathbb{R}^d \to \mathbb{R}^d$ is the unique affine transformation that maps $\hat{a}_i$ on $a_i$ for all $i$:

$$\forall x \in \mathbb{R}^d, \quad x = F_K(\hat{x}) = a_0 + B_K \hat{x},$$

where $B_K$ is a $d \times d$ matrix; column $i$ is given by the coordinates of $a_i - a_0$.

• Since the simplex $K$ is non-degenerated, $B_K$ is invertible and $F_K$ is a one-to-one mapping that maps $\hat{K}$ on $K$ and we have:

$$|K| = |\det(B_K)| |\hat{K}| = \frac{|\det(B_K)|}{d!}.$$

• We use the notation $\hat{\cdot}$ in the reference element and we denote $\hat{q}$ any quantity obtained by transporting a quantity $q$ using the transformation $F_K$. 
Triangular finite elements

- We denote:

\[ \hat{x} = F_K^{-1}(x) = B_K^{-1}(x - a_0) \Leftrightarrow x = F_K(\hat{x}). \]

- For every function \( v \) defined on \( K \), we define \( \hat{v} \) defined on \( \hat{K} \) as:

\[ \hat{v}(\hat{x}) = v \circ F_K(\hat{x}) = v(F_K(\hat{x})), \]

and for every function \( \hat{v} \) defined on \( \hat{K} \), we write:

\[ v(x) = \hat{v} \circ F_K^{-1}(x) = \hat{v}(F_K^{-1}(x)). \]

- Similarly, if \( \psi \) is a linear form acting on the functions defined on \( K \), we define the transported linear form \( \hat{\psi} \) acting on the functions defined on \( \hat{K} \) as:

\[ \hat{\psi}(\hat{v}) = \psi(v). \]

- The barycentric coordinates are preserved by the affine transformation \( F_K \):

\[ \hat{\lambda}_i(\hat{x}) = \lambda_i(x). \]
● **Proposition 7** We denote by $\| \cdot \|$ the Euclidean norm of $\mathbb{R}^d$ and its subordinate norm. Hence, we have:

$$\| B_K \| \leq \frac{h_K}{\rho_{\hat{K}}}, \quad \| B_K^{-1} \| \leq \frac{h_{\hat{K}}}{\rho_K}, \quad |\det(B_K)| = \frac{|K|}{|\hat{K}|}.$$

● Transformation of the derivatives:

$$\nabla_{\hat{x}} \hat{v}(\hat{x}) = (B_K^t \nabla_x v) \circ F_K(\hat{x}), \quad \nabla_x v(x) = ((B_K^t)^{-1} \nabla_{\hat{x}} \hat{v}) \circ F_K^{-1}(x).$$
Triangular finite elements

- Triangle $K$ of type (1) in $\mathbb{R}^2$: $P_K = \mathbb{P}_1$, $\Sigma_K = \{f \mapsto f(a^i); \ 1 \leq i \leq 3\}$
  $\Sigma_K$ is defined on $C^0(K)$. The basis functions are $p_i \in \mathbb{P}_1$ and $p_i(a^j) = \delta_{i,j}$, hence
  \[ p_i = \lambda_i \quad \text{and} \quad \Pi_K(v) = \sum_{i=1}^{3} v(a^i)\lambda_i. \]

- Triangle $K$ of type (2) in $\mathbb{R}^2$: $P_K = \mathbb{P}_2$, $\Sigma_K = \{f \mapsto f(a^i)\} \cup \{f \mapsto f(a^{ij})\}$
  $\Sigma_K$ is defined on $C^0(K)$. $\text{card}(\Sigma_K) = 6$ and $\dim(\mathbb{P}_2) = 6$. The basis functions are
  \[ p_i = 2\lambda_i(\lambda_i - \frac{1}{2}), \quad p_{ij} = 4\lambda_i\lambda_j, \]
  \[ \Pi_K(v) = 2\sum_{i=1}^{3} v(a^i)\lambda_i(\lambda_i - \frac{1}{2}) + 4\sum_{1 \leq i < j \leq 3} v(a^{ij})\lambda_i\lambda_j. \]
Chapter III

The Navier-Stokes Model
Section 3.1
Mathematical and numerical analysis
The Navier-Stokes model

- Suppose the fluid is confined in a domain \( \Omega_t \subset \mathbb{R}^d \) (i.e. an open and connected region in two or three dimensions of space), for \( t \in [0, T] \).

- We recall that the isothermal flow of Newtonian viscous fluids of constant density is modeled by the following Navier-Stokes equations:

\[
\begin{cases}
\text{Given } f, \text{ find } u \text{ such that } \\
\rho \left( \frac{\partial u}{\partial t} + (u \cdot \nabla)u \right) - \mu \Delta u + \nabla p = \rho f, & \text{in } \Omega_t \times (0, T), \\
\text{div } u = 0, & \text{in } \Omega_t \times (0, T),
\end{cases}
\]

where \( u \) represents the velocity, \( p \) the pressure, \( \mu \) is the dynamic viscosity, \( \rho \) is the density and \( f \) denotes a density of volume forces per mass unit.

- The non-linear term \( (u \cdot \nabla)u \) corresponds to the convective acceleration, while the diffusion term \( \Delta u \) models the viscous effects.
The Navier-Stokes model

• By dividing both sides of the first equation by the constant density \( \rho \), we obtain the alternate formulation

\[
\begin{align*}
\frac{\partial u}{\partial t} + (u \cdot \nabla)u - \nu \Delta u + \frac{1}{\rho} \nabla p &= f, & \text{in } \Omega_t \times (0, T), \\
\text{div } u &= 0, & \text{in } \Omega_t \times (0, T),
\end{align*}
\]  

(85)

where \( \nu = \mu/\rho \) is the \textit{kinematic viscosity} coefficient.

• Introducing the Reynolds number \( Re = UL/\nu = \rho UL/\mu \), where \( L \) and \( U \) are characteristic of the length and velocity scales of the flow, and dividing both sides of (85) by \( U^2/L \), allows us to rewrite the equations in a \textit{dimensionless form}:

\[
\begin{align*}
\frac{\partial u}{\partial t} + (u \cdot \nabla)u - \frac{1}{Re} \Delta u + \nabla p &= f, & \text{in } \Omega_t \times (0, T), \\
\text{div } u &= 0, & \text{in } \Omega_t \times (0, T).
\end{align*}
\]  

(86)

where the introduced the change of variables

\[
x = x/L, \quad t = t(U/L), \quad u = u/U, \quad p = p/\rho U^2, \quad f = f(L/U^2).
\]
Initial and boundary conditions

- As such, problems (84) or (86) are not well-posed since these equations may have an infinite number of solutions.
- It is mandatory to introduce further conditions, namely
  (a) an initial condition (i.e. at time $t = 0$):
    \[
    u(x, 0) = u_0(x), \quad \text{in } \Omega
    \]  
    where $u_0$ is a given divergence-free vector field and
  (b) a Dirichlet boundary condition:
    \[
    u(x, t) = g(x, t), \quad \text{on } \partial \Omega \times (0, T),
    \]  
    where $g(x, t)$ is a given function defined over $\partial \Omega \times (0, T)$.

If $\Omega$ is bounded, then we have, thanks to the divergence theorem and the continuity condition,

\[
\int_{\partial \Omega} g(x, t) \cdot n \, ds = \int_{\Omega} \text{div } u(x, t) \, dx = 0,
\]  
where $n$ is the outward unit normal to the domain boundary $\partial \Omega$. 
Initial and boundary conditions

• Other boundary conditions, for example
  
  (c) **mixed boundary conditions:**

  \[ u = g_0, \quad \text{on } \Gamma_D \times (0, T) \quad \sigma n = g_1, \quad \text{on } \Gamma_N \times (0, T) \]  

  \( \Gamma_D \) and \( \Gamma_N \) are such that \( \Gamma_D \cap \Gamma_N = \emptyset \), \( \Gamma_D \cup \Gamma_N = \partial \Omega \), \( g_0 \) and \( g_1 \) are two given functions of \((x, t)\).

  \( \sigma \) is the stress tensor: \( \sigma = 2\mu D(u) - p \text{Id} \)

  and \( D(u) \) is the deformation rate tensor: \( 2D(u) = \nabla u + (\nabla u)^t \).

  

  (d) In applications, other mixed conditions are often preferred:

  \[ u = g_0, \quad \text{on } \Gamma_D \times (0, T) \quad \mu \frac{\partial u}{\partial n} - p n = g_1, \quad \text{on } \Gamma_N \times (0, T) \]  

  ○ Notice that in most applications, the function \( g_1 \) is set to zero in the Neumann boundary condition.
Variational formulation

- We return to the Navier-Stokes equations in primitive variables:

\[
\begin{aligned}
\begin{cases}
\text{Given } f, \text{ find } u \text{ such that } \\
\frac{\partial u}{\partial t} + (u \cdot \nabla) u - \nu \Delta u + \frac{1}{\rho} \nabla p = f, \quad \text{in } \Omega_t \times (0, T), \\
\text{div } u = 0, \quad \text{in } \Omega_t \times (0, T), \\
u = u_0, \quad \text{with } \text{div } u_0 = 0,
\end{cases}
\end{aligned}
\]

endowed with the following mixed boundary conditions:

\[
u = g_0, \quad \text{on } \Gamma_D \times (0, T) \quad \sigma n = g_1, \quad \text{on } \Gamma_N \times (0, T)
\]

where we considered \[\sigma = 2\nu D(u) - \frac{p}{\rho} \text{Id}.\]

- The mathematical theory for the analysis of the Navier-Stokes model and the estimation of the approximation error in finite element method rely on functional analysis and Sobolev spaces.

- In particular, the finite element approximation requires to recast the Navier-Stokes equations into a weak form.
We introduce the Hilbert functional space

\[ V = \{ v \in H^1_0(\Omega_t) ; \, v = 0 \text{ on } \Gamma_D \} , \]

which is a Hilbert space for the scalar product and the norm

\[
(u, v)_{H^1(\Omega)} = \int_\Omega u(x)v(x) \, dx + \int_\Omega \nabla u(x) : \nabla v(x) \, dx
\]

\[
\| u \|_{H^1(\Omega)} = \int_\Omega |u(x)|^2 \, dx + \int_\Omega |\nabla u|^2 \, dx
\]

Space \( V \) will coincide with \( H^1_0(\Omega_t) \) if the Dirichlet boundary \( \Gamma_D \) coincide exactly with \( \partial \Omega_t \). In this case, the prescribed Dirichlet function \( g_0 \) must be compatible with the divergence-free constraint, \( i.e. \)

\[
\int_{\partial \Omega_t} g_0 \cdot n \, ds = \int_{\Omega_t} \text{div} \, u \, dx = 0 .
\]
Weak form of the problem

- By multiplying the first equation by a test function $v \in V$ and integrating over $\Omega_t$, we obtain for almost every $t \in (0, T)$:

$$
\int_{\Omega_t} \frac{\partial u}{\partial t} \cdot v \, dx + \int_{\Omega_t} ((u \cdot \nabla)u) \cdot v \, dx - \int_{\Omega_t} \nu \Delta u \cdot v \, dx + \frac{1}{\rho} \int_{\Omega_t} \nabla p \cdot v \, dx = \int_{\Omega_t} f \cdot v \, dx.
$$

- Green's formula allows to rewrite the partial terms as follows

$$
- \int_{\Omega_t} \nu \Delta u \cdot v \, dx = \int_{\Omega_t} \nu \nabla u : \nabla v \, dx - \int_{\partial \Omega_t} \nu \frac{\partial u}{\partial n} \, ds
$$

$$
\int_{\Omega_t} \nabla p \cdot v \, dx = - \int_{\Omega_t} p \, \text{div} \, v \, dx + \int_{\partial \Omega_t} p v \cdot n \, ds
$$

- Then, taking into account the mixed boundary conditions (93), we have for almost every time $t$ in $(0, T)$

$$
\int_{\Omega_t} \frac{\partial u}{\partial t} \cdot v \, dx + \int_{\Omega_t} ((u \cdot \nabla)u) \cdot v \, dx + \int_{\Omega_t} \nu \nabla u : \nabla v \, dx - \frac{1}{\rho} \int_{\Omega_t} p \, \text{div} \, v \, dx
$$

$$
= \int_{\Omega_t} f \cdot v \, dx + \int_{\Gamma_N} g_1 \cdot v \, ds, \quad \forall v \in V \quad (94)
$$

completed with the Dirichlet boundary condition (93) on $\Gamma_D \times (0, T)$. 
Weak form of the problem

- As expected, the Neumann boundary condition $\sigma n = g_1$ on $\Gamma_N \times (0, T)$ is naturally satisfied by the formulation, which is often called the variational formulation of the momentum equation.

- Introducing the deformation rate tensor $D(u)$ yields the following formulation, for almost any $t \in (0, T)$:

$$
\int_{\Omega_t} \frac{\partial u}{\partial t} \cdot v \, dx + \int_{\Omega_t} ((u \cdot \nabla)u) \cdot v \, dx + 2\nu \int_{\Omega_t} D(u) : D(v) \, dx - \frac{1}{\rho} \int_{\Omega_t} p \, \text{div} \, v \, dx
$$

$$
= \int_{\Omega_t} f \cdot v \, dx + \int_{\Gamma_N} g_1 \cdot v \, ds, \quad \forall v \in V \quad (95)
$$

completed with the Dirichlet boundary condition (93) on $\Gamma_D \times (0, T)$. Likewise, the Neumann boundary condition is satisfied automatically.

- Similarly for the continuity equation, we first multiply the equation by a test function $q$ in a space $Q$ and then we integrate it over $\Omega_t$, to obtain the equation:

$$
\int_{\Omega_t} q \, \text{div} \, u \, dx = 0, \quad \forall q \in Q. \quad (96)
$$
Remark 1  (i) The main difference between this weak formulation and the Stokes weak formulation is related to the convective term. The latter is identified with a trilinear form $c : H^1(\Omega_t)^3 \to \mathbb{R}$ that will be defined later.

(ii) All integrals involved in the bilinear and the trilinear forms are finite.

- The mathematical analysis of incompressible Navier-Stokes equations for viscous flows has been a topic of interest for almost a century since they were introduced by Navier in 1822 and Stokes in 1845.

- The seminal work of Leray proved the existence of solutions when the flow region is the full space $\mathbb{R}^n$.

- An early worthwhile result is due to J.L. Lions established that if the flow domain is two-dimensional then the unsteady Navier-Stokes equations lead a unique solution.

- In three dimensions of space, there is no theoretical result about the uniqueness (and regularity) of a solution to the time-dependent Navier-Stokes equations modelling the flow of incompressible viscous fluids.
The steady Navier-Stokes problem

- We first consider the steady Navier-Stokes equations that describe the motion of an homogeneous incompressible Newtonian fluid.

- The motion is here independent of time.

- The problem is posed in a bounded domain \( \Omega \subset \mathbb{R}^d \) with a Lipschitz continuous boundary \( \partial \Omega \) and reads as follows:

\[
\begin{align*}
\text{Given } f, \text{ find } u \text{ such that } & \\
- \nu \Delta u + (u \cdot \nabla) u + \frac{1}{\rho} \Delta p = f, & \text{in } \Omega \\
\text{div } u = 0, & \text{in } \Omega
\end{align*}
\]  

(97)

In addition, we consider homogeneous Dirichlet boundary conditions,

\[
u = 0, \quad \text{on } \partial \Omega,
\]

(98)

corresponding to the case where the fluid is confined in the domain with fixed boundary.

- This problem close to the Stokes problem, except for the presence of the non-linear convective term \((u \cdot \nabla) u\) that makes it much more difficult to solve.
Mathematical results

- Let consider the functional spaces $V = H^1_0(\Omega)$ and $Q = L^2_0(\Omega)$.

- The variational formulation is almost identical to the formulation (94), except for the integral of the time derivative:
  Given $f \in L^2(\Omega)$, find $(u, p)$ such that for all $v \in V$
  \[
  \int_{\Omega} ((u \cdot \nabla)u) \cdot v \, dx + \int_{\Omega} \nu \nabla u : \nabla v \, dx - \frac{1}{\rho} \int_{\Omega} p \, \text{div} \, v \, dx = \int_{\Omega} f \cdot v \, dx + \int_{\Gamma_N} g_1 \cdot v \, ds
  \]
  \quad (99)

- The weak formulation of the continuity equation is strictly identical to equation (96):
  \[
  \int_{\Omega} q \, \text{div} \, u \, dx = 0, \quad \forall \, q \in Q.
  \]
  \quad (100)

- The formulation (99) can be conveniently recast using bilinear and trilinear forms.
Mathematical results

- We introduce the bilinear forms $a(\cdot, \cdot) : (u, v) \in V \times V \rightarrow a(u, v) \in \mathbb{R}$ and $b(\cdot, \cdot) : (u, q) \in V \times Q \rightarrow b(u, q) \in \mathbb{R}$, (same as the Stokes problem):

\[
a(u, v) = \int_{\Omega} \nu \nabla u : \nabla v \, dx = \int_{\Omega} \nu \sum_{i=1}^{d} \nabla u_i : \nabla v_i \, dx, \quad (101)
\]

\[
b(u, q) = -\int_{\Omega} \text{div} \, u \, q \, dx. \quad (102)
\]

and the trilinear form $c(\cdot; \cdot, \cdot) : (u, w, v) \in V \times V \times V \rightarrow c(w; u, v) \in \mathbb{R}$ associated with the non-linear term and defined as:

\[
c(w; u, v) = \int_{\Omega} ((w \cdot \nabla) u) \cdot v \, dx = \sum_{i,j=1}^{d} \int_{\Omega} w_j \frac{\partial u_i}{\partial x_j} v_i \, dx. \quad (103)
\]

- We have already stated that the bilinear forms $a(\cdot, \cdot)$ and $b(\cdot, \cdot)$ are continuous.

- The next two propositions give useful properties of the trilinear form $c(\cdot; \cdot, \cdot)$. 
Mathematical results

- We have the following results:

**Proposition 8** The trilinear form $c(\cdot;\cdot,\cdot)$ is continuous on $(H^1(\Omega))^d$.

- For any given $u \in V$, the map $v \mapsto c(u,u,v)$ is linear and continuous on $V$.

**Proposition 9** Let $u,v,w$ in $H^1(\Omega)$ and let assume $\text{div} \ w = 0$ and $w \cdot n|_{\partial\Omega} = 0$. Then, the trilinear form $c : (H^1(\Omega))^3 \to \mathbb{R}$ satisfies the following properties, which are equivalent:

\[
\begin{align*}
c(w;v,v) &= 0, \\
c(w;u,v) + c(w;v,u) &= 0.
\end{align*}
\]

- Assume the constant density $\rho = 1$. We can recast the problem (99): given $f \in L^2(\Omega)$, find $u \in V, p \in Q$ such that

\[
\begin{cases}
a(u,v) + c(u;u,v) + b(v,p) = (f,v), & \text{for all } v \in V \\
b(u,q) = 0, & \text{for all } q \in Q.
\end{cases}
\]

A pair of function $(u \in V, p \in Q)$ satisfying these equations, given $f \in L^2(\Omega)$, is a weak solution of the stationary Navier-Stokes problem.
Mathematical results

- Let consider the following divergence-free functional subspaces:

\[ V_{\text{div}} = \{ v \in H^1(\Omega) ; \text{div} \, v = 0 \}, \quad \text{and} \quad V_{\text{div}}^0 = \{ v \in V_{\text{div}} ; v = 0 \text{ on } \Gamma_D \}, \]

- Problem (106) has an equivalent form:

\[
\begin{cases}
\text{find} \ (u, p) \in V_{\text{div}}^0 \times L^2_0(\Omega) \text{ such that } \\
a(u, v) + c(u; u, v) + b(v, p) = (f, v), \quad \text{for all } v \in H^1_0(\Omega).
\end{cases}
\]  

(107)

If \((u, p)\) is a solution to problem (106), then \(u\) is a solution to problem (107). The converse is also true as stated next.

- **Lemma 7** Let \(\Omega\) be a bounded domain of \(\mathbb{R}^d\) with a Lipschitz-continuous boundary \(\partial \Omega\). Given \(f \in L^2(\Omega)\), there exists at least one pair \((u, p) \in V_{\text{div}} \times L^2_0(\Omega)\) which satisfies problem (107) or equivalently problem (106).
Mathematical results

• Let consider the norm of the trilinear form \( (c; \cdot, \cdot) \) of \( V^3 \) and recall the norm of \( V' \):

\[
\sup_{u,v,w \in V} \frac{c(w; u, v)}{|u|_{H^1(\Omega)} |v|_{H^1(\Omega)} |w|_{H^1(\Omega)}} = C, \quad \|f\|_{V'} = \sup_{v \in V} \frac{(f, v)}{|v|_{H^1(\Omega)}}.
\]

• The next result gives the uniqueness of a weak solution \((u, p)\) to the steady problem (107).

**Theorem 12** Under the hypothesis of the previous lemma, and assuming that

\[
\frac{C}{v^2} \|f\|_{V'} \leq 1,
\]

then, problem (107) has a unique solution \((u, p)\) in \( V \times L^2_0(\Omega) \).
Boundary conditions

- From the numerical point of view, we have to consider two types of boundary conditions:
  1. **artificial conditions**: a virtual boundary is considered, e.g., by introducing a wall to account for the symmetry of the flow and reduce the size of half;
  2. **physical conditions**: a boundary exists in the domain and must be treated accordingly, e.g., no-slip condition at wall.

- The **homogeneous Dirichlet boundary condition** \( u = 0 \), (no-slip), is the correct and natural condition for a viscous fluid contained in a rigid domain. Indeed, the viscous effect constrains the fluid particles to adhere to the wall.

- When Dirichlet conditions \( u = g \) are specified on the whole boundary \( \partial \Omega \), the pressure solution is known up to a hydrostatic constant only and then the function \( g \) has to satisfy a **compatibility condition**:

\[
\int_{\partial \Omega} g \cdot n \, ds = 0.
\]
Boundary conditions

- For specific applications or in different situations, more appropriate boundary conditions need to be worked out.

- The problem of well-posed boundary conditions is thus a key issue in many fields.
  - Slip boundary condition: the influence of the wall rugosity on the slip behavior of the fluid is still a matter of debate. In case of impermeable boundary, there is a classical condition for the normal velocity component:
    \[ u \cdot n = 0, \quad \text{on } \partial \Omega \]
  - For the tangential behavior however, the situation is less clear. A complete slip boundary condition has been proposed:
    \[ u \cdot n = [D(u) \cdot n]_{\text{tan}} = 0, \quad \text{on } \partial \Omega, \]
    which means that the fluid is confined in the domain but the fluid particle can slip along the boundary.
Boundary conditions

- Additional boundary conditions have been proposed:
  - **Neumann non-friction boundary conditions**: used to prescribe a force per unit area that involves the normal component of the stress tensor $\sigma$:
    \[
    \sigma n = (2\mu D(u) - p)n = g_1, \quad \text{on } \Gamma_N, 
    \]
    where $\Gamma_N$ is a subset of $\partial \Omega$.
    When $g_1 = 0$, this condition is usually used as an *free outflow* condition on a fictitious boundary, e.g. symmetry wall, or at the interfaces, e.g. free surfaces or between two fluids. In this case, the part of the boundary $\Gamma_N$ is called a *free outflow*. In other situations, $g_1 = -p_e n$, where $p_e$ denotes the external pressure.
  
  - alternate boundary conditions and formulations for the viscous term $-\nu \Delta u$: when the coefficient $\nu$ is constant, the viscous term can be rewritten in different equivalent forms:
    \[
    \nu \Delta u = \text{div}(\nu(\nabla u + (\nabla u)^t)) 
    = -\nu \text{curl(curl } u) = \nu(\nabla(\text{div } u) - \text{curl(curl } u)). 
    \]
Boundary conditions

- Suppose $\Gamma_n$ and $\Gamma_\tau$ are two subsets of $\partial \Omega$, such that $\bar{\Gamma}_n \cap \bar{\Gamma}_\tau = \emptyset$, $\Gamma_n \cup \Gamma_\tau = \partial \Omega$.

- We want to prescribe the normal and the tangent velocities on $\Gamma_n$ and $\Gamma_\tau$, respectively:

  \[ u \cdot n = g_n, \quad \text{on } \Gamma_n \quad \text{and} \quad n \times (u \times n) = g_\tau, \quad \text{on } \Gamma_\tau, \]  
  \[ (110) \]

  where $n$ denotes the outward normal vector to the domain boundary, $u \cdot n$ is the component of $u$ normal to the boundary and $n \times u \times n = u - (u \cdot n)n$ represents the projection of $u$ onto the tangent (plane) to the boundary.

- We introduce the functional spaces:

  \[ V_g = \{ v \in H^1(\Omega); \; v \cdot n = g_n \text{ on } \Gamma_n, \; n \times v \times n = g_\tau \text{ on } \Gamma_\tau \} \]

  \[ V_0 = \{ v \in H^1(\Omega); \; v \cdot n = 0 \text{ on } \Gamma_n, \; v \times n = 0 \text{ on } \Gamma_\tau \] 

  and we consider here $Q = L^2_0(\Omega)$ if $\Gamma_n = \partial \Omega$) or $Q = L^2(\Omega)$ otherwise.
Boundary conditions

- The weak formulation of problem (97) reads:

\[
\begin{cases}
\text{find } (u, p) \in V_g \times Q \text{ such that} \\
a(u, v) + c(u; u, v) + b(v, p) = (f, v) + d(v), \quad \text{for all } v \in V_0 \\
b(u, q) = 0, \quad \text{for all } q \in Q
\end{cases}
\]

(111)

Here, we have introduced in the right-hand side a linear functional \(d(\cdot)\) defined as:

\[
d(v) = \int_{\partial \Omega \setminus \Gamma_n} r v \cdot n \, ds + \int_{\partial \Omega \setminus \Gamma_\tau} s \cdot v \times n \, ds,
\]

where the functions \(r\) and \(s\) have a physical meaning.

- Taking for \(a(\cdot, \cdot)\) the form (101), the natural boundary conditions become

\[
-p + \nu n \cdot \nabla u \cdot n = r, \quad \text{on } \partial \Omega \setminus \Gamma_n,
\]

(112)

\[
\nu n \cdot \nabla u \times n = s, \quad \text{on } \partial \Omega \setminus \Gamma_\tau
\]

(113)
Boundary conditions

- As such, these conditions have no physical meaning, and the boundary conditions (110) must be imposed on all boundaries.

- For more practical conditions, e.g. for free-surface problems or artificial outflow boundaries, one may prefer dealing with the formulation (108) for the viscous term associated with the bilinear form:

  \[ a(u, v) = \frac{1}{2} \int_{\Omega} \nu D(u) : D(v) \, dx, \quad \text{where} \quad 2D(u) = \nabla u + (\nabla u)^t. \]

And the natural boundary conditions write:

\[ -p + 2\nu n \cdot D(u) \cdot n = r, \quad \text{on} \ \partial \Omega \setminus \Gamma_n, \quad (114) \]

\[ 2\nu n \cdot D(u) \times n = s, \quad \text{on} \ \partial \Omega \setminus \Gamma_\tau \quad (115) \]

where \( r \) and \( s \) are then the normal and tangent stresses, respectively.
Non-linear iterative procedures

**Model:** the main difference with Stokes equations is the presence of non-linear terms in the (stationary) Navier-Stokes equations.

- All methods to solve incompressible flows are based on the linearization of the equations and most of them invoke an iterative procedure.

- Given an initial guess, \( u_0 \in H^1(\Omega) \), a sequence of iterates \( \{u_k\}_{k=1,...,n} \in H^1(\Omega) \) is computed which is expected to converge towards the solution of the weak formulation.

- The fixed-point (Picard) linearization of the stationary Navier-Stokes equations (97) leads to a sequence of so-called Oseen problems, *i.e.*, linear problems (endowed with boundary conditions (97)) of the form

\[
\begin{cases}
(w \cdot \nabla)u - \nu \Delta u + \nabla p = f, & \text{in } \Omega, \\
\text{div } u = 0, & \text{in } \Omega \\
\text{div } u = 0, & \text{on } \partial \Omega
\end{cases}
\quad (116)
\]

where \( w \in V \) is the approximation of the solution from the previous Picard iteration.
Non-linear iterative procedures

- The corresponding weak formulation of this problem is:

\[
\begin{aligned}
given \ w \in V, \ find \ u \in V \text{ solving} \\
\ a(u, v) + c(w, u, v) = (f, v), \quad \text{for all } v \in V.
\end{aligned}
\]

- In the Picard sequence, the convective terms \((u \cdot \nabla)u\) are approximated by:

\[
\begin{aligned}
((u \cdot \nabla)u)_k^{+1} &\approx (u_k \cdot \nabla)u_k^{+1}, \\
((u \cdot \nabla)u)_k^{+1} &\approx (u_{k+1} \cdot \nabla)u_k, \\
((u \cdot \nabla)u)_k^{+1} &\approx (u_k \cdot \nabla)u_k,
\end{aligned}
\]

- This leads to solve the following sequence of problems: for \( k = 0, 1, 2, \ldots \)

\[
\begin{aligned}
given \ u_k \in V, \ find \ u_k^{+1} \in V \text{ solving} \\
\ a(u_k^{+1}, v) + c(u_k, u_k^{+1}, v) = (f, v), \quad \text{for all } v \in V
\end{aligned}
\]

or

\[
\int_{\Omega} \nu \nabla u_k^{+1} \cdot \nabla v \, dx + \int_{\Omega} (u_k \cdot \nabla)u_k^{+1} \cdot v \, dx = \int_{\Omega} f v \, dx,
\]

- At each iteration step \( k \), one linear system (116) has now to be solved.
- Lax-Milgram lemma gives the existence and uniqueness of a weak solution to this system.
The rate of convergence of Oseen method is linear.

The Newton method seems appealing because of its theoretical potential to achieve quadratic convergence rate.

Consider the abstract problem of finding the root of a general nonlinear function \( f, f(x) = 0 \).

Suppose \( x_k \) is an approximate solution of the root \( \tilde{x} \).

Then, by a Taylor expansion around an initial guess \( x_k \) we write

\[
0 = f(\tilde{x}) = f(x_k) + f'(x_k)(\tilde{x} - x_k) + o(|\tilde{x} - x_k|) \\
\approx f(x_k) + f'(x_k)(\tilde{x} - x_k),
\]

and thus

\[
\tilde{x} \approx x_k - [f'(x_k)]^{-1} f(x_k).
\]

Newton’s method for solving \( f(x) = 0 \) writes then

\[
x_{k+1} = x_k - [f'(x_k)]^{-1} f(x_k), \quad k = 0, 1, 2, \ldots
\]
Newton’s method

• Likewise, to deal with the nonlinear term of stationary Navier-Stokes model, let define the function

\[ f(u, \nabla u) = (u \cdot \nabla) u. \]

• In this case, a Taylor expansion of \( f(\cdot, \nabla \cdot) \) around \( u_k, \nabla u_k \) reads

\[ f_{k+1}(u, \nabla u) \approx f_k(u_k, \nabla u_k) + (u_{k+1} - u_k) \cdot \frac{\partial f_k}{\partial u} + \nabla(u_{k+1} - u_k) \cdot \frac{\partial f_k}{\partial \nabla u}. \]

• Neglecting the quadratic terms and substituting the definition of \( f(\cdot, \nabla \cdot) \) yields

\[ u_{k+1} \cdot \nabla u_{k+1} \approx u_k \cdot \nabla u_k + \delta_k \cdot \nabla u_k + \nabla \delta_k \cdot u_k \]

\[ = u_{k+1} \cdot \nabla u_k + u_k \cdot \nabla u_{k+1} - u_k \cdot \nabla u_k, \tag{119} \]

which forms the kernel of the classical Newton linearization of the nonlinear term.

• The convergence rate is quadratic; if \( \|u - u_0\| \leq \delta \)

\[ \|u_{k+1} - u\| \leq M\|u_k - u\|^2, \text{ and } \|u_k - u\| \leq \frac{(M\delta)^{2k}}{M}, \]
Newton’s method

- Newton’s method applied to the weak formulation of the stationary Navier-Stokes problem and using identity (119) takes the following form:

  for \( k = 0, 1, 2, \ldots \)
given \( u_k \in V \), find \( u_{k+1} \) solving for all \( v \in V \)

\[
a(u_{k+1}, v) + c(u_k, u_{k+1}, v) + c(u_{k+1}, u_k, v) = (f, v) + c(u_k, u_k, v)
\]
or

\[
\int_{\Omega} \nu \nabla u_{k+1} \cdot \nabla v \, dx + \int_{\Omega} (u_k \cdot \nabla) u_{k+1} \cdot v \, dx + \int_{\Omega} (u_{k+1} \cdot \nabla) u_k \cdot v \, dx \\
= \int_{\Omega} f \, v \, dx + \int_{\Omega} (u_k \cdot \nabla) u_k \cdot v \, dx.
\]

- Another important aspect for Newton-type methods is the starting value, since one only obtains local convergence in general.

- One possibility is to use some steps of an Oseen-iteration first, and to use the result as a starting value for a Newton method.
Section 3.2

Discretization procedures
Unsteady Navier-Stokes equations

- We return to the unsteady Navier-Stokes equations:

\[
\begin{aligned}
\begin{cases}
\text{Given } f, \text{ find } u \text{ such that } \\
\frac{\partial u}{\partial t} + (u \cdot \nabla) u - \nu \Delta u + \frac{1}{\rho} \nabla p = f, \quad \text{in } \Omega_t \times (0, T), \\
\text{div } u = 0, \quad \text{in } \Omega_t \times (0, T), \\
u = u_0, \quad \text{in } \Omega_t \times (0, T), \quad \text{with } \text{div } u_0 = 0,
\end{cases}
\end{aligned}
\tag{121}
\]

deeded with the conditions (to be well-posed):

\[
u = g, \quad \text{on } \partial \Omega \times (0, T) \quad \text{with } \int_{\partial \Omega} g(t) \cdot n \, ds = 0, \quad \text{on } (0, T) \tag{122}
\]

- Then, we have for almost every time \( t \) in \( (0, T) \)

\[
\int_{\Omega_t} \frac{\partial u}{\partial t} \cdot v \, dx + \int_{\Omega_t} ((u \cdot \nabla) u) \cdot v \, dx + \int_{\Omega_t} \nu \nabla u : \nabla v \, dx - \frac{1}{\rho} \int_{\Omega_t} p \text{div } v \, dx \\
= \int_{\Omega_t} f \cdot v \, dx, \quad \forall v \in V \tag{123}
\]

completed with the Dirichlet boundary condition on \( \partial \Omega \times (0, T) \).
Discretization in space

- We perform the space discretization of the NS system as previously, to find \((u_h, p_h)\) in \(V_h \times Q_h\) for all \(t \in (0, T)\) such that:

\[
\begin{align*}
\int_{\Omega_t} \frac{\partial u_h}{\partial t} \cdot v_h \, dx + \int_{\Omega_t} ((u_h \cdot \nabla) u_h) \cdot v_h \, dx + \int_{\Omega_t} \nu \nabla u_h : \nabla v_h \, dx \\
- \frac{1}{\rho} \int_{\Omega_t} p_h \operatorname{div} v_h \, dx = \int_{\Omega_t} f_h \cdot v_h \, dx, \quad \forall v_h \in V_0h \\
\int_{\Omega_t} \operatorname{div} u_h q_h \, dx = 0 \quad \forall q_h \in Q_h \\
u_h = g_h(t) \quad \text{on } \partial \Omega_t \\
u_h(0) = u_{0h}
\end{align*}
\]

- The finite element spaces \(V_h\) and \(Q_h\) are classically

\[
V_{0h} = V_h \cap \left( H^1_0(\Omega_t) \right)^2 = \{ v_h \in V_h ; \; v_h = 0 \text{ on } \partial \Omega_t \},
\]

- The function \(f_h, u_{0h}\) and \(g_h\) are convenient approximations of \(f, u_0\) and \(g\) and \(g_h\) must verify

\[
\int_{\partial \Omega_t} g_h(t) \cdot n \, ds = 0 \quad \text{on } (0, T).
\]
Operator splitting

We introduce a \( \theta \)-scheme (Glowinski, 1985) to discretize an initial value problem:

\[
\frac{\partial u}{\partial t} + A(u, t) = 0, \quad \text{with } u(0) = u_0.
\]

- Let \( \theta \) be a number in the open interval \((0, 1/2)\),
- the scheme applied to the solution of this problem, when \( A = A_1 + A_2 \), solves \( u^{n+\theta} \), \( u^{n+1-\theta} \) and \( u^{n+1} \)
- it writes:

\[
\begin{align*}
  u^0 &= u_0 \\
  \frac{u^{n+\theta} - u^n}{\theta \delta t} + A_1(u^{n+\theta}, (n + \theta)\delta t) + A_2(u^n, n\delta t) &= 0 \\
  \frac{u^{n+1-\theta} - u^{n+\theta}}{(1 - 2\theta)\delta t} + A_1(u^{n+\theta}, (n + \theta)\delta t) + A_2(u^{n+1-\theta}, (n + 1 - \theta)\delta t) &= 0 \\
  \frac{u^{n+1} - u^{n+1-\theta}}{\delta t} + A_1(u^{n+1}, (n + 1)\delta t) + A_2(u^{n+1-\theta}, (n + 1 - \theta)\delta t) &= 0
\end{align*}
\]
The Navier-Stokes model

Time discretization procedures

- The full discretization of unsteady Navier-Stokes equations by a $\theta$-scheme leads to the following systems:
  - initializations and parameter settings:
    \[
    \begin{aligned}
    u_h^0 &= u_{0h} \\
    \theta &= 1 - 1/\sqrt{2}, \quad \alpha = (1 - 2\theta)/(1 - \theta), \quad \beta = \theta/(1 - \theta),
    \end{aligned}
    \]
  - and then, $u_h^n$ being know,
  - compute first $(u_h^{n+\theta}, p_h^{n+\theta}) \in V_h \times Q_h$ by solving the elliptic systems, for $n \geq 0$:
    \[
    \forall v_h \in V_0h, \quad \begin{cases}
    \int_{\Omega_t} \frac{u_h^{n+\theta} - u_h^n}{\theta \delta t} \cdot v_h \, dx + \alpha \nu \int_{\Omega_t} \nabla u_h^{n+\theta} : \nabla v_h \, dx - \int_{\Omega_t} p_h^{n+\theta} \text{div} v_h \, dx \\
    = \int_{\Omega_t} f_h^{n+\theta} \cdot v_h \, dx - \beta \nu \int_{\Omega_t} \nabla u_h^n : \nabla v_h \, dx - \int_{\Omega_t} (u_h^n \cdot \nabla) u_h^n \cdot v_h \, dx \\
    \int_{\Omega_t} \text{div} u_h^{n+\theta} q_h = 0, \quad \forall q_h \in Q_h \\
    u_h^{n+\theta} = g_h^{n+\theta} \text{ on } \partial \Omega_t
    \end{cases}
    \]
Time discretization procedures

- compute $u_h^{n+1-\theta} \in V_h$, for all $v_h \in V_0h$

$$
\begin{align*}
&\int_{\Omega_t} \frac{u_h^{n+1-\theta} - u_h^{n+\theta}}{(1 - 2\theta)\delta t} \cdot v_h \, dx + \beta \nu \int_{\Omega_t} \nabla u_h^{n+1-\theta} : \nabla v_h \, dx \\
&\quad + \int_{\Omega_t} (u_h^{n+1-\theta} \cdot \nabla) u_h^{n+1-\theta} \cdot v_h \, dx \\
&= \int_{\Omega_t} f_h^{n+\theta} \cdot v_h \, dx - \alpha \nu \int_{\Omega_t} \nabla u_h^{n+\theta} : \nabla v_h \, dx - \int_{\Omega_t} p_h^{n+\theta} \text{div} \, v_h \, dx \\
&u_h^{n+1-\theta} = g_h^{n+1-\theta} \quad \text{on } \partial \Omega_t,
\end{align*}
$$

- and finally compute $(u_h^{n+1}, p_h^{n+1}) \in V_h \times Q_h$

$$
\begin{align*}
&\int_{\Omega_t} \frac{u_h^{n+1} - u_h}{(\theta)\delta t} \cdot v_h \, dx + \alpha \nu \int_{\Omega_t} \nabla u_h^{n+1} : \nabla v_h \, dx - \int_{\Omega_t} p_h^{n+1} \text{div} \, v_h \, dx \\
&= \int_{\Omega_t} f_h^{n+1} \cdot v_h \, dx - \beta \nu \int_{\Omega_t} \nabla u_h^{n+1-\theta} : \nabla v_h \, dx \\
&\quad - \int_{\Omega_t} (u_h^{n+1-\theta} \cdot \nabla) u_h^{n+1-\theta} \cdot v_h \, dx \\
&\int_{\Omega_t} \text{div} \, u_h^{n+1} \, q_h \, dx = 0 \quad \forall q_h \in Q_h \\
&u_h^{n+1} = g_h^{n+1} \quad \text{on } \partial \Omega_t,
\end{align*}
$$
The Navier-Stokes model

Time discretization procedures

- Until now, the term \((u \cdot \nabla)u\) was treated as a general non-linearity;
- However, in Navier-Stokes equations, the term

\[
\rho \left( \frac{\partial u}{\partial t} + (u \cdot \nabla)u \right)
\]

models the transport of the momentum \(\rho u\) by the velocity field \(u\).

- The idea of combining FE approximations and methods of characteristics for numerical simulations of incompressible viscous fluid flows is quite old (1970’s): J.P. Benque, O. Pironneau, Gresho, . . . but still considered a bit "exotic".

- Despite the simplicity of the principle that underlies the methods of characteristics, their practical implementation is delicate and not fully understood;

- "Any method for approximating hyperbolic equations sacrifices a good deal if it takes no account of the method of characteristics" (Morton, 1992), "Numerical schemes that follow characteristics backward in time and then interpolate at their feet have a history stretching back in the very early days of computational fluid dynamics" (Courant et al., 1952).
Solution of transport problems

• We introduce the method of characteristics to solve a transport problem:

\[
\begin{cases}
\frac{\partial \varphi}{\partial t} + u \cdot \nabla \varphi = 0 \quad \text{in } \Omega \times (0, T) \\
\varphi(0) = \varphi_0 \\
\varphi = g \quad \text{on } \Gamma_- \times (0, T)
\end{cases}
\]

with

\[
\frac{\partial u}{\partial t} = 0 \quad \text{div } u = 0 \quad \frac{\partial g}{\partial t} = 0
\]

and

\[
\Gamma_- = \{ x \in \partial \Omega, \ u(x) \cdot n(x) < 0 \}.
\]

• Let consider \((x_*, t_*) \in \Omega \times (0, T)\); we associate the solution of the system

\[
\begin{cases}
\frac{dX}{dt} = u(x) \\
X(t_*) = x_*
\end{cases}
\]

the solution of this system is \(X(\cdot; x_*, t_*)\)
The Navier-Stokes model

Solution of transport problems

• The curve $C_{x*,t*}$ described by the points $\{X(t; x_*, t_*), t\}$ as $t$ varies is called the characteristic curve associated to the transport equation (124) and to $\{x_*, t_*\}$.

• Suppose $\varphi$ is solution of (124) and restrict it to the curve $C_{x*,t*}$, we have then

$$\frac{d}{dt} \varphi(X(t; x_*, t_*), t) = \left(\frac{\partial \varphi}{\partial t} + u \cdot \nabla \varphi\right) (X(t; x_*, t_*), t) = 0.$$

• this shows that $\varphi$ is constant along the curve $C_{x*,t*}$.

• it follows that if $\varphi$ solve (124), we have, for $\tau > 0$ sufficiently small, the relation

$$\varphi(x, t) = \varphi(X(t - \tau; x, t), t - \tau).$$
The Navier-Stokes model

Time discretization procedures

- The backward method of characteristics can be used to solve Navier-Stokes problem:
  - initialization
    \[ u^0 = u_0 \]
  - and for \( n \geq 0 \):
    \[
    \begin{cases}
    \frac{u^{n+1} - u_n^*}{\delta t} - \nu \Delta u^{n+1} + \nabla p^{n+1} = f^{n+1} & \text{in } \Omega_t \\
    \text{div } u^{n+1} = 0, & \text{in } \Omega_t \\
    u^{n+1} = g^{n+1} & \text{on } \partial \Omega_t ;
    \end{cases}
    \]
  - where \( u_n^* \) is obtained by solving the system:
    \[
    \begin{cases}
    \frac{dX}{dt} = u^n(X) & \text{on } (0, \delta t) \\
    X(\delta t) = x
    \end{cases}
    \]
    we denote by \( X^n(\cdot, x) \) the solution of the problem (126).
  - We take \( u_n^*(x) = u^n(X^n(0, x)) \)
Section 3.3
Numerical examples
2d lid-driven cavity

- classical test case: simulations for Reynolds number ranging from 400 up to 10000.
  - problem involves a primary vortex at the cavity center and vortices in the corners,
  - number of vortices increases with $Re$ and position of the center of primary vortex moves toward the center,
  - 4 meshes have been used: a regular triangulation (2,461 nodes, 5,000 elements); a uniform triangulation (2,143 nodes, 4,136 elements); a refined uniform triangulation (8,421 nodes, 16,544 elements); and a regular triangulation (10,201 nodes, 20,000 elements);
  - results in good accordance with experimental results (Ghia et al.).
Section 4.1
Modelling
Problem statement

The numerical modelling and resolution of bifluid problems

- is needed because of small time and length scales, reliable experiments are impossible,
- investigate and understand physical phenomena,
- requires the accurate discretization and the tracking of the interface separating two immiscible fluids.

The major challenges are related to
- the evolution of the interface and
- the induced changes of its geometry and topology.

Here, we consider the dynamics of interface deformation in low Reynolds number flow. This topic has interest in wide variety of fields: chemical and petroleum engineering, geophysics, biology, ...
Simulation of bifluid flows

Several difficulties may jeopardize the resolution:

- large jumps of viscosity and density between the fluids must be properly taken into account and resolved to satisfy momentum balance,
- mass conservation is especially important in interfacial flows,
- the surface tension force must be considered in the model and accurately evaluated,
- the resolution of the interface must be preserved at all stages, even in the extreme cases of folding, merging and breaking.
Since the seminal work of Harlow and Welch (1965), numerous methods have been proposed. Comprehensive surveys written by:


Algorithms for fluid flows are subdivided into two classes, namely Lagrangian and Eulerian approaches.
Lagrangian vs. Eulerian approaches

1. **Lagrangian methods: interface tracking**
   - follow the interface evolution using a set of markers and deform the grid.
   - each grid cell contains the same fluid part throughout the whole computation.
   - face difficulties in handling markers when the interface becomes highly stretched or distorted, and when the topology changes.

2. **Eulerian techniques: interface capturing**
   - introduce a scalar valued level set function to define the interface manifold,
   - fixed coordinate system, the fluid travels from one grid cell to another.
   - topology changes easily handled, but mass conservation may be a real concern.
   - expressions of the interface normal and curvature from the level set function.

Interface capturing

The numerical resolution strategy is

- set in the context of Eulerian and interface-capturing methods.
- involving an important feature: mesh adaptation using unstructured (anisotropic) triangulations.

Our choice is motivated by the following arguments:

1. necessity to deal with complex interfacial motions and topology changes;
2. flow resolution decoupled from the advection part;
3. anisotropic mesh adaptation for accurate representation of the interface with a minimal number of unknowns;
4. bifluid resolution allows large viscosity ratios;
5. the advection term treated by the method of characteristics combined with a Galerkin FE scheme.

Furthermore, fluid coalescence and detachment can be efficiently treated with the sharp interface definition.
Hypothesis and notations

• Suppose $\Omega$ is an open bounded computational domain in $\mathbb{R}^d$, the outer boundary is denoted $\Sigma$,

• the subdomains denoted $\Omega^1(t)$ and $\Omega^2(t)$, $\partial \Omega^i(t)$ is the boundary of $\Omega^i(t)$,

• the interface between the fluids by: $\Gamma(t) = \partial \Omega^1(t) \cap \partial \Omega^2(t)$.

• suppose also that: $\overline{\Omega^1(t)} \cup \overline{\Omega^2(t)} = \overline{\Omega}$ and $\Omega^1(t) \cap \Omega^2(t) = \emptyset$.

• the domains $\Omega^i$ can have several connected components, and the interface $\Gamma(t)$ can possibly intersect the outer boundary $\Sigma$.

\[
\Omega^1(t) \
\Omega^2(t) \\
\Gamma(t) \\
\Sigma
\]

\[
\text{Notice that there is no temporal derivative in this equation. Physically however, this does not mean that the flow is steady. This only reflects that the forces exerted on the fluid are in a state of dynamic equilibrium as a result of a rapid diffusion of the momentum. Hence, the transient character of the solution is related to the motion of the two fluids and of the interface.}

We have assumed that the surface tension effect must be taken into account at the interface. Therefore, this system is endowed with conditions on the continuity of the velocity and on the balance of the normal stress with the surface tension across the interface [GLM06]:

\[
\begin{align*}
\mathbf{u}^1 - \mathbf{u}^2 &= 0 \\
\frac{\sigma^1 - \sigma^2}{\mathbf{n}^1} &= -\gamma \kappa \mathbf{n}^1
\end{align*}
\]

where

– $\sigma^i = \mu \left( \nabla \mathbf{u}^i + (\nabla \mathbf{u}^i)^T \right) - p \mathbf{I}$ denotes the stress tensor,

– $\mathbf{n}^1$ is the unit exterior normal vector to $\Gamma(t)$ of $\Omega^1(t)$ pointing from $\Omega^1(t)$ to $\Omega^2(t)$ (we assume that $\Gamma(t)$ is sufficiently smooth),

– $\gamma > 0$ is the surface tension coefficient assumed to be constant along the interface,

– $\kappa$ is the signed mean curvature of the interface, being positive if the interface curve/surface bends towards $\Omega^1(t)$ and negative otherwise.

These equations are completed with some appropriate boundary conditions on the outer boundary $\Sigma$. We can set the classical Dirichlet, Neumann or mixed...
Model equations

- we introduce two scalar functions $\mu$ and $\rho$ for the viscosity and density of the fluid defined on the whole domain $\Omega$ as follows:

$$
\mu = \chi^1 \mu^1 + \chi^2 \mu^2, \quad \rho = \chi^1 \rho^1 + \chi^2 \rho^2
$$

where $\chi^i$ is the characteristic function of the domain $\Omega^i$, $\mu^i$ and $\rho^i$ are dynamic viscosities and densities of each fluid ($i=1,2$), respectively.

- we also assume that $\rho(x,t) = \rho(x)$ and $\mu(x,t) = \mu(x)$.

The flow of the incompressible viscous fluid is governed by the Navier-Stokes equations:

$$
\begin{cases}
\rho \left( \frac{\partial u}{\partial t} + (u \cdot \nabla)u \right) - \mu \Delta u + \nabla p = \rho f & \text{in each } \Omega^i \ (i = 1, 2) \\
\text{div } u = 0 & \text{in } \Omega
\end{cases}
$$

(127)

where $(u \cdot \nabla)u = \sum_{i=1}^d u_i \partial_i u$ and

- $u = u^i$, $p = p^i$ are the velocity and pressure unknown of the flow in $\Omega^i$ ($i=1,2$)
- $f$ is an internal force exerted on the fluid (e.g. gravity).
To ensure the well-posedness of the problem, Equations (127) need to be complemented by boundary conditions:

- \( u = u_D \) on \( \Sigma_D \) (Dirichlet)
- \( \sigma n = u_N \) on \( \Sigma_N \) (Newmann)
- \( u.n = 0 \), or \( \alpha u.\tau + \tau.\sigma n = 0 \) on \( \Sigma_S \) (Slip)

assuming \( \Sigma = \Sigma_D \cup \Sigma_N \cup \Sigma_S \) is split into a finite number of components corresponding to different types of boundary conditions,

- \( \tau, n \) are unit tangent vector and unit exterior normal vector to \( \Gamma \) and

- \( \sigma \) is the stress tensor defined as:

\[
\sigma = \mu \left( \nabla u + t^\nabla u \right) - p \text{Id}.
\]
Boundary conditions: numerical issues

Note: in practice,

- the slip condition can be imposed by adding the integral
  \[ \int_{\Sigma_S} \alpha(u.\tau).(v.\tau)ds \]
  in the left hand side of the variational formulation of problems (see after) and

- the condition \( u.n|_{\Sigma_S} = 0 \) can be treated as a Dirichlet condition on \( u \).

- in case of \( \alpha = 0 \) (slip without friction), this integral term vanishes and we can implement the condition \( u.n|_{\Sigma_S} = 0 \) generally by adding
  \[ \int_{\Sigma_S} A \ast (u.n).(v.n)ds \]
  in the left hand side and multiplying the components on \( \Sigma_S \) of right hand side with \( A \) where \( A \) is so-called penalization number (values about \( 10^6 \)).
Initial and interfacial condition

- **initial condition**: divergence-free velocity field \( u_0(x) \) specified over the domain \( \Omega \) at time \( t = 0 \), i.e.
  \[
  u(x, 0) = u_0(x).
  \]

- at the **interface** \( \Gamma \), two conditions are imposed:
  \[
  \begin{align*}
  u^1 - u^2 &= 0 \\
  (\sigma^1 - \sigma^2) \cdot n^1 &= -\gamma \kappa n^1
  \end{align*}
  \]
  where
  - \( n^1 \) is the unit normal vector to \( \Gamma(t) \), exterior to \( \Omega^1(t) \),
  - \( \gamma > 0 \) is the constant surface tension coefficient along the interface,
  - \( \kappa \) is the algebraic mean curvature of the interface, being positive if the interface curve/surface bends towards \( \Omega^1 \) and negative otherwise.
Section 4.2

Evolution of the interface
level set formalism
Auxiliary function

The first problem we face is to describe the interface $\Gamma$ in proper way. Following [Set99], we introduce an auxiliary function, level set function, defined as the signed distance function to the interface $\Gamma$, i.e.,

$$
\phi(x) = \pm d(x, \Gamma),
$$

(129)

and we have

$$
\begin{cases}
\phi(x) > 0, & \text{if } x \in \Omega^1, \\
\phi(x) < 0, & \text{if } x \in \Omega^2, \\
\phi(x) = 0, & \text{if } x \in \Gamma.
\end{cases}
$$


Signed distance function to the interface $\Gamma$. 
Level sets and interface capturing

• the evolution of the interface may induce geometry and topology changes.

• fortunately, with the level set formulation, at each time step $t$, the fluid interface $\Gamma$ is associated with the zero isocontour of the continuous function $\phi$:

$$\Gamma(t) = \{x \in \Omega : \phi(x, t) = 0\}, \quad \phi(x, t) = \pm \min_{y \in \Gamma(t)} \|x - y\|. \quad (130)$$

• the interface $\Gamma$ is solution, at each $t$, of the advection equation:

$$\begin{cases}
\frac{\partial \phi}{\partial t}(x, t) + \tilde{u}(x, t) \cdot \nabla \phi(x, t) = 0, & \forall (x, t) \in \Omega \times \mathbb{R}^+ \\
\phi(x, 0) = \phi_0(x), & \forall x \in \Omega
\end{cases} \quad (131)$$

where $\phi_0(x)$ is the signed distance function to $\Gamma^0$ and $\tilde{u}$ is the vector field defined from the velocity field along the interface (solving the fluid equations).
Definition of the velocity field

• the evolution of the interface depends only on the flow field in its vicinity, and not on the whole domain $\Omega$.

• moreover, in case of complex displacements, sharp velocity variations may cause uncontrolled oscillations and jeopardize the numerical stability of subsequent algorithms.

• hence, we extend and regularize the velocity of flows $u$ (taken only along $\Gamma$) into a vector field $\tilde{u}$ (defined on $\Omega$) before solving the advection:

$$
\begin{align*}
-\alpha \Delta \tilde{u} + \tilde{u} &= 0 \quad \text{in } \Omega \\
\tilde{u} &= 0 \quad \text{on } \Sigma \\
\tilde{u} &= u \quad \text{on } \Gamma
\end{align*}
$$

(132)

where small $\alpha > 0$ can be interpreted as a regularization lengthscale (balance between keeping $u$ and the level of regularization).

It is a more regular inner product than $L^2(\partial \Omega)$ over functions on $\Omega$. 
Extension of the velocity field

- let $V$ be a Hilbert space which is composed of functions enjoying the desired regularity for $\tilde{u}$ (usually $V = H^1(\Omega)$),

- let $a(\cdot, \cdot)$ be a coercive bilinear form on $V$ which is close to $I$, so that $\tilde{u}$ is close to $u$:

  $$\forall \phi \psi \in V, \quad a(\phi, \psi) = \alpha \int_{\Omega} \nabla \phi \cdot \nabla \psi + \int_{\Omega} \phi \psi = \int_{\Gamma} \phi \psi,$$

  for a small $\alpha > 0$ which can be interpreted as a regularization lengthscale.

- then, $\tilde{u}$ is searched as the unique solution in $V$ to the variational problem:

  $$\forall \phi \in V, \quad a(\tilde{u}, \phi) = \int_{\Gamma} u \phi ds.$$

- this problem can easily be solved using a finite element method.
Numerical scheme

The bifluid problem can be numerically resolved using the following general scheme:

1. Initialization: level set function $\phi_0$, velocity $u_0$, mesh $T^0_h$.

2. At each time step $t^n = n\Delta t$:
   (a) solve the Navier-Stokes equations for $(u, p)^n$;
   (b) define a regularized velocity field $\tilde{u}$ (close to $u$, matching $u$ along $\Gamma$);
   (c) solve the advection equation for $\phi$ (new location of the interface);
   (d) generate a conforming mesh $T^{n+1}_h$ refined in the vicinity of $\Gamma$);
   (e) regularize $\phi$, update density $\rho$ and viscosity $\mu$;
   (f) interpolate $(u, p)^n$ onto $T^{n+1}_h$.

3. Resume step 2 until the final time is reached.

Note: this scheme is masking some complex and necessary numerical routines: solution interpolation, error estimate and mesh adaptation.
Section 4.3
Numerical resolution
Numerical methods

- The method of characteristics is known to be very efficient for solving advection-diffusion problems, including the Navier-Stokes equations.

- Here, we use this method not only for solving the advection of the interface, but also for solving the nonlinear convective term in Navier-Stokes equations.

- a Cauchy problem:
  - given an initial function $\phi^0(x) : \Omega \rightarrow \mathbb{R}$ and
  - given a velocity field $u(x, t) : \Omega \rightarrow \mathbb{R}^d$ defined on $\Omega$,
  - find $\phi(x, t) : \Omega \times [0, T] \rightarrow \mathbb{R}$ solving:

$$
\begin{aligned}
\frac{\partial \phi}{\partial t}(x, t) + u(x, t) \nabla \phi(x, t) &= 0 \quad \forall (x, t) \in \Omega \times (0, T) \\
\phi(x, 0) &= \phi^0(x) \forall x \in \Omega.
\end{aligned}
$$

(134)
Method of characteristics

The problem (134) is solved by following backward the characteristic curves of the fluid particles:

- given a particle $x \in \Omega$ at time $s$, its curve is described by the following equations:

$$
\begin{cases}
\frac{dX(x, s; t)}{dt} = u(X(x, s; t), t) \forall t \in (0, t) \\
X(x, s; s) = x
\end{cases}
$$

where $X(x, s; t)$ is the position of $x$ at the time $t$.

- the first equation of (134) implies that $\phi(x, t)$ is constant along the characteristic lines $X(x, s; t)$,

- hence the solution of the Cauchy problem (134) writes:

$$
\phi(x, t) = \phi^0(X(x, t; 0), 0) \quad \forall (x, t) \in \Omega \times [0, T]
$$
Method of characteristics: discretization

- the interval \([0, T]\) is divided into a finite number of intervals \(\Delta t\) of the form \((t^{n-1}, t^n)\) with \(t^n = n\Delta t\),

- the discretization in time of the equations (135), for all \(n\), reads:
  \[
  \begin{align*}
  \frac{dX(x, t^n; t)}{dt} &= u(X(x, t^n; t), t) \quad \forall t \in (t^{n-1}, t^n) \\
  X(x, t^n; t^n) &= x
  \end{align*}
  \]  
  (137)

- we compute only \(\phi(x, t^n)\) for all \(n\), so if denote \(\phi(x, t^n)\) by \(\phi^n(x)\), by substituting the time interval \([t^{n-1}, t^n]\) into (136) we obtain the following result:
  \[
  \phi^n(x) = \phi^{n-1}(X(x, t^n; t^{n-1})) \quad \forall x \in \Omega
  \]  
  (138)

  where \(X(x, t^n; t^{n-1})\) is the position at the time \(t^{n-1}\) of the characteristic emerging from \(x\) at the time \(t^n\).

- numerically, the expression (138) can be solved using a Lagrange interpolation, i.e \(\phi^n(x)\) is computed by taking into account the values of \(\phi^{n-1}(x)\) at the degree of freedoms of element \(K\) which contains \(X(x, t^n; t^{n-1})\).
Numerical approximation of the characteristic curves

- only an approximation \( u_h \) of \( u \) is known at the vertices of a triangulation \( T_h \),

- compute an approximation of \( X_h(x, t^n; t^{n-1}) \), solution at the time \( t^{n-1} \) of the approximated characteristic curve:

\[
\begin{align*}
\frac{dX_h(x, t^n; t)}{dt} &= u_h(X_h(x, t^n; t), t) \\
X_h(x, t^n; t^n) &= x
\end{align*}
\]  \hspace{1cm} (139)

that implies the "formal" expression:

\[
X_h(x, t^n; t) = x - \int_t^{t^n} u_h(X_h(x, t^n; t), t) \, dt
\]  \hspace{1cm} (140)

- the simplest algorithm is obtained directly from (140) as:

\[
X_h(x, t^n; t^{n-1}) = x - \Delta t u_h(x)
\]  \hspace{1cm} (141)
Numerical approximation of the characteristic curves

- the characteristic curve is considered as a straight line connecting point \( x \) and the foot \( X_h(x, t^n; t^{n-1}) \).

  in practice, given a substep \( \delta t \) in \([t^{n-1}, t^n]\), suppose there exists \( \Delta t = M\delta t \):

  (0) \( X_h(x, t^n; t^n) = x \)

  (i) Euler’s scheme: \( m = 1, \ldots, M \)

  \[
  X_h(x, t^n; t^n - m\delta t) = X_h(x, t^n; t^n - (m - 1)\delta t) - \delta t u_h(X_h(x, t^n; t^n - m\delta t));
  \]

  (ii) Runge-Kutta 4 scheme:

  \[
  X_h(x, t^n; t^n - m\delta t) = X_h(x, t^n; t^n - (m - 1)\delta t) - \frac{\delta t}{6}(v_1 + 2v_2 + 2v_3 + v_4)
  \]

  with

  \[
  \begin{align*}
  v_1 &= u_h(X_h(x, t^n; t^n - m\delta t)) \\
  v_2 &= u_h(X_h(x, t^n; t^n - m\delta t) - \frac{\delta t}{2}v_1) \\
  v_3 &= u_h(X_h(x, t^n; t^n - m\delta t) - \frac{\delta t}{2}v_2) \\
  v_4 &= u_h(X_h(x, t^n; t^n - m\delta t) - \delta tv_3).
  \end{align*}
  \]
Navier-Stokes: discretization in time

- time discretisation of the Navier-Stokes equations based on method of characteristics.

- choice motivated by the fact that this scheme is known to be unconditionally stable, and very efficient on adapted meshes.

- main idea: hide the nonlinear convective part of Navier-Stokes equations in the Cauchy problem (135),

- the operator $\frac{\partial}{\partial t} + u \cdot \nabla$ may be turned into a total derivative $\frac{d}{dt}$,

- so the equation (127) can be recast into the following form:

$$\rho \frac{du(X(x, s; t), t)}{dt} - \mu \Delta u + \nabla p = \rho f$$

(142)
Navier-Stokes: discretization in time

- time-dependent Navier-Stokes problem is now rewritten as follows, with $u^n(x) = u(x, t^n)$, in each $\Omega^i (i = 1, 2)$:

$$\begin{cases}
\rho \frac{u^n(x) - u^{n-1} \circ X^{n-1}(x)}{\Delta t} - \mu \Delta u^n(x) + \nabla p^n(x) = \rho f^n \\
\text{div } u^n(x) = 0 \text{ in } \Omega
\end{cases} \quad (143)$$

or equivalently, in each $\Omega^i (i = 1, 2)$:

$$\begin{cases}
\rho \frac{u^n(x)}{\Delta t} - \mu \Delta u^n(x) + \nabla p^n(x) = \rho f^n + \rho \frac{u^{n-1} \circ X^{n-1}(x)}{\Delta t} \\
\text{div } u^n(x) = 0 \text{ in } \Omega
\end{cases} \quad (144)$$

where $X^{n-1}(x)$ denotes $X(x, t^n; t^{n-1})$ and $u^{n-1} \circ X^{n-1}(x)$ corresponds to the velocity at this location at time $t^{n-1}$.
To summarize, the resolution consists in performing two consecutive steps:

1. approximate the characteristic curves $X^{n-1}(x)$.
2. solve the resulting generalized Stokes system.

Remarks:

- the approximation of characteristic curves $X^{n-1}(x)$ in each time interval $[t^{n-1}, t^n]$ for the Navier-Stokes problem is implemented as before, except that here $u^n$ is unknown, and thus $X^{n-1}(x)$ is associated with $u^{n-1}$ (and not $u^n$ as in the advection equation).
- it is worth to forecast some difficulties: characteristic curves may cross the boundary and thus it is needed to retain the last integration point.
Navier-Stokes: variational formulation

- in each time interval \([t^{n-1}, t^n]\), we have to solve an unsteady Stokes problem (144).
- suppressing the dependency on \(n\), the equation (144) becomes:

\[
\begin{align*}
\alpha \rho u - \mu \Delta u + \nabla p &= \rho f + \alpha \rho w \quad \text{in each } \Omega^i (i = 1, 2) \\
d\div \mathbf{u} &= 0 \quad \text{in } \Omega,
\end{align*}
\]

(145)

where \(\alpha\) denotes \(\frac{1}{\Delta t}\) and \(w\) represents \(u^{n-1} \circ X^{n-1}\).

- we set the spaces \(V = (H^1_0(\Omega))^d\) if \(\Sigma_D \equiv \Sigma\) and \(M = L^2_0(\Omega)\) for Dirichlet conditions or \(M = L^2(\Omega)\) if \(\Sigma \setminus \Sigma_D\) is not empty.

- without loss of generality, we consider here the variational formulation for Stokes problem in case of homogenous Dirichlet boundary condition, i.e. \(u|_\Gamma = 0\).
Navier-Stokes: variational formulation

- Introducing $D(u) = \frac{1}{2}(\nabla u + {}^t\nabla u)$, the symmetric gradient of $u$ (the rate of deformation tensor) and thanks to the incompressibility condition, we have: $\text{div}(2D(u)) = \Delta u$.

- Using Green's formula on each domain $\Omega^i$ yields:

$$- \int_{\Omega^i} \text{div}(2\mu D(u^i))v^i \, dx = \int_{\Omega^i} 2\mu D(u^i) : \nabla v^i \, dx - \int_{\partial\Omega^i} 2\mu D(u^i)n^i \cdot v^i \, ds$$

$$= \int_{\Omega^i} 2\mu D(u^i) : D(v^i) \, dx - \int_{\partial\Omega^i} 2\mu D(u^i)n^i \cdot v^i \, ds$$

$$\int_{\Omega^i} \nabla p^i \cdot v^i \, dx = -\int_{\Omega^i} p^i \text{div} v^i \, dx + \int_{\partial\Omega^i} p^i n^i \cdot v^i \, ds$$

where $A : B = \sum_{i,j=1}^{n} A_{ij} B_{ij}$.

- Finally, we come to the following equation:

$$\alpha \int_{\Omega} \rho u \cdot v \, dx + \int_{\Omega} 2\mu D(u) : D(v) \, dx - \int_{\Omega} p \text{div} v \, dx = \int_{\Omega} \rho f \cdot v \, dx + \alpha \int_{\Omega} \rho w \cdot v \, dx$$

$$+ 2 \sum_{i=1}^{2} \int_{\partial\Omega^i} 2\mu D(u^i)n^i \cdot v^i \, ds - 2 \sum_{i=1}^{2} \int_{\partial\Omega^i} p^i n^i \cdot v^i \, ds. \quad (146)$$
Navier-Stokes: variational formulation

- using the interface and geometric conditions $n^2 = -n^1$ on $\Gamma$, $v = 0$ on $\Sigma$ we have:

\[
\sum_{i=1}^{2} \int_{\partial\Omega^i} 2\mu D(u^i)n^i \cdot v^i \, ds - \sum_{i=1}^{2} \int_{\partial\Omega^i} p^i n^i \cdot v^i \, ds = \sum_{i=1}^{2} \int_{\Gamma} (2\mu D(u^i) - p^i I)n^i v^i \, ds
\]

\[
= \int_{\Gamma} (\sigma^1 n^1 v^1 - \sigma^2 n^1 v^2) \, ds
\]

\[
= \int_{\Gamma} -\gamma \kappa n^1 \cdot v \, ds
\]

- the variational formulation of the homogeneous problem reads: given the functions $f, w, \mu, \rho$ (supposed constant in each time interval) and the constant $\alpha$; find $u \in V$ and $p \in M$ solving:

\[
\begin{cases}
\alpha \int_{\Omega} \rho u \cdot v \, dx + \int_{\Omega} 2\mu D(u) : D(v) \, dx - \int_{\Omega} p \, \text{div} \, v \, dx = \int_{\Omega} \rho f \cdot v \, dx \\
+ \alpha \int_{\Omega} \rho w \cdot v \, dx - \int_{\Gamma} \gamma \kappa n^1 \cdot v \, ds \quad \forall v \in V \\
\int_{\Omega} q \, \text{div} \, u \, dx = 0 \quad \forall q \in M
\end{cases}
\]
Navier-Stokes: spatial discretization

- Galerkin finite element approximation leads to the discrete problem:

\[
\text{find } (u_h, p_h) \in V_h \times M_h \text{ s.t. :}
\]

\[
\begin{cases}
  a(u_h, v_h) + b(v_h, p_h) = l(v_h), & \forall v_h \in V_h, \\
  b(u_h, q_h) = 0 & \forall q_h \in M_h
\end{cases}
\]

(147)

with \( V_h \subset V \) and \( M_h \subset M \) two families of finite dimensional subspaces and \( a(u_h, v_h) \), \( b(v_h, p_h) \), \( l(v_h) \) are bilinear and linear forms defined on \( V_h \times V_h \), \( V_h \times M_h \) and \( V_h \) respectively as follows:

\[
\begin{cases}
  a(u_h, v_h) = \sum_{K \in T_h} \alpha \int_K \rho u_h \cdot v_h dx + \sum_{K \in T_h} \int_K 2\mu \mathcal{D}(u_h) : \mathcal{D}(v_h) dx \\
  b(v_h, p_h) = \sum_{K \in T_h} \int_K -p_h \text{div } v_h \\
  l(v_h) = \sum_{K \in T_h} \int_K \rho f_h \cdot v_h dx + \sum_{K \in T_h} \alpha \int_K \rho \omega_h \cdot v_h dx + l_{\Gamma_h}(v_h)
\end{cases}
\]

(148)

where the term \( l_{\Gamma_h}(v_h) = -\int_{\Gamma_h} \gamma \kappa n_h^1 \cdot v_h \) is a discretization of the surface tension \(-\int_{\Gamma} \gamma \kappa n^1 \cdot v.\)
Variational formulation: existence and uniqueness

- the existence and the uniqueness of a solution to the weak formulation of the generalized Stokes problem can be established, see [EG04] or [Qua09].

- proof relies on:
  i) the ellipticity of the form \( a(., .) \) (Poincaré-Friedrichs inequality);
  ii) the compatibility of the spaces, velocity and pressure (Babuska-Brezzi inf-sup condition on the form \( b(., .) \)), i.e. there existing a positive constant \( C \) such that:

\[
\inf_{q} \sup_{v} b(v, p) \frac{\|v\|_{1}}{\|q\|_{0}} \geq C > 0
\]

where \( \|v\|_{1} = \left( \sum_{i=1}^{d} \|v_{i}\|_{1}^{2} \right)^{1/2} \) and \( \|\cdot\|_{1}, \|\cdot\|_{0} \) are standard notations of norms in the Sobolev spaces \( H^{1}(\Omega), L^{2}(\Omega) \) respectively.

References

Variational formulation: compatibility condition

- The generalized Stokes problem also involves a compatibility condition: the discrete spaces for the velocity and the pressure need to be compatible (see section on Stokes).

- In practice, mini elements ($\mathbb{P}_1$-bubble/$\mathbb{P}_1$) or Taylor-Hood elements ($\mathbb{P}_2$/$\mathbb{P}_1$) are used to solve it.

- The problem (147) leads to solve the sparse symmetric linear system:

\[
\begin{pmatrix}
A & B^t \\
B & 0
\end{pmatrix}
\begin{pmatrix}
U \\
P
\end{pmatrix}
=
\begin{pmatrix}
F \\
0
\end{pmatrix}
\]  

(150)

where $A$, $B$ and $F$ correspond to the bilinear forms $a_h$, $b_h$ and to the right-hand side, respectively.

- This linear system is solved using Uzawa’s method or by a penalty method.
Approximation of surface tension term

- $\Gamma_h$ is a piecewise affine approximation of $\Gamma$ included in $T_h$.

- Using a quadrature formula on each edge $E \subset \Gamma_h$, we write:

$$\int_{\Gamma_h} \gamma \kappa v_h \cdot n^1 \, ds = \sum_{E \subset \Gamma_h} \int_E \gamma \kappa v_h \cdot n^1 \, ds = \sum_{E \subset \Gamma_h} \frac{|E|}{2} \sum_{x_i \in E} \gamma \kappa(x_i) \, v_h(x_i) \cdot n^1(x_i)$$

$$= \sum_{x_i \in \Gamma_h} \gamma \kappa(x_i) \, v_h(x_i) \cdot n^1(x_i) \sum_{E \ni x_i} \frac{|E|}{2}.$$

where $n^1$ is the unit normal vector with respect to $\Omega^1$.

- how to evaluate $n^1$ and $\kappa$ at each point of $\Gamma_h$?

- in principle, with level sets, function $\phi$ is used to compute normal and curvature along $\Gamma$:

$$n = \frac{\nabla \phi}{|\nabla \phi|}_{\phi=0}, \quad \kappa = \text{div} \, n = \text{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right)_{\phi=0}.$$

but it is too sensitive to numerical artifacts.
Approximation of surface tension term

We propose another scheme to evaluate the differential quantities.

- Suppose $x_{i-1}, x_i, x_{i+1}$ are three successive points along $\Gamma_h$.
- Define $n^1(x_i)$ as the unit vector orthogonal to the edge vector $e_{i+1,i-1}$, considered as the approximation of the tangent $\tau(x_i)$.

- The local radius of curvature $\rho(x_i)$ is then approximated by:

$$\rho(x_i) = \frac{1}{4} \left( \frac{\langle e_{i,i-1}, e_{i,i-1} \rangle}{\langle -n^1(x_i), e_{i,i-1} \rangle} + \frac{\langle e_{i,i+1}, e_{i,i+1} \rangle}{\langle -n^1(x_i), e_{i,i+1} \rangle} \right)$$

and the local mean curvature is defined as

$$\kappa(x_i) = 1/\rho(x_i).$$
**Numerical scheme**

- time interval $[0, T]$ is divided into $N$ intervals $[t^{n-1}, t^n]$,

- at the iteration $n$, given $T_h^n$, solve Navier-Stokes and advection equations to obtain the solutions $(u^n, p^n)$ and $\phi^n$,

- the scheme is an iterative procedure that involves mesh adaptation, it reads:
  1. **init** $t = 0$: $u^0 = u_0(x), p^0 = p_0(x), \phi^0 = \phi_0(x), T_h^0$,
  2. **for** $n = 1, ..., N$ **do**

<table>
<thead>
<tr>
<th>mesh</th>
<th>input</th>
<th>output</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_h^{n-1}$</td>
<td>$(u^{n-1}, p^{n-1}, \phi^{n-1})$</td>
<td>$(u^n, p^n)$</td>
</tr>
<tr>
<td>$T_h^n$</td>
<td>$(u^n</td>
<td>_{\Gamma})$</td>
</tr>
<tr>
<td>$T_h^{n-1}$</td>
<td>$(\tilde{u}^n, \phi^{n-1})$</td>
<td>$\phi^n$</td>
</tr>
<tr>
<td>$T_h^n$</td>
<td>$(\tilde{u}^n, \phi^n)$</td>
<td>$T_h^n$</td>
</tr>
<tr>
<td>$T_h^{n-1}, T_h^n$</td>
<td>$(\tilde{u}^n, p^n, \phi^n)$</td>
<td>$T_h^n, (u^n, p^n, \phi^n)$</td>
</tr>
</tbody>
</table>

*Navier-Stokes, velocity extension, level set advection, mesh adaptation, $L^2$-projection*
Section 4.4

Numerical results
2d rising bubble

- study of the rising and the deformation of a single bubble under gravity in a fluid confined in a rectangular domain \( \Omega = [0, 4] \times [0, 10] \).
- initial configuration: circular bubble of radius \( r = 0.5 \) centered at \([2, 1.5]\).
- initial mesh contains 2,494 nodes.
- boundary conditions: no-slip condition \((u=0)\) on the horizontal walls and free-slip condition \((\tau \cdot \sigma n = 0 \text{ and } u \cdot n = 0)\) on the vertical walls.
- Reynold number and Bond number (or Eotvos number) characterize the simulations:

\[
Re = \frac{\rho_1 \sqrt{g(2r)^{3/2}}}{\mu_1}, \quad Bo = \frac{4\rho_1 gr^2}{\gamma} \tag{153}
\]

- parameters: constant densities and viscosities
  \( \rho_1 = 100 \text{kg.m}^{-3}, \mu_1 = 0.1 \text{kg.m}^{-1}.s^{-1}, \rho_2 = 1.0 \text{kg.m}^{-3}, \mu_2 = 0.01 \text{kg.m}^{-1}.s^{-1}, \)
  gravity \( g = 9.81 \times 10^{-3} \text{m.s}^{-2} \),
  surface tension \( \gamma = 6.10^{-3} \text{N.m}^{-1} \).
2d rising bubble

Rising bubble: 2d domain

- $u_x = u_y = 0$
- $u_x = 0$
- $10$
- $1.5$
- $2.0$
- $4.0$

$t = 5s$

$t = 10s$
2d rising bubble: surface tension coefficient

\[ \gamma = 6e^{-5} \]

\[ \gamma = 6e^{-3} \]

\[ \gamma = 2.5e^{-2} \]

\[ \gamma = 9e^{-2} \]
3d rising bubble

- Extension to 3d of the rising bubble problem: bubble with diameter 0.5 m initialized at [0.75, 0.75, 1.0] in a domain $\Omega = [0, 1.5] \times [0, 1.5] \times [0, 4.5]$.

- At low Reynolds number, the shape of the bubble deforms slowly and becomes dimpled ellipsoidal and more distorted as time increases.

$t = 0$ s  $t = 2.0$ s  $t = 4.0$ s  $t = 8.0$ s  $t = 10.0$ s
Section 6.1

Residual and geometric estimates
Error estimate, adaptative scheme

**Context:** we compute a solution \( u_h \) for a model problem using a triangulation \( T_h \).

The analysis of this solution allows to decide whether \( u_h \) is an **accurate approximation** of \( u \) or not. In such case, we deduce from \( u_h \) a sizing information that allows to **modify the triangulation** \( T_h \) such as to compute a **more accurate** solution \( u_h \).

Such process is called an **adaptative scheme**.

It assumes that the **error estimate** can be performed only based on known quantities: diameter \( h \), solution \( u_h \), right-hand side \( f \).

This estimate is called **a posteriori estimate**, to make a distinction with **a priori** estimates like:

\[
\|u - u_h\|_{H^1(\Omega)} \leq c h^k \|u\|_{H^{k+1}(\Omega)},
\]

which are quite limited in practice since \( \|u\|_{H^{k+1}(\Omega)} \) is **unknown**.
Error estimate, adaptative scheme

Motivations:
we consider the model problem posed in $\Omega$ of the form:

$$\text{find } u \in W \text{ such that } a(u,v) = f(v), \text{ for all } v \in V,$$

where $V$ and $W$ are Hilbert spaces, $f \in V'$ and $a \in \mathcal{L}(W \times V, \mathbb{R})$.

We suppose also that $a(\cdot, \cdot)$ satisfies the hypothesis of the Nečas (or Lax-Milgram) theorem, and that the problem is well-posed.

For now, we consider given a triangulation $\mathcal{T}_h$ of the domain and the approximation spaces $W_h$ and $V_h$. The approximate problem is then:

$$\text{find } u_h \in W_h \text{ such that } a(u_h,v_h) = f(v_h), \text{ for all } v_h \in V_h,$$

which is also considered to be well-posed.
**Error estimates, adaptative scheme**

**Definition 13** a function $e(h, u_h, f)$ is called an **a posteriori error** if it provides an upper bound on the approximation error: $\|u - u_h\|_W \leq e(h, u_h, f)$. Furthermore, if $e(h, u_h, f)$ is such that

$$e(h, u_h, f) = \left( \sum_{T \in \mathcal{T}_h} e_T(u_h, f)^2 \right)^{1/2},$$

then $e_K(u_h, f)$ is called an **error indicator**.

It remains to connect the error indicator and the mesh adaptation.

Intuitively, if the value $e_T(u_h, f)$ is large (resp. small), we can locally (i.e., in the neighborhood of $T$) **refine** (resp. **coarsen**) the triangulation.

The aim is thus to **equidistribute** the error on $\mathcal{T}_h$ so as to achieve a maximal accuracy for a given number of degrees of freedom (nodes).
Error estimate, adaptative scheme

One main concern about the error estimate: is it optimal?

The optimality of the estimate is a guarantee that the evaluation $e_T(u_h, f)$ is not too pessimistic and does not lead to over-refine the triangulation. It is thus interesting to introduce estimates of the form:

$$c_1 e_T(u_h, f) \leq \|u - u_h\|_{W,K} \leq c_2 e_T(u_h, f), \quad \text{pour tout } T \in \mathcal{T}_h.$$

for the norm $\|u - u_h\|_W$ taken as $(\sum_{T \in \mathcal{T}_h} \|u - u_h\|_{W,T}^2)^{1/2}$. This type of inequality means that the error indicator $e_T(u_h, f)$ is equivalent to the local error $\|u - u_h\|_{W,T}$.

The objective is thus to obtain:

- "anisotropic" bounds on the derivatives, so as to prescribe the size, the shape and the orientation of the simplices of $\mathcal{T}_h$.
- estimates in $L^1$, $L^2$ norms or for the $H^1$ seminorm.
A priori error estimate

**Theorem 13** Let $M$ be the constant associated with the continuity of the bilinear form $a$: $a(u, v) \leq M \|u\|\|v\|$ and $\alpha$ be the constant associated with the hypothesis of ellipticity (coercivity), we have:

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

If $a(\cdot, \cdot)$ is symmetric, $a(u - u_h, v_h) \equiv 0$ for all $v_h \in V_H$ means that $u_h$ is the projection of $u$ in $V_h$ according to the inner product $a$. Thus, we have:

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

The result follows:

**Theorem 14** For Lagrange $P_k$ finite elements, and considering a sufficiently smooth solution $u \in H^{k+1}(\Omega)$, we have the error bound:

$$\|u - u_h\|_{H^1(\Omega)} \leq C h^k |u|_{H^{k+1}(\Omega)}$$

where $h$ is the maximal diameter of the simplices in $\mathcal{T}_h$ and $C$ is a constant independent of $h$. 
A priori error estimate

We look at the one dimensional problem. Let consider a discretisation of $I = [a, b]$ in $N$ subintervals $T_i = [x_{i-1}, x_i]$ and suppose $V_h$ is the space of continuous piecewise affine functions. We rely on the classical bound:

$$\|u - u_h\|_{H^1(I)} \leq \frac{M}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_{H^1(I)}, \quad \forall v_h \in V_h.$$  

and then we deduce the following results:

**Theorem 15** Let denote $h = \max_i |x_i - x_{i-1}|$ and for every $x \in [a, b]$ we have:

$$|u(x) - P_h u(x)| \leq \frac{h^2}{8} \max_{x \in [a, b]} |u''(x)|$$

$$|u'(x) - P'_h u(x)| \leq \frac{h}{2} \max_{x \in [a, b]} |u''(x)|$$

$$\|u(x) - P_h u(x)\|_{H^1(I)} \leq Ch \max_{x \in [a, b]} |u''(x)|$$

$$\|u(x) - u_h(x)\|_{H^1(I)} \leq Ch \max_{x \in [a, b]} |u''(x)|$$

By noticing that: $\|u - P_h u\|_{H^1(\Omega)} = \left( \sum_T \|u - P_h u\|_{H^1(\Omega)}^2 \right)^{\frac{1}{2}}$. 

Error estimates
A priori error estimate

**Remark 2** The previous bounds are only available if the forms $a(\cdot, \cdot)$ and $l(\cdot)$ are employed in the exact and in the approximated problems (hence, in the exact evaluation of the integrals).

**Proof:** For each $[x_{i-1}, x_i]$, we introduce the two basis functions of $V_h$ associated with the points $x_{i-1}$ and $x_i$, respectively. Let us recall that their restriction in $[x_{i-1}, x_i]$ are written as:

$$w_{i-1}(x) = \frac{x_i - x}{x_i - x_{i-1}} \quad w_i(x) = \frac{x - x_{i-1}}{x_i - x_{i-1}}$$

(154)

We have the following relations for all $x$ in $[x_{i-1}, x_i]$: 

$$w_{i-1}(x) + w_i(x) = 1 \quad x_{i-1}w_{i-1}(x) + x_iw_i(x) = x$$

and

$$P_h u(x) = u(x_{i-1})w_{i-1}(x) + u(x_i)w_i(x)$$

and thus

$$(P_h u)'(x) = u(x_{i-1})w_{i-1}'(x) + u(x_i)w_i'(x).$$

Now, we use the second order Taylor expansions as follows:

$$u(x_{i-1}) = u(x) + (x_{i-1} - x)u'(x) + \frac{(x_{i-1} - x)^2}{2}u''(\xi_i)$$

$$u(x_i) = u(x) + (x_i - x)u'(x) + \frac{(x_i - x)^2}{2}u''(\eta_i)$$
A priori error estimate

and we obtain:

\[ P_h u(x) = u(x) + \frac{1}{2} [(x_{i-1} - x)^2 w_{i-1}(x) u''(\xi_i) + (x_i - x)^2 w_i(x) u''(\eta_i)] \]

and

\[ (P_h u)'(x) = u'(x) + \frac{1}{2} [(x_{i-1} - x)^2 w_{i-1}'(x) u''(\xi_i) + (x_i - x)^2 w_i'(x) u''(\eta_i)] \]

On the one hand, we introduce a bound on:

\[ \frac{1}{2} [(x_i - x)^2 w_i(x) u''(\xi_i) + (x_i - x)^2 w_i(x) u''(\eta_i)] \leq \frac{h^2}{8} \]

and then

\[ \frac{1}{2} [(x_{i-1} - x)^2 w_{i-1}(x) + (x_i - x)^2 w_i(x)] \max_{x \in [a,b]} |u''(x)| \]

and similarly for the derivatives:

\[ \frac{1}{2} [(x_{i-1} - x)^2 w_{i-1}'(x) u''(\xi_i) + (x_i - x)^2 w_i'(x) u''(\eta_i)] \leq \frac{h^2}{8} \]

\[ \frac{1}{2} [(x_{i-1} - x)^2 |w_{i-1}'(x)| + (x_i - x)^2 |w_i'(x)|] \max_{x \in [a,b]} |u''(x)| \]
A priori error estimate

**Remark 3** In each simplex $T_i$ of length $h_i$, we obtain the bounds on the local error:

$$\|u - P_h u\|_{L^2(T_i)} \leq C \frac{h_i^2}{8} \max_{x \in T_i} |u''(x)|$$

$$\|u - P_h u\|_{H^1(T_i)} \leq C h_i \max_{x \in T_i} |u''(x)|$$

And we deduce the choice of the sizes $h_i$ to equidistribute the errors:

- in $L^2$ norm: $h_i \propto 1/\sqrt{\max_{x \in T_i} |u''(x)|}$
- in $H^1$ norm: $h_i \propto 1/\max_{x \in T_i} |u''(x)|$

It can be shown that each choice minimizes the global errors $L^2$ and $H^1$, respectively.
A priori error estimate

Hence, for the $L^2$ error, by using for example:

$$P_h u(x) = u(x) + \frac{1}{2}[(x_{i-1} - x)^2 w_{i-1}(x) u''(\xi_i) + (x_i - x)^2 w_i(x) u''(\eta_i)],$$

we have then

$$\|u - P_h u\|_{0,2,\Omega} \leq \left[ \frac{1}{4} \sum_{i=1}^{i=N} \left( \int_{x_{i-1}}^{x_i} [(x_{i-1} - x)^2 w_{i-1}(x) + (x_i - x)^2 w_i(x)]^2 dx \right) \max_{[x_{i-1},x_i]} |u''|^2 \right]^{\frac{1}{2}}.$$

An explicit computation yields to write

$$[(x_{i-1} - x)^2 w_{i-1}(x) + (x_i - x)^2 w_i(x)] = (x_i - x)(x - x_{i-1})$$

And we obtain, by posing $h_i = x_i - x_{i-1}$:

$$\|u - P_h u\|_{0,2,\Omega} \leq \left[ \frac{1}{120} \sum_{i=1}^{i=N} h_i^5 \max_{[x_{i-1},x_i]} |u''|^2 \right]^{\frac{1}{2}}.$$
**A priori error estimate**

The minimum is obtained by computing the gradient of this error with respect to the $x_i$. This yields the following relation:

$$h_i \sqrt{\max_{[x_{i-1}, x_i]} |u''|} = h_{i+1} \sqrt{\max_{[x_i, x_{i+1}]} |u''|} \quad \forall i = 1, \ldots, N - 1$$

Figure 2: Optimal point placement for an integration (Matlab) and to equidistribute the interpolation error in $L^2$ norm.
Remark 4 Suppose \( u \) is of class \( C^2 \), a polynomial of degree 2. Then, \( u'' \) is constant. In this case, we observe that the maximum of the interpolation error is achieved at the midpoint of the interval.

Furthermore, the derivative of \( u \) is an affine function. The derivative of its interpolate, \( P_h'u \), is a constant function which is equal to the value of \( u' \) at the midpoint of the interval.

In summary, at the midpoint of the interval \([x_{i-1}, x_i]\), we have both a maximum of the interpolation error on \( u \) and no interpolation error on \( u' \). This explains why it is desirable to estimate the derivatives of the solution at the midpoints of the intervals.
Two dimensional error estimate

Consider a polygonal planar domain \( \Omega \) covered by a triangulation \( T_h \). Consider the space of continuous piecewise affine functions \( V_h \) (affines on the triangles \( T_i \)). Again, we introduce the bound:

\[
\| u - u_h \|_{H^1(\Omega)} \leq \frac{M}{\alpha} \inf_{v_h \in V_h} \| u - v_h \|_{H^1(\Omega)} \quad \forall v_h \in V_h
\]  

(155)

Similarly, we take for \( v_h \) the interpolate \( P_h u \) of \( u \) in the space of finite elements \( V_h \). We bound the approximation error \( \| u - u_h \| \) by a constant multiplied by the interpolation error \( \| u - P_h u \| \).

We denote by \( h \) the diameter and by \( \theta_0 \) the smallest vertex angle among each triangles of \( T_h \); The Hessian matrix \( D^2 v \) of a \( C^2 \) continuous function \( v \) is then:

\[
D^2 v(x, y) = \begin{pmatrix}
\frac{\partial^2 v}{\partial x^2} & \frac{\partial^2 v}{\partial x \partial y} \\
\frac{\partial^2 v}{\partial x \partial y} & \frac{\partial^2 v}{\partial y^2}
\end{pmatrix}
\]

(156)

and \( |D^2 v(x, y)| \) will then denote its spectral norm.
Two dimensional error estimate

We have the following result:

**Theorem 16** for all $x, y \in \Omega$:

$$|u(x, y) - P_h u(x, y)| \leq \frac{h^2}{2} \sup_{x,y \in \Omega} |D^2 v(x, y)|$$

$$|\nabla u(x, y) - \nabla P_h u(x, y)| \leq 3\frac{h}{\sin(\theta_0)} \sup_{x,y \in \Omega} |D^2 v(x, y)|$$

hence the interpolation error is bounded:

$$\|u - P_h u\|_{H^1(\Omega)} \leq C h \sup_{x,y \in \Omega} |D^2 v(x, y)|$$

and the approximation error is bounded:

$$\|u - u_h\|_{H^1(\Omega)} \leq C h \sup_{x,y \in \Omega} |D^2 v(x, y)|$$
Two dimensional error estimate

**Remark 5** The previous result raises several comments and remarks:

1. The dominant term in the error bound is related to the gradients.
   
   It will be minimal if \( \frac{h}{\sin(\theta_0)} \) is minimal; this yields to consider the equilateral triangle as the optimal shape. It is also the largest triangle inscribed in a circle of given radius.

2. Suppose \( u \) is a polynomial of degree 2, then \( D^2u \) is a matrix with constant coefficients.
   
   In the case of isotropic and homogeneous problem, the Hessian matrix \( D^2u \) is a scalar matrix.

   In each triangle, the maximum of the interpolation error is attained at the center of the circumscribed circle.

   The gradient of \( u \) is an affine function and the gradient of its interpolate \( \nabla P_h u \) is a constant function equal to the value of \( \nabla u \) at the center of gravity of the triangle.

   At the circumcenter of the equilateral triangle, we have a maximum for the interpolation error on \( u \) and the interpolation error on \( \nabla u \) vanishes.
2. Two dimensional error estimate

3. For non-isotropic problems, the previous analysis is incomplete.
   We have to consider the eigenvalues and eigenvectors of the Hessian matrix in each triangle and to find the largest triangle inscribed in an ellipse as the optimal triangle.

Hence, we go back to the expressions:

\[
P_h u(x, y) = u(x, y) + \frac{1}{2} \sum_{i=1}^{3} \lambda_i(x, y) \overrightarrow{M A_i} \cdot D^2 u(\xi, \eta) \overrightarrow{MA_i}
\]

\[
\nabla P_h u(x, y) = \nabla u(x, y) + \frac{1}{2} \sum_{i=1}^{3} \nabla \lambda_i(x, y) \overrightarrow{MA_i} \cdot D^2 u(\xi, \eta) \overrightarrow{MA_i}.
\]

However, the Hessian matrices diagonalized in their eigenbasis are no longer scalar matrices. We have to consider taking the absolute values of the eigenvalues.
Two dimensional error estimate

We define a new inner product and then a new metric for which the isolines of error become concentric ellipses. Let us denote

\[ D^2 = \begin{pmatrix} |d_1| & 0 \\ 0 & |d_2| \end{pmatrix} \]

la the symmetric and positive definite matrix associated with the inner product in this metric. We still have the bounds:

\[
|u(x, y) - P_h u(x, y)| \leq \frac{1}{2} \sum_{i=1}^{i=3} \lambda_i(x, y) \overrightarrow{MA}_i \cdot D^2 \overrightarrow{MA}_i
\]

\[
|\nabla u(x, y) - \nabla P_h u(x, y)| \leq \frac{1}{2} \sum_{i=1}^{i=3} |\nabla \lambda_i(x, y)| \overrightarrow{MA}_i \cdot D^2 \overrightarrow{MA}_i
\]

**Remark 6** Again, we observe that the error on \( u \) vanishes along the ellipse circumscribed to the triangle and is maximal at its center. Regarding the error on \( \nabla u \), it is the opposite.
Two dimensional error estimate

Figure 3: Isolines of the error on the equilateral triangle: isotropic case (left), anisotropic case with adapted (middle) and non-adapted triangle (right).

**Remark 7** We can conclude that the choice of the optimal shape for the elements is related to the error on the gradients. The interpolation error on \( u \) is only related to the size. Regarding the approximation error, it depends on both the size and the shape of the elements, as it is known by the \( H^1 \) norm.
Residual based error estimates

We consider a model problem, involving a second order scalar PDE. Let $\Omega$ be a polyhedral domain in $\mathbb{R}^d$ and consider the problem posed for $f$ in $L^2(\Omega)$

\[
\begin{cases}
\text{find } u \in H^1_0(\Omega), \text{ such that } \\
a(u, v) = \int_{\Omega} fv, \quad \forall v \in H^1_0(\Omega).
\end{cases}
\]

where $a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v$ is elliptic. Let $\mathcal{T}_h$ denotes a family of affine triangulations of $\Omega$ composed of Lagrange finite element Lagrange $\mathbb{P}_k$. We denote $V_h \subset H^1_0(\Omega)$ the approximation space and the approximate problem is then:

\[
\begin{cases}
\text{find } u_h \in V_h, \text{ such that } \\
a(u_h, v_h) = \int_{\Omega} fv_h, \quad \forall v_h \in V_h.
\end{cases}
\]

We recall here the \textit{a priori} error estimate:

\[
\|u - u_h\|_{H^1(\Omega)} \leq c_1 \inf_{v_h \in V_h} \|u - v_h\|_{H^1(\Omega)} \leq c_2 \left( \sum_{T \in \mathcal{T}_h} h_T^{2k} \|u\|^2_{H^{k+1}(T)} \right)^{1/2}.
\]
Residual based error estimates

To be slightly more general, we do not suppose here the ellipticity of the form but only its stability of $a(\cdot, \cdot)$ according to Nečas theorem:

$$\inf_{u \in H^1_0(\Omega)} \sup_{v \in H^1_0(\Omega)} \frac{a(u, v)}{\|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}} \geq \alpha. \quad (157)$$

We can the deduce:

$$\alpha \|u - u_h\|_{H^1(\Omega)} \leq \sup_{v \in H^1_0(\Omega)} \frac{a(u - u_h, v)}{\|v\|_{H^1(\Omega)}} \leq \sup_{v \in H^1_0(\Omega)} \frac{\langle \Delta(u - u_h), v \rangle_{H^{-1}(\Omega), H^1_0(\Omega)} }{\|v\|_{H^1(\Omega)}} \leq \|f + \Delta u_h\|_{H^{-1}(\Omega)},$$

that is now a true a posteriori estimate.

**Proposition 10** We have, for any $h$ (discretization parameter):

$$\|u - u_h\|_{H^1(\Omega)} \leq \frac{1}{\alpha} \|f + \Delta u_h\|_{H^{-1}(\Omega)}. \quad (158)$$
Residual based error estimates

**Remark 8** The previous result appeals two comments:

1. the error estimate (158) is indeed an a posteriori estimate, but in practice, the norm $\| \cdot \|_{H^{-1}}$ is quite tedious to evaluate.

2. the intuitive idea of integrating by parts is interesting since it leads to put the unknown $u$ in the right-hand side. It is possible to avoid the norm $\| \cdot \|_{H^{-1}}$ by performing the integration on each element $T$ of $T_h$.

**Theorem 17** Suppose the family $(T_h)$ is regular. Then it exists $c > 0$ such that for any $h$

$$
\| u - u_h \|_{H^1(\Omega)} \leq c \left( \sum_{T \in T_h} e_T(u_h, f)^2 \right)^{1/2},
$$

where we introduced the error estimate

$$
e_T(u_h, f) = h_T \| f + \Delta u_h \|_{L^2(T)} + \frac{1}{2} \sum_{e \in E_T} h_e^2 \| [\partial_n u_h] \|_{L^2(I)},
$$

with $e$ a face of $T$, $E_T$ the set of (non boundary) faces of $T$, $h_e$ the diameter of $e$ and $[\partial_n u_h]$ the jump of the normal derivative of $u_h$ across $e$. 
Residual based error estimates

A family of triangulations \((\mathcal{T}_h)_h\) is regular if there exists a constant \(\sigma_0\) such that

\[
\forall h, \forall T \in \mathcal{T}_h, \quad \sigma_T = \frac{h_T}{\rho_T} \leq \sigma_0.
\]

Before writing the proof, we need to introduce Clément’s operator.

**Proposition 11** Let \((\mathcal{T}_h)_h\) be a regular family of affine triangulations on \(\Omega\). Let \(V_h\) be an approximation space built with \(\mathcal{T}_h\) supposed \(H^1\)-conforming. Let \(T\) be a simplex. We denote \(V(T)\) the set of all simplices of \(\mathcal{T}_h\) that share a common intersection with \(T\). Let \(e\) be a face of \(T\) and \(V(e)\) bet the set of all simplices having an intersection with \(e\). There exists an operator \(C_h : H^1(\Omega) \to V_h\) and a constant \(c > 0\) such that:

\[
\forall h, \forall T \in \mathcal{T}_h, \quad \left\{ \begin{array}{l}
\|v - C_h v\|_{L^1(T)} \leq c h_T \|v\|_{H^1(V(T))} \\
\|v - C_h v\|_{L^1(e)} \leq c e^{1/2} \|v\|_{H^1(V(e))}
\end{array} \right.
\]
Residual based error estimates

**Proof:** We notice that $a(u - u_h, v_h) = 0$ for all $v_h \in V_h$, and thus the stability inequality allows to write

$$
\|u - u_h\|_{H^1(\Omega)} \leq \frac{1}{\alpha} \sup_{v \in H^1_0(\Omega)} \frac{a(u - u_h, v - v_h)}{\|v\|_{H^1(\Omega)}}.
$$

On the right-hand side, we develop the upper part:

$$
a(u - u_h, v - v_h) = \int_{\Omega} (-\Delta u)(v - v_h) - \nabla u_h \cdot \nabla (v - v_h)
$$

$$
= \sum_{T \in T_h} \left[ \int_T (f + \Delta u_h)(v - v_h) - \sum_{e \in \partial T} \int_e (\partial_n u_h)(v - v_h) \right].
$$

Since $v - v_h$ vanishes on $\partial \Omega$, the sum on $e$ involves only the common faces between two simplices. Since $v - v_h$ is continuous across each interface $e$, it comes

$$
a(u - u_h, v - v_h) \leq \sum_{T \in T_h} \left( \|f + \Delta u_h\|_{L^2(T)} \|v - v_h\|_{L^2(T)} + \sum_{e \in E_T} \frac{1}{2} \|\partial_n u_h\|_{L^2(e)} \|v - v_h\|_{L^2(e)} \right)
$$

$$
\leq \sum_{T \in T_h} e_T(u_h, f) \max_T \left( h^{-1}_T \|v - v_h\|_{L^2(T)}, \max_{e \in E_T} (h^{-1/2}_e \|v - v_h\|_{L^2(e)}) \right).
$$
Residual based error estimates

We chose \( v_h = C_h v \) and thus we have

\[
a(u - u_h, v - C_h v) \leq c \sum_{T \in T_h} e_T(u_h, f) \eta_T,
\]

where \( c \) is the interpolation constant associated with \( C_h \) and \( \eta_T = \max(\|v\|_{H^1(V(T))}, \max_{e \in E_T} (\|v\|_{H^1(V(e))})) \).

We introduce the two integers

\[
M = \max_{T \in T_h} \text{card}\{T' \in T_h; T \in V(T')\}, \quad N = \max_{e \in E_h} \text{card}\{e \in E_h; T \in V(e)\},
\]

where \( E_h \) denotes the set of internal (non boundary) faces of \( T_h \). The numbers \( M \) and \( N \) depend only on the smoothness of \( T_h \) and are independent of the granularity \( h \). It appears clearly that

\[
\sum_{T \in T_h} \eta_T^2 \leq \max(M, N) \|v\|^2_{H^1(\Omega)},
\]

and thus we have:

\[
\frac{a(u - u_h, v - C_h v)}{\|v\|_{H^1(\Omega)}} \leq c \left( \sum_{T \in T_h} e_T(u_h, f)^2 \right)^{1/2}
\]

that can be combined with the bound on \( \|u - u_h\|_{H^1(\Omega)} \) deduced from the stability of \( a(\cdot, \cdot) \) to conclude.
Residual based error estimates

Remark 9  This result leads to several comments:

1. \( e_T(u_h, f) \) is called a residual based error estimate; indeed, the value \( f + \Delta u_h \) is the residual of the equation \( f = -\Delta u \).

2. if \( V_h \) is build with \( \mathbb{P}_1 \) Lagrange finite elements, \( \Delta u_h|_T = 0 \) and \( \partial_n u_h|_e \) is constant along \( e \), which leads to the result:

\[
e_T(u_h, f) = h_T\|f\|_{L^2(T)} + \frac{1}{2} \sum_{e \in E_T} h_e^{1/2} \text{mes}(e)^{1/2} \|\partial_n u_h\|_e.
\]

3. the coercivity of the bilinear form \( a(\cdot, \cdot) \) is not really involved, but we relied on the stability inequality (157). Hence, it this analysis can be generalized to non-coercive problems, provided that they fulfill the hypothesis of of Nečas theorem.
Residual based error estimates

The approach can be generalized to the following type of problem:

\[ \mathcal{L}u = -\nabla \cdot (\sigma \cdot \nabla u) + (\beta \cdot \nabla)u + \mu u, \]
supplied with mixed Dirichlet-Neumann boundary conditions.

Suppose \( \partial \Omega = \Gamma_D \cup \Gamma_N \), \( f \in L^2(\Omega) \) and \( g \in L^2(\Omega) \). We consider the problem

\[
\begin{aligned}
\text{find } u \in V, & \quad \text{such that} \\
\quad a(u, v) &= \int_{\Omega} f v + \int_{\Gamma_N} g v, \quad \forall v \in V.
\end{aligned}
\]

where \( V = \{ v \in H^1(\Omega); \quad v|_{\Gamma_D} = 0 \} \) and \( a(u, v) = \int_{\Omega} \nabla u \cdot \sigma \cdot \nabla v + v(\beta \cdot \nabla u) + \mu uv. \)

Suppose the problem is well posed. Its solution satisfies

\[
\begin{aligned}
\mathcal{L}u &= f \quad \text{a.e. in } \Omega \\
u &= 0 \quad \text{a.e. on } \Gamma_D, \quad \\
n \cdot \sigma \nabla u &= g \quad \text{a.e. on } \Gamma_N.
\end{aligned}
\]
Residual based error estimates

Consider a regular family \((T_h)_h\) of affine triangulations, a Lagrange \(\mathbb{P}_k\) finite element, and \(V_h \subset V\) the corresponding approximation space. The discrete problem reads:

\[
\begin{cases}
\text{find } u_h \in V_h, \text{ such that } \\
a(u_h, v_h) = \int_\Omega fv_h + \int_{\Gamma_N} gv_h, \forall v_h \in V_h.
\end{cases}
\]

**Proposition 12** Under these assumptions, we derive the following error indicator:

\[
e_T(u_h, f, g) = h_T \| f - \mathcal{L}u_h \|_{L^2(T)} + \frac{1}{2} \sum_{e \in E_T} h_e^{1/2} \| [n \cdot \sigma \cdot \nabla u_h] \|_{L^2(e)} + \sum_{e \in N_T} h_e^{1/2} \| g - n \cdot \sigma \cdot \nabla u_h \|_{L^2(e)},
\]

where \(N_T\) is the set of faces of \(T\) located on \(\Gamma_N\).
Geometric Error estimates

The analysis is here very similar to the study of parameterized surfaces. We have the result:

**Theorem 18** Let $\mathcal{T}$ be a triangulation of the domain $\Omega \subset \mathbb{R}^d$ ($d = 2$ or $3$) and $u$ be a $C^2$ function on $\Omega$. Let $V_h$ the space of continuous functions on $\Omega$ whose restriction to every simplex of $\mathcal{T}$ is a $\mathbb{P}^1$ function, and denote by $\pi_h : C(\Omega) \to V_\mathcal{T}$ the usual $\mathbb{P}^1$ finite elements interpolation operator. Then for every simplex $K \in \mathcal{T}$,

$$
\|u - \pi_h u\|_{L^\infty(K)} \leq \frac{1}{2} \left( \frac{d}{d + 1} \right)^2 \max_{x \in K} \max_{y,z \in K} \langle |\mathcal{H}(u)|(x)xy, xy \rangle
$$

where, for a symmetric matrix $S \in S_d(\mathbb{R})$, which admits the following diagonal shape in orthonormal basis $S = P \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_d \end{pmatrix}^t P$, we denote

$$
|S| := P \begin{pmatrix} |\lambda_1| & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & |\lambda_d| \end{pmatrix}^t P.
$$
Geometric Error estimates

This result can be rewritten under the following (more convenient) form:

\[
\| u - \pi_h u \|_{L^\infty(K)} \leq \frac{1}{2} \left( \frac{d}{d+1} \right)^2 \max_{x \in K} \max_{e \in E_K} \langle |\mathcal{H}(u)|(x)e, e \rangle
\]

However, the right-hand side is still tedious to evaluate.

Hence, we are searching for a metric tensor \( \tilde{M}_K \) on \( K \in \mathcal{T} \) such that:

\[
\max_{x \in K} \langle e, |\mathcal{H}(u)|(x)e, e \rangle \leq \langle \tilde{M}_K e \rangle,
\]

and such that the region defined by \( \{ \langle v, \tilde{M}_K v \rangle / \forall v \subset K \} \) has a minimal surface. This yields the result:

\[
\| u - \pi_h u \|_{\infty,T} \leq \frac{1}{2} \left( \frac{d}{d+1} \right)^2 \max_{e \in E_T} \langle e, \tilde{M}_K e \rangle.
\]
Geometric Error estimates

**Remark 10**  1. The relation shows that the interpolation error is locally proportional to the squared value of the diameter \( h \) of a simplex \( K \), with respect to the metric \( \tilde{M}_K \).

2. Hence, a control on the edge lengths for all \( K \in \mathcal{T}_h \) allows to bound the interpolation error on \( \mathcal{T}_h \).

3. The aim is then to equidistribute the error to control the accuracy of the solution.

Let \( \epsilon \) be the tolerance on the error on each mesh element. All edges in the mesh must be such that:

\[
\epsilon = c \langle e, \tilde{M}_K e \rangle, \quad \text{for all } e \in E_K,
\]

by introducing the metric tensor \( \mathcal{M}_T = \frac{c}{\epsilon} \tilde{M}_T \), edges must satisfy:

\[
\langle e, \mathcal{M}_Te \rangle = 1 \implies \text{length}_{\mathcal{M}_T}(e) = 1.
\]
Continuous metric

- The metric notion allows to define an equivalence class among triangulations: 2 unit triangulations are equivalent.

- The triangulation becomes then an unknown of the problem, given $\mathcal{M}$:
  
  - the density $d = \prod_{i=1}^{n} \frac{1}{h_i}$ where $h_i$ is the local size,
  
  - the complexity of a triangulation is given by:
    
    $\mathcal{C}(\mathcal{M}) = \int_{x \in \Omega} d(x) dx = \int_{x \in \Omega} \prod_{i=1}^{n} \frac{1}{h_i}(x) dx$,

- for given $N = Card(T_h)$, we search for $\mathcal{M}$ minimizing the $L^p$ norm of the error, which yields:
  
  $\min_{\mathcal{M}} \int_{\Omega} (e_{\mathcal{M}}(x))^p dx = \min_{h_i} \int_{\Omega} \left( \sum_{i=1}^{n} h_i^2 \left| \frac{\partial^2 u}{\partial \alpha_i^2}(x) \right| \right)^p dx$,

  under the constraint: $\int_{\Omega} d(x) dx = N$. 
Section 4.2
Mesh adaptation
Anisotropic triangulations

Numerous physical problems exhibit anisotropic features or behaviors. Examples of such features include for instance shocks waves, boundary layers, interfaces or singularities of a flow or vector field.

Figure 4: Anisotropic features in physical phenomena developing shocks. Left: bullet travelling at Mach 2.45, showing Mach cone and shock wavefronts. Right: Rayleigh-Taylor and Kelvin-Helmholtz instabilities on a shock wave propagating through a gas (courtesy N. Patterson, U. of Michigan).
A 2d example

Consider the domain $\Omega = [-1, 1]^2$ and the triangulation $\mathcal{T}$ defined on $\Omega$, with $|\mathcal{T}| = 144$. Consider the function $u$ defined on $\Omega$ as $u : \Omega \to \mathbb{R}^2$, $u(x, y) = \sqrt{1 - x^2}$. We denote by $\pi_h : C(\Omega) \to V_\mathcal{T}$ the usual $P^1$ finite elements interpolation operator:

$$\pi_h u(x) = \sum_{i=1}^{n} u_i \phi_i(x).$$
A 2d example

We define the interpolation error on $\mathcal{T}_h$ as:

$$e(u) = \|u - \pi_h u\|_{L^p(\Omega)}$$

and we have for this example the following table:

<table>
<thead>
<tr>
<th></th>
<th>$L^1$</th>
<th>$L^2$</th>
<th>$L^\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.029</td>
<td>0.059</td>
<td>0.133</td>
</tr>
</tbody>
</table>

Table 1: Interpolation errors on $\mathcal{T}_h$.

Problem:
We are looking for a triangulation $\mathcal{T}_h$ having the same cardinality but minimizing the complexity (number of vertices). This is equivalent to solve the problem:

$$\begin{cases}
\text{find } T_h = \text{argmin}(\|u - \pi_h u\|_{L^p}) \\
|T_h| = 144.
\end{cases}$$
Anisotropic triangulations

**Remark 11** Concerning the function $u$, we can state the following:

1. the largest curvature variation follows the $x$;
2. the function $u$ depends on $x$ only

Hence, in our example, function $u$ exhibits anisotropic features. This may help us to find the best discretization of $u$:

![Discretization of a curved segment (left) and of a line (right).](image)

Figure 5: Discretization of a curved segment (left) and of a line (right).
Optimal anisotropic triangulation

Figure 6: *Isotropic (top) vs. anisotropic (bottom) triangulations for function $u*. 
Optimal anisotropic triangulation

If we look at the interpolation errors in $L^p$ norms, we observe:

<table>
<thead>
<tr>
<th></th>
<th>$L^1$</th>
<th>$L^2$</th>
<th>$L^\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>iso</td>
<td>0.029</td>
<td>0.059</td>
<td>0.133</td>
</tr>
<tr>
<td>aniso</td>
<td>0.008</td>
<td>0.005</td>
<td>0.014</td>
</tr>
</tbody>
</table>

Table 2: Interpolation errors on isotropic and anisotropic $\mathcal{T}_h$.

This raises two questions:

1. how to measure mesh size and anisotropy ?
2. how to connect this information to error estimate and mesh generation ?
Anisotropic triangulations

- we can imagine constructing an *anisotropic* triangulation $\mathcal{T}_h$ in which elements will be aligned along specific directions to fit at best the *geometry* of the solution and will exhibit large *aspect ratios* (eccentricities).

- at the difference of isotropic triangulations, this peculiarity requires the ability to accurately control and handle the *eccentricity*, the *size* and the *orientation* of the simplices.

- this leads to the notion of *metric tensor*, a matrix-valued field.

- in the continuous geometric viewpoint, the metric defining an element is represented by an ellipsoid. Hence, *size*, *shape* and *orientation* notions are associated with its volume, length ratios between the lengths of its semi-axes and its principal axis vectors, respectively.
Euclidean metric space

Defined as a vector space of finite dimension in which the standard inner (dot) product is defined by:

\[ \langle u, v \rangle_M = ^t u M u \implies \text{length}(a, b) = \|ab\| = \sqrt{^t ab M ab} \]

where \( M \) is a \( d \times d \) symmetric and positive definite matrix, i.e.:

(i) \( \forall (u, v) \in \mathbb{R}^d \times \mathbb{R}^d, \langle u, M v \rangle = \langle v, M u \rangle \), (symmetry);

(ii) \( \forall u \in \mathbb{R}^d, \langle u, M u \rangle \geq 0 \), (positive);

(iii) \( \langle u, M u \rangle = 0 \Rightarrow u = 0 \), (definite).

\( M \) is a so-called metric tensor and we will use the term metric in short.

The Canonical Euclidean space correspond to the case where \( M = I_d \).
Anisotropic triangulations

- the metric tensor denoted as $M(x)$ allows to create a quasi-uniform triangulation in the metric related to $M$.

- more precisely, the mesh elements are equilateral in the metric

$$\text{tr} \left( (J_T)^t M_T J_T \right) = n \det \left( (J_T)^t M_T J_T' \right)^{\frac{1}{d}}, \quad \forall T \in \mathcal{T}_h,$$  

where $M_T$ is an average of $M(x)$ on the simplex $T$ and $J_T$ is the Jacobian matrix of the affine mapping $T_T$.

- moreover, the volume of the element $T$ in $\mathcal{T}_h$ is unitary:

$$\int_T \sqrt{\det(M(x))} \, dx = 1, \quad \forall T \in \mathcal{T}_h,$$  

or, in a discrete formulation:

$$|T| \sqrt{\det(M_T)} = 1, \quad \forall T \in \mathcal{T}_h.$$

where $M_T$ is an average of $M(x)$ on the simplex $T$ and $J_T$ is the Jacobian matrix of the affine mapping $T_T$. 
Anisotropic triangulations

- in the metric given by $M(x)$ for any $x$ in $\Omega$, the length of a curve $\gamma$ is

$$|\gamma|_M = \int_0^1 \sqrt{(\gamma'(t), \gamma'(t))_{M(\gamma(t))}} \, dt = \int_0^1 \sqrt{(\gamma'(t), M(\gamma(t))\gamma'(t))} \, dt$$  \hspace{1cm} (163)$$

where $\gamma(t) : [0, 1] \to \mathbb{R}^d$ is a parametrization of $\gamma$.

- in $\mathbb{R}^d$, the metric $M_T$ associated with an element $T$ can be represented by its \textit{unit sphere}, the ellipsoid $B(T')$ defined as

$$B(T') = \left\{ P \in \mathbb{R}^d, \|OP\|_{M_T} = \sqrt{(OP, M_T OP)} = 1 \right\}$$

where $O$ denotes the center of $B(T')$.

- we assume that the metric tensor $M(x)$ is a \textit{symmetric} and \textit{positive definite} $d \times d$ matrix. Then, the spectral value decomposition theorem leads to express matrix $M$ in terms of its $d$ eigenvalue-eigenvector pairs $(D_i, e_i)$ as:

$$M = PDP^t = \sum_{i=1}^d \lambda_i e_i e_i^t \quad \text{with} \quad P = [e_1, \ldots, e_d].$$

with $PP^t = P^tP = I_d$. 

Anisotropic triangulations

The geometric meaning of this decomposition is easy to understand and the matrix $P$ clearly allows to define the orientation and the matrix $D$ to prescribe the size and the shape of any simplex $T$ in $\mathcal{T}_h$.

Figure 7: The geometric interpretations of the notion of metric tensor in dimension 2, from left to right: an isotropic metric is represented as a circle, an anisotropic metric as an ellipse aligned with the canonical basis or aligned with any arbitrary orthogonal vectors.
Riemannian metric space

Defined as a vector space of finite dimension supplied with a metric tensor function:

\[ M = (M(x))_{x \in \Omega} \]

in which the standard inner (dot) product is defined by:

\[ \langle u, v \rangle_M = ^t u M u \implies \text{length}(a, b) = \|ab\| = \sqrt{^t ab M ab} \]

**Remark 12** Notice that such metric space is not a Riemannian manifold (a real differentiable manifold in which each tangent space is equipped with an inner product \( M \), a Riemannian metric, which varies smoothly from point to point).

Here, we do not know any manifold.

Indeed, there is no global definition of inner product.

It can be assimilated to a function that can represent a set of Cartesian surface (graph surface).
Riemannian metric space

Suppose $\Omega \subset \mathbb{R}^d$.
A Cartesian surface $\Sigma$ of $\mathbb{R}^{d+1}$ is a sub-manifold of $\mathbb{R}^{d+1}$ (surface) defined locally by:

$$\Sigma = (x_1, \ldots, \sigma(x_1, \ldots, x_d)) \in \mathbb{R}^{d+1}$$

where $\sigma$ is a $C^2$-continuous function defined on $\mathbb{R}^d$.

Evaluating geometrical quantities in a Riemannian metric space is equivalent to evaluate these quantities on the underlying Cartesian surfaces.

Hence, a Riemannian metric space curves space (i.e, the parametrization space)
Evaluating geometric quantities on $\Sigma$ is then equivalent to computing geometric quantities in Riemannian metric space

$$M = (M(x))_{x \in \Omega}$$
Metric based operations

We may need to perform metric interpolations and metric intersections.

1. **intersection**: suppose that $M_1(p)$ and $M_2(p)$ are specified at a vertex $p \in \mathcal{T}_h$.
   - geometrically, intersecting metrics consists in defining the largest ellipsoid $E$ included in the intersection of the two ellipsoids $E_1$ and $E_2$ associated with $M_1$ and $M_2$;
   - algebraically, we use the simultaneous reduction of the two quadratic forms to find a basis of the vector space in which the matrices $M$ and $N$ associated with $E_1$ and $E_2$ are represented by $I_3$ and $D$, a diagonal matrix of $M_3(\mathbb{R})$. This is equivalent to finding a basis $(e_1, e_2, e_3)$ in which $M$ and $N$ are congruent to a diagonal matrix and then to deduce the metric tensor.

Denoting by $P = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix}$ the invertible matrix of $GL_3(\mathbb{R})$ formed by the eigenvectors $e_i$ of the matrix $N' = M^{-1}N$ leads to

$$M_\cap = (P^t)^{-1}D P^{-1},$$

where $D$ is diagonal with the coefficients $\max(\lambda_i, \mu_i)$.

Notice that $M_\cap$ is a symmetric positive definite matrix since $\det M_\cap = (P^{-1})^2 \det D$. 
2 interpolation: consider a parametrization of an edge $pq$ as $c : [0, 1] \rightarrow \mathbb{R}^3$, $c(t) = (1 - t)p + tq$ and $M_p$ and $M_q$. We look for $M(t)$, for any value of the parameter $t \in [0, 1]$.

- this procedure allows one to define a continuous metric field along the segment as follows

$$M(t) = \left((1 - t)M_p^{-\frac{1}{2}} + tM_q^{-\frac{1}{2}}\right)^{-2}, \quad 0 \leq t \leq 1. \quad (164)$$

- finding the interpolated metric tensor $M_t$ requires us to express the two matrices in a basis $\{e_i\}$ in which both are congruent to a diagonal matrix and then to deduce the metric tensor at point $t$ (cf. above).
Metric based operations

Figure 8: Examples of metric intersection (top) and metric interpolation (bottom) operations on anisotropic tensors in dimension two.
Anisotropic Delaunay triangulations

We consider the extension of the Delaunay empty circumscribed ball criterion to generate anisotropic simplices.

- finding all simplices having a circumsphere containing a given point $p$ in $\mathbb{R}^d$ can be performed using a test

$$\alpha(p, T) = \frac{d(O, p)}{\rho(T)} < 1,$$

where $\rho(T)$ is the circumradius of $T$ and $O$ is the center of the circumsphere.

- the anisotropic measure of $\alpha$ is derived from this formula

$$\alpha_M(p)(p, K) = \frac{\|Op\|_{M(p)}}{(r_K)_{M(p)}} < 1.$$

The Delaunay cavity criterion relies on distance evaluations to define the cavity of the point $p$ to be inserted in an existing Delaunay triangulation.

**Lemma 8** The anisotropic Delaunay measure, $\alpha_M(p)(p, K) < 1$, allows to construct a valid, star-shaped with respect to $p$, anisotropic Delaunay cavity.
Quality of a triangulation

Quality measures are needed to measure the discrepancy between the prescriptions and the current values of shape, size and orientation given by \( M(x) \) at any \( x \) in \( \Omega \).

- the geometric shape quality measure of a simplex \( T \) can be defined based on the Jacobian matrix as:

\[
Q_{geo}(T) = \left( \frac{\text{tr}\left((J_T)^t J_T\right)}{n \det\left((J_T)^t J_T\right)} \right)^{\frac{d}{2(d-1)}} = \left( \frac{\|J_T\|_F}{\sqrt{d} \det(J_T)^{\frac{1}{d}}} \right)^{\frac{d}{d-1}}
\]

(166)

where \( \|\cdot\|_F \) denotes the Frobenius matrix norm. Notice that \( Q_{geo}(T) \geq 1 \) and \( Q_{geo}(T) = 1 \) for an equilateral element \( T \). We have, for all \( T \in \mathcal{T}_h \)

\[
\left( \frac{1}{d(d-1)} \left( \left( \frac{\mu_{max}}{\mu_{min}} \right)^{\frac{1}{2d}} - 1 \right)^2 + 1 \right)^{\frac{d}{2(d-1)}} \leq Q_{geo}(T) \leq \frac{\mu_{max}}{\mu_{min}},
\]

(167)

where \( \mu_{min} \) and \( \mu_{max} \) are the minimal and maximal singular values of \( J_K \), respectively. i.e., the geometric quality of \( T \) is equivalent to the aspect ratio of \( T \), for small \( d \).
Quality of a triangulation

- by extension, the geometric quality of $\mathcal{T}_h$ is defined as

$$Q_{geo}(\mathcal{T}_h) = \max_{T \in \mathcal{T}_h} Q_{geo}(T)$$  \hspace{0.5cm} (168)

- similar quality measures can be defined to assess the alignment and equidistribution;

- however, a single measure is better to assess the quality of an element, like

$$Q_{ani}(T) = \alpha \left( \frac{\sum_{i=1}^{k} (e_i, M_T e_i)}{|K| \sqrt{\det(M_T)}} \right)^d$$  \hspace{0.5cm} (169)

where $e_i$ denotes here any of the $k$ edges of $T$ and $\alpha$ is a normalization coefficient such that $Q_{ani}(T) = 1$ for an equilateral element.

- an efficiency index is defined

$$\tau(\mathcal{T}_h) = \exp \left( \frac{1}{N_e} \sum_{i=1}^{N_e} \min(l_i, l_i^{-1}) - 1 \right)$$  \hspace{0.5cm} (170)

where $l_i$ denotes the length of edge $i$ in the current mesh.
Anisotropic triangulation

Consider the domain $\Omega = [0, 1]^2$ and the function $f_1$ defined by:

$$f_1(x, y) = \tanh(-100(y - 0.5 - 0.25 \sin(2\pi x))) + \tanh(100(y - x))$$

Figure 9: Locally adapted anisotropic triangulation for the function. Initial triangulation $|T_h| = 5,417$ (left) and final adapted triangulation $|T_h| = 5,132$ (right)

$$\|f_1 - \pi_h f_1\|_{L^\infty(\Omega)} \approx 0.463$$

$$\|f_1 - \pi_h f_1\|_{L^\infty(\Omega)} \approx 0.00674$$
Anisotropic triangulation

Consider the domain \( \Omega = [0, 1]^2 \) and the function \( f_1 \) defined by:

\[
f_1(x, y) = \tanh(-100(y - 0.5 - 0.25 \sin(2\pi x))) + \tanh(100(y - x))
\]

![Anisotropic triangulation](image)

\|
\| f_1 - \pi_h f_1 \|_{L^\infty(\Omega)} \approx 0.463
\]

\|
\| f_1 - \pi_h f_1 \|_{L^\infty(\Omega)} \approx 0.00674
\]

Figure 10: Locally adapted anisotropic triangulation for the function. Initial triangulation \( |T_h| = 5,417 \) (left) and final adapted triangulation \( |T_h| = 5,132 \) (right)
Anisotropic triangulation

Consider the domain $\Omega = [0, 1]^2$ and the function $f_1$ defined by:

$$f_1(x, y) = \tanh(-100(y - 0.5 - 0.25\sin(2\pi x))) + \tanh(100(y - x))$$

Figure 11: Locally adapted anisotropic triangulation for the function. Initial triangulation $|T_h| = 5,417$ (left) and final adapted triangulation $|T_h| = 5,132$ (right)

$$\|f_1 - \pi_h f_1\|_{L^\infty(\Omega)} \approx 0.463$$

$$\|f_1 - \pi_h f_1\|_{L^\infty(\Omega)} \approx 0.00674$$
**Anisotropic triangulation**

Consider the domain $\Omega = [0, 1]^2$ and the function $f_1$ defined by:

$$f_1(x, y) = \tanh(-100(y - 0.5 - 0.25 \sin(2\pi x))) + \tanh(100(y - x))$$

Figure 12: Locally adapted anisotropic triangulation for the function. Initial triangulation $|T_h| = 5,417$ (left) and final adapted triangulation $|T_h| = 5,132$ (right)

$$\|f_1 - \pi_h f_1\|_{L^\infty(\Omega)} \approx 0.463$$

$$\|f_1 - \pi_h f_1\|_{L^\infty(\Omega)} \approx 0.00674$$
Anisotropic triangulation

Consider the domain $\Omega = [0, 1]^2$ and the function $f_1$ defined by:

$$f_1(x, y) = \tanh(-100(y - 0.5 - 0.25 \sin(2\pi x))) + \tanh(100(y - x))$$

Figure 13: Isotropic triangulation $|T_h| = 5,158$ (left) vs. anisotropic triangulation $|T_h| = 5,132$ (right)

$$\|f_1 - \pi_h f_1\|_{L^\infty(\Omega)} \approx 0.0952$$

$$\|f_1 - \pi_h f_1\|_{L^\infty(\Omega)} \approx 0.00674$$
Mesh adaptation

For adapting triangulations to a given size or metric field, $h$-adaptation is most widely used.

- the main idea consists in generating quasi-uniform meshes in the metric supplied by an error estimate that prescribes the size, shape and orientation of the elements.

- a typical error estimate:
  - suppose we want to solve the problem:
    \[
    \text{Find } u \in W \text{ such that } a(u, v) = f(v), \text{ for all } v \in V,
    \]
    where $V$ and $W$ are Hilbert spaces, $f \in V'$ and $a \in \mathcal{L}(W \times V, \mathbb{R})$.

  - suppose also that the bilinear form $a$ satisfies the hypothesis of the Nečas theorem (well-posed problem).

  - given $\mathcal{T}_h$ of $\Omega$ and approximation spaces $W_h$ and $V_h$, we consider
    \[
    \text{Find } u_h \in W_h \text{ such that } a(u_h, v_h) = f(v_h), \text{ for all } v_h \in V_h,
    \]
    that is also assumed to be well-posed.
Mesh adaptation

- a function \( e(h, u_h, f) \) is called an a posteriori error if it provides an upper bound on the approximation error: \( \| u - u_h \|_W \leq e(h, u_h, f) \). Furthermore, if \( e(h, u_h, f) \) is such that

\[
e(h, u_h, f) = \left( \sum_{T \in T_h} e_T(u_h, f)^2 \right)^{1/2},
\]

then \( e_K(u_h, f) \) is called an error indicator.

- the aim is to obtain an anisotropic bound where the physical derivatives are related to the size, shape and orientation of the elements.

- interest in obtaining estimates for the classical \( L^1 \) and \( L^2 \) norms or the \( H^1 \) seminorm.

- such estimates have been provided for the interpolation error on linear Lagrange finite elements and involve the eigenvalues and eigenvectors of the Jacobian matrix of the affine mapping between the reference element and a mesh element or are based on the Hessian matrix of the solution.
Mesh adaptation

- A local error model can be defined at a vertex $p$ as

$$e_M(p) = \sum_{i=1}^{d} h_i^2 \left| \frac{\partial^2 u}{\partial \alpha_i^2} \right|,$$

where the $\alpha_i$ are the coefficients of the diagonal matrix $D$ and the $h_i$ indicate the local sizes in the directions of the eigenvectors of the Hessian matrix.

- We define a local error indicator for the $L_\infty$ norm as

$$e(T) = \|u - P_h u\|_{\infty,T} \leq C \max_{y \in T} \max_{\vec{c} \subset T} \langle \vec{v}, |D^2 u(y)| \vec{v}\rangle,$$

where $D^2 u$ represents the Hessian matrix of the function $u$ and $C = \frac{1}{2} \left( \frac{d}{d+1} \right)^2$ is a constant independent of $h$. 
Consider a general advection-diffusion problem in dimension one, for $0 < x < L$

$$\left\{ \begin{array}{l}
-\nu \frac{d^2 u}{dx^2}(x) + c \frac{du}{dx}(x) = f(x) \\
u(0) = 0 \quad u(L) = 0
\end{array} \right.$$ 

we consider here $L = 1$, $\nu = 0.01$, $c = 0.5$ and $f(x) = 1$.

the solution exhibits a boundary layer at $L$. 

Optimal point distribution, second order derivative, error graph and solution
Mesh adaptation algorithm

- weak formulation, we take \( v \in H^1_0(\Omega) \)

\[
a(u, v) = \nu \int_0^L u'(x)v'(x)\,dx + c \int_0^L u'(x)v(x)\,dx = \int_0^L f(x)v(x)\,dx
\]

- \( \Omega \) is splitted in \( N \) segments of length \( h \), we have

\[
u_h(x) = \sum_{j=1}^{N-1} u_j w_j
\]

and the system

\[
\sum_{j=1}^{N-1} a(w_j, w_i) u_j = \int_0^L f(x)w_i(x)\,dx, \quad \forall i = 1, \ldots, N - 1
\]

- and the stiffness and mass matrices are

\[
\frac{\nu}{h} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \frac{c}{2} \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} = \frac{h}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}
\]
Mesh adaptation algorithm

- interpolation error, given a discretization \( T_h \)

\[
\|u - P_h u\|_{H^1(\Omega)} = \left( \sum_{T} \|u - P_h u\|_{H^1(\Omega)}^2 \right)^{1/2}
\]

- posing \( h = \max_i |x_i - x_{i-1}| \) yields, for all \( x \in \Omega \)

\[
|u(x) - P_h u(x)| \leq \frac{h^2}{8} \max_{x \in \Omega} |u''(x)|
\]

\[
|u'(x) - P_h' u(x)| \leq \frac{h}{2} \max_{x \in \Omega} |u''(x)|
\]

- and we have

\[
\|u(x) - P_h u(x)\|_{L^2} \leq C\frac{h^2}{8} \max_{x \in \Omega} |u''(x)|
\]

\[
\|u'(x) - P_h' u(x)\|_{H^1} \leq Ch \max_{x \in \Omega} |u''(x)|
\]

- in \( H^1 \) norm: \( h_i \propto 1/\max_{x \in T_i} |u''(x)| \)
Mesh adaptation algorithm

Given a triangulation $\mathcal{T}_h$ and a metric tensor field $M$ defined at the vertices of $\mathcal{T}_h$.

Enforce the desired size variation, do

1. loop over mesh edges (edge analysis)
   compute $l_M(\vec{e})$ for every edge $\vec{e}$
   if $(l_M(\vec{e}) < l_{min})$ then
     collapse $\vec{e}$
   else if $(l_M(\vec{e}) > l_{max})$ then
     split $\vec{e}$ and insert new vertex (Delaunay criterion)
   endif
   while ( $\mathcal{T}_h$ is modified )
2. loop over mesh elements (quality optimization)
   perform edge flips
   perform vertex relocation
   end for
   while ( $\mathcal{T}_h$ is modified ).
Mesh adaptation algorithm

**Remark 13** Some comments on the adaptation scheme:

1. the algorithm always terminates in a finite number of iterations. 
   Mesh modifications are only applied if they result in a mesh quality improvement.

2. the complexity of the procedure is in $O(|T|)$ and the overall complexity is $cO(|T|)$, i.e.,
   linear in number of elements.

Figure 14: Example of mesh adaptation in dimension 2 for a computational fluid dynamics computation around a Naca airfoil. Adapted mesh (left) and corresponding density distribution (right). Notice how the mesh lines follow the isolines of the density function, thanks to the error estimate.
Adaptation scheme

The mesh adaptation loop is composed of 4 steps:

1. **Solver**: compute the solution $u_h$ on the triangulation $\mathcal{T}_h$.
2. **Error indicator**: analyze $u_h$ and supplies a metric $\mathcal{M}$ in order to equidistribute the error.
3. **Adaptation**: generate a triangulation, unit for $\mathcal{T}_h$.
4. **Interpolation**: projects $u_h$ on $\mathcal{T}_h$.

Classical mesh adaptation scheme.
Consider a supersonic jet simulation at Mach 1.6, angle = 3 deg., alt. 45 000 feet. Aircraft size 36 m, mesh size 1 mm to 30 cm; domain size (m): $[-225, 2025] \times [-1200, 1200]^2$.

Figure 15: Volume adapted mesh with $L^2$ norm on the Mach number (3.8 M vertices).
3d mesh adaptation example

Consider a supersonic jet simulation at Mach 1.6, angle = 3 deg., alt. 45 000 feet. Aircraft size 36 m, mesh size 1 mm to 30 cm; domain size (m): $[-225, 2025] \times [-1200, 1200]^2$

Figure 16: Volume adapted mesh with $L^2$ norm on the Mach number (3.8 M vertices).
Consider the inviscid computation around a transonic speed cruise (Falcon, Dassault Aviation).

Figure 17: *Volume adapted mesh with $L^2$ norm on the Mach number (12 M vertices).*
3d mesh adaptation example

Consider the inviscid computation around a transonic speed cruise (Falcon, Dassault Aviation).

Figure 18: Volume adapted mesh with $L^2$ norm on the Mach number (12 M vertices).
3d mesh adaptation example

Consider the inviscid computation around a transonic speed cruise (Falcon, Dassault Aviation).

Figure 19: *Volume adapted mesh with $L^2$ norm on the Mach number (12 M vertices).*
3d mesh adaptation example

Consider the inviscid computation around a transonic speed cruise (Falcon, Dassault Aviation).

Figure 20: Volume adapted mesh with $L^2$ norm on the Mach number (12 M vertices).
Section 4.3

Mesh adaptation for level sets
Level sets to describe evolution

- **A paradigm** for describing surface evolution: use implicit function representation.
  
  Given a domain $\Omega \subset \mathbb{R}^d$, define a function $\phi$ on $\mathbb{R}^d$ such that:
  
  $$\phi(x) < 0, \text{ if } x \in \Omega, \quad \phi(x) = 0, \text{ if } x \in \partial \Omega, \quad \phi(x) > 0, \text{ if } x \in c\Omega.$$  
  
- This representation gives access to intrinsic properties of $\Omega$: normals, fundamental forms;
  
- Suppose that the evolution of the domain $\Omega(t)$ represented by $\phi(t, \cdot)$ is defined by the velocity field $v(t, x) \in \mathbb{R}^d$:
    
    $$\forall t, \forall x \in \mathbb{R}^d, \quad \frac{\partial \phi}{\partial t}(t, x) + v(t, x) \cdot \nabla \phi(t, x) = 0, \quad (171)$$  
    
    often, $v(t, \cdot) = V(t, x) \cdot \frac{\nabla \phi(t, x)}{\|\nabla \phi(t, x)\|}$ is normal to $\partial \Omega(t)$, and (171) becomes a Hamilton-Jacobi equation:
    
    $$\forall t, \forall x \in \mathbb{R}^d, \quad \frac{\partial \phi}{\partial t}(t, x) + V(t, x)\|\nabla \phi(t, x)\| = 0.$$  

- **Interface evolution** is now equivalent to solving a PDE problem.
Solving the advection equation

- Given an interval of time $[t^n, t^{n+1}]$ and $[t^n, t^{n+1}] \times \mathbb{R}^d \ni (t, x) \mapsto v(t, x) \in \mathbb{R}^d$ a velocity field, we consider the advection equation of a scalar value $\phi(t, .)$:

$$\begin{cases}
\frac{\partial \phi}{\partial t}(t, x) + v(t, x) \cdot \nabla \phi(t, x) = 0 & \text{for } (t, x) \in (t^n, t^{n+1}) \times \mathbb{R}^d \\
\phi(t^n, x) = \phi^n(x) & \text{if } x \in \mathbb{R}^d,
\end{cases}$$

(172)

where $\phi^n$ is the (known) scalar value at time $t^n$ ($\phi$ is typically a level set function).

- The interface is the 0-level set of the advected function $\phi$, and need thus to be discretized as a continuous function on the computational domain, for example a $\mathbb{P}^1$-finite element function (no finite volume methods, nor discontinuous Galerkin methods);

- "Classical" finite element methods are known to behave poorly to solve this equation and thus, we use the method of characteristics, pionneered by O. Pironneau.
Numerical method for solving the advection equation

• the characteristic curve emerging from point \( x \in \mathbb{R}^d \) at time \( t \in (t^n, t^{n+1}] \) is the solution \( s \mapsto X(s, t, x) \) to the ODE, for \( t^n < s < t \):

\[
\begin{align*}
\frac{dX}{dt}(s, t, x) &= v(s, X(s, t, x)), \\
X(t, t, x) &= x
\end{align*}
\]  

(173)

• let \( v : [t^n, t^{n+1}] \times \mathbb{R}^d \to \mathbb{R}^d \) be of class \( C^1 \) and such that \( \|v(t, x)\| \leq \kappa (1 + \|x\|) \) \((\kappa > 0)\), then if the initial state \( \phi^n \) is of class \( C^1 \), the advection equation has a unique \( C^1 \) solution over \( \mathbb{R}^d \):

\[
\forall x \in \mathbb{R}^d, \quad \phi(t^{n+1}, x) = \phi^n(X(t^n, t^{n+1}, x)).
\]

• each function \( \phi(t^n, \cdot) \) is approximated by a \( \mathbb{P}^1 \)-finite element function and the desired approximation \( \hat{\phi}^{n+1} \) of \( \phi(t^n, \cdot) \) is obtained as the \( \mathbb{P}^1 \)-finite element function such that for each mesh vertex \( x \):

\[
\hat{\phi}^{n+1}(x) = \hat{\phi}^n(X(t^n, t^{n+1}, x))).
\]
A priori estimate for the advection equation

- consider equation (172) for a period of time $[0, T]$ over the whole space $\mathbb{R}^d$, endowed with simplicial mesh $\mathcal{T}$ (approximations $\tilde{\phi}$ of $\phi(t, x)$ and $\tilde{v}$ of $v(t, x)$).

- denote $V_T$ the space of continuous functions over $\mathbb{R}^d$, whose restriction to every simplex $K \in \mathcal{T}$ is a $\mathbb{P}^1$ polynomial function.

Then, we have the error estimate:

**Theorem 19** Under some hypothesis on the Lipschitz continuity of the fields $v$ and $\tilde{v}$, and on $\phi^0$, such that $\sup_{x \in \mathbb{R}^d} \|v(x) - \tilde{v}(x)\|$ and $\sup_{K \in \mathcal{T}, x \in K} \|\nabla v(x) - \nabla \tilde{v}\|$ are bounded.

Let $\delta t$ be a time step, and consider the sequence $t^n = n\delta t$, $0 = t^0 < t^1 < \cdots < t^N = T$. Denote $\phi^N \in V_T$ the sought approximation of $\phi(T, \cdot)$ defined as

For each node $x$ of $\mathcal{T}$, $\phi^N(x) = \pi_T(\phi^0)(\tilde{X}(0, T, x))$, where $\tilde{X}(0, T, x)$ is the approximation of the solution $X(0, T, x)$ to the ODE (173).

Then, there exists a constant $C$ which only depends on $v$, such that:

$$
\|\phi(T, \cdot) - \phi^N\|_{L^\infty(\mathbb{R}^d)} \leq \|\phi(T, \cdot) - \pi_T\phi(T, \cdot)\|_{L^\infty(\mathbb{R}^d)} + \|\phi^0 - \pi_T\phi^0\|_{L^\infty(\mathbb{R}^d)}
$$

$$
+ \frac{k'}{k} (e^{kT} - 1) \|v - \tilde{v}\|_{L^\infty(\mathbb{R}^d)} + \frac{k'}{k} C T e^C \delta t \delta t
$$

(174)
A priori estimate for Hausdorff distance

• assume that $\phi^0$ is a level set function associated to a regular bounded domain $\Omega^0 \subset \mathbb{R}^d$, and that $\phi^0$ does not admit any critical point in a vicinity of $\partial \Omega^{in}$.

• it follows that for all $t \in [0, T]$, $\Omega(t) := \{ x \in \mathbb{R}^d \setminus u(t, x) < 0 \}$ is a bounded regular domain, with smooth boundary $\partial \Omega(t) := \{ x \in \mathbb{R}^d \setminus u(t, x) = 0 \}$, and $u(t, \cdot)$ does not admit any critical point within a vicinity of $\partial \Omega(t)$.

• denote, for $n = 0, \ldots, N$, $\Omega^n$ and $\partial \Omega^n$ the piecewise affine reconstructions of $\Omega(t^n)$ and $\partial \Omega(t^n)$ obtained as

$$\Omega^n := \{ x \in \mathbb{R}^d \setminus u^n(x) < 0 \} ; \quad \partial \Omega^n := \{ x \in \mathbb{R}^d \setminus u^n(x) = 0 \} \quad (175)$$

**Lemma 9** Let $\phi \in C^1(\mathbb{R}^d)$, without any critical point within a certain tubular neighbourhood $W$ of $\partial \Omega$, so that $\partial \Omega$ is a submanifold of $\mathbb{R}^d$, and $\Omega$ is a bounded subdomain of $\mathbb{R}^d$ with $C^1$ boundary. For any point $x \in W$ we have the estimate:

$$d(x, \partial \Omega) \leq \frac{\sup_{z \in W} \| \nabla \phi(z) \|}{\inf_{z \in W} \| \nabla \phi(z) \|^2} |\phi(x)| \quad (176)$$
A priori estimate for Hausdorff distance

A formal use of the lemma yields:

\[
\sup_{x \in \partial \Omega(T)} d \left( x, \partial \Omega^N \right) \leq \sup_{x \in \partial \Omega(T)} \frac{\sup_{K \in T} \| \nabla \phi^N |_K \|}{\inf_{K \in T} \| \nabla \phi^N |_K \|^2} | \phi^n(x) |
\]

\[
= \sup_{x \in \partial \Omega(T)} \frac{\sup_{K \in T} \| \nabla \phi^N |_K \|}{\inf_{K \in T} \| \nabla \phi^N |_K \|^2} | \phi^N(x) - \phi(T, x) |
\]

\[
\leq \frac{\sup_{K \in T} \| \nabla \phi^N |_K \|}{\inf_{K \in T} \| \nabla \phi^N |_K (z) \|^2} \| \phi^N - \phi(T, \cdot) \|_{L^\infty(\mathbb{R}^d)}
\]

And now, symmetrically:

\[
d^H \left( \partial \Omega(T), \partial \Omega^N \right) \leq \sup \left( \frac{\sup_{z \in \mathbb{R}^d} \| \nabla \phi(T, z) \|}{\inf_{z \in \mathbb{R}^d} \| \nabla \phi(T, z) \|^2}, \frac{\sup_{K \in T} \| \nabla \phi^N |_K \|}{\inf_{K \in T} \| \nabla \phi^N |_K \|^2} \right) \| \phi^N - \phi(T, \cdot) \|_{L^\infty(\mathbb{R}^d)}.
\]

(177)

The estimate of Theorem (19) allows for a control over the discrepancy, measured in terms of Hausdorff distance, between \( \partial \Omega(T) \) and its piecewise affine approximation \( \partial \Omega^N \).
Simplicial mesh adaptation

• the goal of mesh adaptation is to modify a given mesh $\mathcal{T}$ in such a way its elements' size, orientation and eccentricity allow to perform the computation with optimal efficiency.

• since the work of Hecht-Vallet, the idea of metric-based mesh adaptation has been increasingly popular: the local desired size, shape and orientation related information at a node $x$ of mesh $\mathcal{T}$ are stored in a Riemannian metric tensor field $M(x)$.

• given $M(x)$, defined at each point $x \in \mathbb{R}^d$, we consider respectively the length $l_M(\gamma)$ of a curve $\gamma : [0, 1] \rightarrow \mathbb{R}^d$, the volume $V_M(K)$ of a simplex $K$, and the distance $d_M(x, y)$ between two points $x, y \in \mathbb{R}^d$ in the Riemannian space $(\mathbb{R}^d, M)$:

$$l_M(\gamma) = \int_0^1 \sqrt{\langle M(\gamma(t))\gamma'(t), \gamma'(t) \rangle} dt, \quad V_M(K) = \int_K \sqrt{\det(M(x))} dx,$$

$$d_M(x, y) = \inf_{\gamma \in C([0,1],\mathbb{R}^d)} l_M(\gamma).$$
Simplicial mesh adaptation

- the aim is to modify the mesh $\mathcal{T}$ so as to make it quasi-unit with respect to the metric $M(x)$, i.e., all the simplices $K$ have edges lengths lying in $[\frac{1}{\sqrt{2}}, \sqrt{2}]$.

- we expect the anisotropic quality measure:

$$Q_M(K) := \alpha_d \frac{V_M(K)^2}{\left(\sum_{i=1}^{na} l_M(e_i)^2\right)^{d}}$$

of all elements of the mesh (where $na = d(d + 1)/2$ is the number of edges of a $d$-dimensional simplex, $e_i$ are the edges of $K$ and $\alpha_d$ is a normalization factor) to be as close to 1 as possible.

- for our purposes, it is desirable to adapt a mesh at the same time to several a priori independent information, supplied by two (or more) metric tensor fields $M_1, M_2$. This is classically achieved resorting to a so-called metric intersection procedure: operating on the simultaneous reductions of $M_1(x)$ and $M_2(x)$ at any point $x \in \mathbb{R}^d$. 
to achieve such a mesh, we recall the classical $L^\infty$ error estimate for the Lagrange finite element $\mathbb{P}^1$-interpolation error of a function $u$ of class $C^2$:

**Theorem 20** Let $\mathcal{T}$ a simplicial mesh of $\mathbb{R}^d$ (or a polyhedral subset of it) and $u$ a $C^2$ function on $\mathbb{R}^d$. Then for every simplex $K \in \mathcal{T}$,

$$
||u - \pi_{\mathcal{T}}u||_{L^\infty(K)} \leq \frac{1}{2} \left( \frac{d}{d+1} \right)^2 \max_{x \in K} \max_{y,z \in K} \langle |\mathcal{H}(u)|(x)yz, yz \rangle
$$

where $\mathcal{H}(u)$ is the Hessian matrix of $u$ and, for a symmetric matrix $S \in S_d(\mathbb{R})$ which admits a diagonal shape in orthonormal basis $S = P \text{diag}(\{\lambda_i\}_{1 \leq i \leq d})^t P$.

a mesh $\mathcal{T}$ suitable for the Lagrange $\mathbb{P}^1$-interpolation of a smooth function $u$ (such that $||u - \pi_{\mathcal{T}}u||_{L^\infty(K)} \leq \epsilon$) can be obtained as a quasi-unit mesh for the metric $M_u$ defined at each node $x$ of $\mathcal{T}$. 


The adaptation method

- we intend to use the error estimates (174) and (177) to infer a mesh adaptation method which allows for a good discrete approximation $\partial \Omega_T$ of $\partial \Omega(T)$.

- From formula (174), the control over $\|u(T, .) - u_T\|_{L^\infty(\mathbb{R}^d)}$ (or $d^H(\partial \Omega(T), \partial \Omega_T)$) consists of three independent contributions:
  1. the term (we try to dampen by adaptation):

     $$\left(\|\phi(T, .) - \pi_T \phi(T, .)\|_{L^\infty(\mathbb{R}^d)} + \|\phi^{in} - \pi_T \phi^{in}\|_{L^\infty(\mathbb{R}^d)}\right)$$

     is only linked to the way mesh $T$ is adapted to the interpolation of $\phi(T, .)$ and $\phi^0$.
  2. the second term:

     $$\frac{k'}{k} C T e^{C \delta t} \delta t$$

     is solely related to the time discretization of interval $[0, T]$ with the substep $\delta t$.
  3. and the last term:

     $$\frac{k'}{k} (e^{kT} - 1) \|v - \tilde{v}\|_{L^\infty(\mathbb{R}^d)}$$

     is connected to the quality of the approximation of velocity field $v$. 
A simple algorithm for mesh adaptation

The objective is to adapt mesh $\mathcal{T}$ so that both interpolation errors $\|u(T, .) - \pi_T u(T, .)\|_{L^\infty(\mathbb{R}^d)}$ and $\|u^{in} - \pi_T u^{in}\|_{L^\infty(\mathbb{R}^d)}$ are made small. This is actually very intuitive.

1: Start with an approximation $\phi_0$ (e.g. $\mathbb{P}^1$-interpolate) of function $\phi^0$ on mesh $\mathcal{T}$.
2: for $k = 0, ..., m - 1$ do
3: if $k = 0$ then
4: Set $\mathcal{T}^0 = \mathcal{T}$ and $\tilde{\phi}^{0,0} = \phi_0$.
5: end if
6: Solve the advection equation with velocity field $V$ and initial state $\tilde{\phi}^{0,k}$, over $[0, T]$, on $\mathcal{T}^k$.
7: Compute the intersected metric $M^k := M_{\tilde{\phi}^k_0} \cap M_{\tilde{\phi}^k_T}$
8: Adapt $\tilde{T}^k$ with respect to $M^k$.
9: if $k + 1 < m$ then
10: Project $\phi_0$ on mesh $\mathcal{T}^{k+1}$.
11: else
12: Project $\tilde{\phi}^k_T$ on $\mathcal{T}^{k+1}$.
13: Adapt $\mathcal{T}^m$ with respect to $\tilde{\phi}^m_T$, and project this function on the final mesh.
14: end if
15: end for
16: return ($\mathcal{T}_T, \phi_T$)
Mesh adaptation

• in all applications we have in mind, we are interested in modifying mesh $\mathcal{T}$ so that it becomes adapted to both Lagrange $\mathbb{P}^1$ interpolation of functions $\phi^0$ and $\phi(T,.)$.

• such a mesh is built as a quasi-unit mesh according to the intersection $M_{\phi^0} \cap M_{\phi(T,.)}$ of metrics $M_{\phi^0}$ and $M_{\phi(T,.)}$, respectively adapted to $\phi^0$ and $\phi(T,.)$.

Figure 21: Rotation of Zalesak’s slotted disk of angle $\frac{\pi}{4}$: (a) adapted mesh at time $t = \pi$, (b) adapted mesh at time $\frac{5\pi}{8}$, (c) mesh adapted to both interfaces $\partial \Omega^0$ and $\partial \Omega(T)$ (displayed in red), (d) zoom on the mesh.
1. **mesh gradation** control: the mesh $\mathcal{T}$ is adapted with respect to the metric $M_\phi$ defined on a neighborhood of $\partial\Omega := \{ x \in \mathbb{R}^d \setminus \phi(x) = 0 \}$, and $\mathcal{T}$ to show very large isotropic elements ‘far’ from $\partial\Omega$. Hence, we impose a control over mesh gradation near the interface:

$$\frac{1}{r} \leq \frac{l_M(e_1)}{l_M(e_2)} \leq r.$$ 

2. **redistancing**: too steep or too loose variations in the level sets of the function $\phi(t, .)$ under evolution may jeopardize the accuracy of the computation. To overcome this feature, $\phi(t, .)$ is restored as the signed distance function to its 0-level set $\partial\Omega(t)$ at least near $\partial\Omega(t)$.

Here we replace the computed approximation $\phi_T$ of $\phi(T, .)$ by the signed distance function $\tilde{\phi}_T$ to $\partial\Omega_T$. Given a small time step $dt$, $\tilde{\phi}_T$ is computed as the steady state of the sequence of $\mathbb{P}^1$-FE functions $\phi^n$, based on the properties of the unsteady Eikonal equation.
Numerical examples of interface capturing

We deal now with two classes of examples in 2 and 3 dimensions:

1. solving **advection equation** on classical analytical examples and reconstructing the 0-level set of the function $\phi$; 
   the objective is to evaluate the accuracy of the computation (reconstruction);

2. capturing an interface evolving under an external velocity field $v(t, x)$ prescribed by a fluid flow problem; 
   the aim is to show the coupling between a Stokes problem and the advection equation.

2d examples were run on laptop, 2.66 Ghz (4 Go), and 3d examples on an Opteron 2.1 Ghz.
Classical analytical problems

We intend to solve the advection equation on a domain \( \Omega(t) \subset \mathbb{R}^d \).

- An initial interface \( \partial \Omega^{in} \) evolves under a (deforming) velocity field \( V \), \( 0 \leq t \leq T \), so that the final surface \( \partial \Omega(T) \) coincides with \( \partial \Omega^{in} \).

- All functions are approximated by \( \mathbb{P}^1 \) Lagrange finite element functions, and the ODE (5) is discretized with a 4-th order Runge-Kutta method. A redistancing procedure is used at each step, even for the cases that do not theoretically introduce any distortion of the level sets of the evolving function.

- The accuracy of the computation is evaluated in terms of the Hausdorff distance between the numerical interface associated with \( \partial \Omega^{in} \) and the computed interface \( \partial \Omega(T) \). Although no particular attention is paid to mass conservation, the loss of mass between initial and final step is also computed.
Rotation of Zalesak’s slotted disk

Consider a disk of radius 0.15 centered at \((0.5, 0.75)\), with a slot of length 0.25, in a unit square, submitted to a uniform rotation of center \((0.5, 0.5)\), corresponding to:

\[
V(t, x, y) = \left(\frac{-y}{x - 0.5}\right), \quad 0 \leq t \leq T = 2\pi.
\]

The time interval is subdivided into 8 time periods on which algorithm 1 is applied.

- **Purpose**: assessment of the well-preservation of the sharp features of the interface throughout the evolution.
- **Error estimates**: (see Table 3)
  1. Hausdorff distance between the initial and final interfaces;
  2. \(L^1\)-error measure: \(E_{sd}(\Omega^{in}, \Omega_T) := \left|(\Omega^{in} \cup \Omega_T) \setminus (\Omega^{in} \cap \Omega_T)\right|\);
  3. \(L^\infty\)-error measure between numerical level-set functions \(u^{in}\) and \(u_T\):

\[
E_\infty(u^{in}, u_T) := \sup_{K \in \mathcal{K}} \|u^{in} - u_T\|,
\]

where \(K \subset \mathcal{K}\) denotes the set of simplices crossed by the 0-level set of \(u^{in}\) or \(u_T\).
Rotation of Zalesak’s slotted disk

Figure 22: Superposition of the slotted disk after one complete revolution (blue line) over the initial one (red line), and zoom on the surfaces (middle and right).

<table>
<thead>
<tr>
<th>Test</th>
<th>ε</th>
<th>h_{min}</th>
<th>np</th>
<th>Volume loss (%)</th>
<th>Volume loss (initial vol.)</th>
<th>d^H(∂Ω^in, ∂Ω_T)</th>
<th>E_{sd}(∂Ω^in, ∂Ω_T)</th>
<th>E_∞(u^{in}, ua_T)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1e^{-3}</td>
<td>1e^{-3}</td>
<td>5,699</td>
<td>−0.889</td>
<td>2.08e^{-3}</td>
<td>3.36e^{-3}</td>
<td>2.60e^{-3}</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1e^{-4}</td>
<td>1e^{-4}</td>
<td>14,477</td>
<td>−0.299</td>
<td>8.60e^{-4}</td>
<td>2.18e^{-3}</td>
<td>1.34e^{-3}</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Details on the two-dimensional test of Zalesak’s slotted disk.
Rotation of Zalesak’s sphere

Very similar to the 2d case, consider a sphere of radius 0.15 centered at \((0.5, 0.5, 0.25)\), with a slot of width 0.15, in a unit cube, submitted to a uniform rotation w/r to the \(x\)-axis:

\[
V(t, x, y) = \begin{pmatrix}
0 \\
-(z - 0.5) \\
(y - 0.5)
\end{pmatrix}, \quad 0 \leq t \leq T = 2\pi.
\]

We set parameters \(\epsilon = 0.005\), \(h_{\text{min}} = 0.005\) and the computational mesh has about 200,000 points \((\approx 1,200,000\) tetrahedra).

Figure 23: Elargement on initial Zalesak’s sphere and on final sphere after a rotation of \(2\pi\). Cut through the adapted mesh at times \(t = T/2\) and \(t = 5T/8\).
2d deformation test flow

A more serious test case for shape distortion is now proposed. In a unit square domain, a disk of radius 0.15, centered at (0.5, 0.75) is evolved according to the following velocity field:

\[
V(t, x, y) = \begin{pmatrix}
\sin(4\pi(x + \frac{1}{2}))\sin(4\pi(y + \frac{1}{2}))\cos\left(\frac{\pi t}{T}\right) \\
\cos(4\pi(x + \frac{1}{2}))\cos(4\pi(y + \frac{1}{2}))\cos\left(\frac{\pi t}{T}\right)
\end{pmatrix},
\]

for \(0 \leq t \leq T = 3\), with periodicity of the domain w/r top and bottom sides.

Figure 24: (a) Maximum elongation step \((t = T/2)\), (b) corresponding 0-level set, (c) intersected mesh, adapted to both steps \(t = 0.9\) and \(t = 1.2\), (d) zoom on the intersected mesh (the surface associated to \(t = 1.2\) is displayed in red).
3d deformation test flow

- in a unit cube, a sphere of radius 0.15, centered at (0.35, 0.35, 0.35) is made evolved according to the following velocity field:

\[
V(t, x, y, z) = \begin{pmatrix}
2\sin^2(\pi x)\sin(2\pi y)\sin(2\pi z)\cos\left(\frac{\pi t}{T}\right) \\
-\sin(2\pi x)\sin^2(\pi y)\sin(2\pi z)\cos\left(\frac{\pi t}{T}\right) \\
-\sin(2\pi x)\sin(2\pi y)\sin^2(\pi z)\cos\left(\frac{\pi t}{T}\right)
\end{pmatrix}
\]

for \(0 \leq t \leq T = 3\), time interval is split into 10 subperiods.

- results presented in next figure correspond to a computation held with isotropic adaptation with a minimum size parameter \(h_{min} = 0.002\) (which amounts to anisotropic adaptation with very small precision parameters). The largest mesh of the computation is worth 3,763,497 vertices, and the whole computation took about 21 hours.

- the final interface is not exactly as smooth as the initial one, notably near its horizontal diameter; actually, this area corresponds to the most stretched zone at the maximum elongation time \(t = T/2\), and is the most difficult to track accurately.
Figure 25: Deformation test case: sequence of computed isosurfaces.
Figure 11. 3d deformation test case: sequence of computed surfaces.

Figure 12. A cut in the most stretched interface, at time $t = 1.5$ (left); a zoom on the cut (right).

7. APPENDIX

Here is a variation of classical Gronwall’s lemma for the estimation of the discrepancy between the solutions of two ODEs associated to different vector fields (see [37] for proof).

Lemma 7.1

Let $V_1, V_2 : \mathbb{R}^{d+1} \rightarrow \mathbb{R}^d$ two vector fields, $V_1$ being Lipschitz continuous in the second variable.
Figure 27: Deformation test case: a cut in a mesh adapted to the evolving interface at time $t = 1.5$ using isotropic mesh adaptation ($\approx 1,500,000$ points) and anisotropic mesh adaptation ($\approx 700,000$ points).
Viscous flow model equations

We assume that, at each time step $t$, the flow of each fluid is governed by the quasi-static incompressible Stokes equations, written in each subdomain $\Omega^i$:

$$
\begin{cases}
\rho^i \frac{\partial u^i}{\partial t} - \mu^i \Delta u^i(t, x) + \nabla p^i(t, x) = \rho^i f^i(t, x), \\
div u^i(t, x) = 0
\end{cases}
$$

where $u^i(t, .)$, $p^i(t, .)$ are the velocity and pressure in $\Omega^i(t)$ and $f^i(t, .)$ is a body force.

- Introducing the stress tensor $\sigma = \mu (\nabla u + \nabla^t u) - pI$, the boundary conditions read:

$$
\begin{cases}
u^1 - u^2 = 0 \text{ on } \partial\Omega(t) \quad \text{(velocity continuity)} \\
(\sigma^1 - \sigma^2).n^1 = -\gamma \kappa n^1 \text{ on } \partial\Omega(t) \quad \text{(surface tension for the normal constraint)}
\end{cases}
$$

and for the boundary of the computational domain $\Sigma$:

$$
u = u_D \text{ on } \partial\Omega_D(t); \quad \nu = u_N \text{ on } \partial\Omega_N(t);
$$

$$
\partial\Omega_D(t) \cap \partial\Omega_N(t) = \emptyset; \quad \partial\Omega_D(t) \cup \partial\Omega_N(t) = \partial\Sigma
$$

- Initial conditions read:

$$
\begin{cases}
\partial\Omega(0) = \partial\Omega_0 \text{ given} \\
u(0, .) = u_0 \text{ given on } \Sigma
\end{cases}
$$
Bifluid interface evolution

At each time $t \in [0, T]$, $\Gamma(t)$ is represented by a level set function $\phi(t)$. $\phi$ is solution of the following level set equation:

$$
\begin{align*}
\frac{\partial \phi}{\partial t}(t, x) + u(t, x) \cdot \nabla \phi(t, x) &= 0 \forall (t, x) \in (0, T) \times \Sigma \\
\phi(0, x) &= \phi_0(x) \forall x \in \Sigma
\end{align*}
$$

where $u(t, x) = u^1(t, x)$ if $x \in \Omega^1(t)$, $u(t, x) = u^2(t, x)$ otherwise, is solution of (178).

This model being far too difficult to solve as it reads, we propose to discretize the time interval $0 = t_0 < \ldots < t_N = T$, and to solve for each time $t_n < t < t_{n+1}$ the approximated problem

$$
\begin{align*}
\frac{\partial \phi}{\partial t}(t, x) + u(t^n, x) \cdot \nabla \phi(t, x) &= 0 \quad \forall (t, x) \in (t^n, t^{n+1}) \times \Sigma \\
\phi(t^n, x) &= \phi^n(x) \quad \forall x \in \Sigma
\end{align*}
$$

where $\phi^n$ denotes the approximation of $\phi(t^n, .)$ at time $t^n$.

This allows to implement the pure advection case described.
Coalescence of droplets

Illustrate the change of topology of an interface \( \partial \Omega(t) \) in \( \Omega = [-0.6, 0.6]^2 \).

The model consists in two droplets of a viscous fluid \((\mu^2 = 10^{-1} \text{ kg/m s})\) embedded in a less viscous fluid \((\mu^1 = 10^{-3} \text{ kg/m s})\) in a zero-gravity medium.

The surface tension coefficient is \( \gamma = 0.03 \text{ N/m} \).

The droplets have initial elliptic shapes: \( a_1 = 0.4 \text{ mm}, b_1 = 0.2 \text{ mm} \), centered at \((-0.1, 0)\), and \( a_2 = 0.23 \text{ mm}, b_2 = 0.15 \text{ mm} \), centered at \((0.27, 0)\).

The incompressibility condition makes the fluids hard to coalesce.

Mesh size is \( h_{min} = 0.005 \) and \( T^n_h \) contains about 1,500 nodes.

\( V_0 = 0.358937 \) \((V_{ex} = 0.359712)\), corresponds to that of a circle of radius \( r = 0.338 \).

Final volume of the drop is \( V = 0.348124 \) and the loss of mass is \(|V - V_0|/V_0 \approx 0.03\).

\( X_{g0} = (0.004097, 0) \) and \( X_{gf} = (0.002775, 0) \), \( \delta \approx 1.3 \times 10^{-3} \).
Coalescence of two droplets (2)

$\begin{array}{ccc}
t = 0 \text{ ms} & t = 0.016 \text{ ms} & t = 0.032 \text{ ms} \\
t = 0.048 \text{ ms} & t = 0.08 \text{ ms} & t = 0.36 \text{ ms}
\end{array}$
Coalescence of two droplets (3)

\begin{align*}
t & = 0 \text{ ms} \\
t & = 0.016 \text{ ms} \\
t & = 0.032 \text{ ms} \\
t & = 0.048 \text{ ms} \\
t & = 0.08 \text{ ms} \\
t & = 0.36 \text{ ms}
\end{align*}
Rising bubble

We consider the rise of a gas bubble in a fluid under gravity and the coalescence of the bubble with its bulk phase.

Here, $\Omega = [0, 4] \times [0, 5]$, the bubble is centered at $(2.0, 1.5)$ with a radius of 0.5 and the upper interface is at $y = 3$.

Parameters are set as follows:

- fluid: $\rho^1 = 100 \text{ kg/m}^3$, $\mu^1 = 10^{-1} \text{ kg/(m s)}$;
- gas: $\rho^2 = 1 \text{ kg/m}^3$, $\mu^2 = 10^{-2} \text{ kg/(m s)}$;
- surface tension: $\gamma = 6 \times 10^{-5} \text{ N/m}$ and
- gravity: $g = 9.81 \text{ m/s}^2$.

Boundary conditions are:

- $\mathbf{u} = 0$ on the lower wall;
- free-slip condition on the vertical walls: $\tau \cdot \sigma \cdot \mathbf{n} = 0$ and $\mathbf{u} \cdot \mathbf{n} = 0$;
- $\sigma \cdot \mathbf{n} = 0$ on the top wall.
Rising bubble (2)

Mesh adaptation for level sets
Rising bubble (3)

Figure 7.10: Surface position as a function of time for the rising bubble.

Figure 7.11: Diameter of the neck $R_n$ as a function of time after the connection of the bubble and its bulk phase.
This is the end...
This is the end...

"You're not allowed to use the sprinkler system to keep your audience awake."
Appendix

Variational approximation of linear problems
Linear functionals

• We introduce a linear continuous functional \( \ell : V \rightarrow \mathbb{R} \), i.e. an element \( \ell \) of the topological dual space \( V' = \mathcal{L}(V, \mathbb{R}) \) of \( V \), endowed with the dual norm:

\[
\| \ell \|_{V'} = \sup_{v \in V, v \neq 0} \frac{\ell(v)}{\| v \|}.
\]

• One of the most remarkable result in Hilbert spaces theory, states that the dual space \( V' \) of continuous linear functionals on \( V \) can be identified with \( V \) via the inner product \((\cdot, \cdot)\) in \( V \):

**Theorem 21 (Riesz' Representation Theorem)** Let \( \ell \in V' \) be a continuous linear functional on \( V \). Then there is a unique element \( u \) of \( V \) for which

\[
\ell(v) = (u, v), \quad \text{for all } v \in V.
\]

In addition,

\[
\| \ell \| = \| u \|.
\]
Adjoint operators

The first application of the Riesz representation theorem is to determine the existence of the adjoint of a linear transformation.

- Assume $V$ and $W$ are Hilbert spaces and $L \in \mathcal{L}(V, W)$ is a linear operator. We use theorem 21 to define a new operator $L' : W \to V$, called the adjoint of $L$.

- Given $w \in W$, define a linear function $\ell_w \in V'$ by
  \[ \ell_w(v) = (Lv, w)_W , \quad \forall v \in V . \]

- Riesz’ representation theorem states that there is a unique element (the adjoint), denoted $L'(w) \in V$ such that
  \[ \ell_w(v) = (Lv, w)_W = (v, L'(w))_V , \quad \forall v \in V , w \in W . \]

- In the particular case $V = W$ and $L = L'$, $L$ is called a self-adjoint operator. When $L$ is a self-adjoint operator from $\mathbb{R}^n$ to $\mathbb{R}^n$, it is represented by a symmetric matrix in $\mathbb{R}^{n \times n}$. Equations of the form $Lv = w$ involving a self-adjoint operator occur in many physical settings.
Linear variational problems

Linear spaces and linear operators provide a convenient setting for the analysis of such problems that are also interesting for the analysis of nonlinear operators.

- We consider the following formulation:

$$\text{Find } u \in V, \text{ such that } a(u, v) = \ell(v), \quad \forall v \in V.$$  \hspace{1cm} \text{(179)}

- The Lax-Milgram lemma provides an answer about the existence and uniqueness of the solution of such problem.

- Consider a linear operator $A : V \to V'$ and a bilinear form $a : V \times V \to \mathbb{R}$ such that

$$ (Au, v) = a(u, v), \quad \forall u, v \in V. $$ \hspace{1cm} \text{(180)}

**Theorem 22** Relation (180) provides a one-to-one correspondence between linear continuous operators $A : V \to V'$ and continuous bilinear forms $a(\cdot, \cdot)$ defined on $V \times V$. 
Linear variational problems

Many properties of the bilinear form $a(\cdot, \cdot)$ can be deduced from those of the linear operator $A$, and conversely, e.g.:

- $a$ is bounded (i.e. $a(u, v) \leq M\|u\|\|v\|$, for all $u, v \in V$) if and only if $A$ is bounded (i.e. $\|Au\| \leq M\|v\|$, for all $v \in V$).

- $a$ is positive (i.e. $(a(v, v) \geq 0$, for all $v \in V, v \neq 0)$ if and only if $A$ is positive (i.e. $(Av, v) \geq 0$, for all $v \in V$).

- $a$ is strictly positive (i.e. $a(v, v) > 0$, for all $v \in V, v \neq 0$) if and only if $A$ is strictly positive (i.e. $(Av, v) > 0$, for all $v \in V, v \neq 0$).

- $a$ is strongly positive or $V$-elliptic (i.e. $a(v, v) \geq \alpha\|v\|^2$, for all $v \in V, \alpha > 0$) if and only if $A$ is strongly positive (i.e. $(Av, v) \geq \alpha\|v\|^2$, for all $v \in V$).

- $a$ is symmetric (i.e. $a(u, v) = a(v, u)$ for all $u, v \in V$) if and only if $A$ is symmetric (i.e. $(Au, v) = (Av, u)$ for all $u, v \in V$).
A minimization principle

**Theorem 23** Suppose $K \subset V$ is a non-empty, closed and convex subset of the Hilbert space $V$, $a(\cdot, \cdot)$ is a bilinear, symmetric, bounded and $V$-elliptic form on $V$ and $\ell \in V'$. Let

$$J(v) = \frac{1}{2}a(v, v) - \ell(v), \quad v \in V.$$  

Then, there exists a unique minimizer $u \in K$ such that

$$J(u) = \inf_{v \in K} J(v),$$  \hfill (181)

which is also the unique solution of the variational inequality

$$u \in K, \quad a(u, v - u) \geq \ell(v - u), \quad \forall v \in K,$$  \hfill (182)

or equivalently, if $K$ is a subspace of $V$

$$u \in K, \quad a(u, v) = \ell(v), \quad \forall v \in K.$$  \hfill (183)

**Remark 14** When the bilinear form $a(\cdot, \cdot)$ is not symmetric, there is no longer a corresponding minimization problem.
**Well-posed problems**

**Definition 14** The problem (179) is said to be well-posed if there is a unique solution to this problem and if the following linear stability property is satisfied:

\[
\text{there exists } c > 0 ; \text{ for all } \ell \in V', \quad \|u\|_V \leq c \|\ell\|_{V'}.
\]

The next result will allows to attest the well-posedness character of a variational problem.

**Lemma 10 (Lax-Milgram)** Let \( V \) be a Hilbert space. Suppose \( a(\cdot, \cdot) \) is bounded, \( V \)-elliptic bilinear form on \( V \), \( \ell \in V' \). Then, there is a unique solution to the problem

\[
u \in V , \quad a(u, v) = \ell(v) , \quad \forall v \in V .
\]

Furthermore, the following estimate holds:

\[
\text{for all } \ell \in V' , \quad \|u\|_V \leq \frac{1}{\alpha} \|\ell\|_{V'}.
\]
Corollary 3 \hspace{0.5cm} Under the same hypothesis of theorem 10 and assuming that the bilinear form $a(\cdot, \cdot)$ is symmetric, then the solution $u$ to problem (184) is characterized by the property
\[ u \in V, \quad J(u) = \frac{1}{2} a(u, u) - \ell(u) = \inf_{v \in V} \left( \frac{1}{2} a(v, v) - \ell(v) \right). \tag{185} \]

Remark 15 \hspace{0.5cm} (i) When the bilinear form $a(\cdot, \cdot)$ is symmetric, the problem (184) corresponds to the minimization of a quadratic functional on a Hilbert space $V$, which is the abstract formulation of numerous problems in calculus of variations. This explains why this problem is called a variational problem.

(ii) When the bilinear form $a(\cdot, \cdot)$ is symmetric and $V$-elliptic, the Lax-Milgram theorem indicates that the optimization problem $\inf_{v \in V} J(v)$ has a unique solution. The $V$-ellipticity of $a$ can be seen as a property of strong convexity of the functional $J$.

(iii) In many physical applications, the functional $J$ corresponds to an energy term (e.g. the deformation of an elastic membrane).
Internal approximation of variational problems

• We consider the variational problem (179)

\[
\text{Find } u \in V, \text{ such that } a(u, v) = \ell(v), \quad \forall v \in V. \tag{186}
\]

and assume that all hypothesis of Lax-Milgram theorem 10 are satisfied: \(a(\cdot, \cdot)\) is a bounded, \(V\)-elliptic bilinear form on \(V\), \(\ell \in V'\) is a continuous linear form.

• In order to obtain a numerical approximation of the solution \(u\), we will consider replacing this problem by a discrete problem posed in a functional space of finite dimension.

• General principle:
  1. introduce an approximation: sequence of subspaces of increasing finite dimension.
  2. the resulting discrete problem is then more easy to solve than the initial one.
  3. finally, we consider the limit case when the dimension of the subspaces tends toward infinity to construct a solution to the initial problem.
Variational problems

Internal approximation of variational problems

- Consider a family \( \{V_h\} \) of closed subspaces of the infinite dimensional Hilbert space \( V \), for a given parameter \( h \) related to the discretization of the domain and intended to tend towards zero in the analysis.

- The notion of internal approximation is justified by the choice of subspaces \( V_h \subset V \), for all \( h \).

- We assume that for every \( v \in V \), there exists an element \( r_hv \in V_h \) such that

\[
\lim_{h \to 0} \|r_hv - v\| = 0.
\]

- The bilinear form \( a(\cdot, \cdot) \) and the linear form \( \ell(\cdot) \) are defined on \( V_h \times V_h \) and \( V_h \), respectively and the problem (179) can be approximated by the following discrete problem:

\[
\text{Find } u_h \in V_h, \text{ such that } a(u_h, v_h) = \ell(v_h), \quad \forall v_h \in V_h.
\]

In this case, the discrete problem is called a Galerkin approximation.
Internal approximation of variational problems

**Theorem 24** Under the hypothesis of theorem 10, problem (187) has a unique solution $u_h$ in $V_h$ and the following estimate holds for all $\ell \in V'$:

$$\|u_h\|_V \leq \frac{1}{\alpha} \|\ell\|_{V'}.$$ 

**Remark 16** Usually in applications, it is necessary to replace, using numerical integration, the bilinear form $a(\cdot, \cdot)$ and the linear form $\ell(\cdot)$ respectively by a bilinear form $a_h(\cdot, \cdot)$ and a linear form $\ell_h(\cdot)$.

Furthermore, for obvious reasons, we have the following result:

**Corollary 4** Under the hypothesis of theorem 24 and assuming the bilinear form $a(\cdot, \cdot)$ is symmetric, then the solution $u_h$ to problem (187) is characterized by the property:

$$u_h \in V_h, \quad J(u_h) \leq J(v_h), \quad \text{for all } v_h \in V_h,$$

where the functional $J(\cdot)$ is defined as previously by (185).
Convergence and a priori estimates

- Since we have replaced a variational problem by an approximation of this problem, it is interesting to question the error related to the substitution of $V$ by a subspace $V_h$.

**Proposition 13** Under the previous hypothesis, we have the orthogonality identity:

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

- If $a(\cdot, \cdot)$ is a continuous, $V$-elliptic and symmetric bilinear form on $V \times V$ of ellipticity constant $\alpha$, then it defines an inner product and an energy norm associated to it as:

$$\|u\|_e = (a(u, u))^{1/2}, \quad \text{for all } v \in V$$  \hspace{1cm} (189)

which is equivalent to the natural norm of $V$:

$$\alpha^{1/2} \|u\|_V \leq \|u\|_e \leq \|A\|^{1/2} \|u\|_V, \quad \text{for all } u \in V.$$  

- The approximate solution $u_h$ is the orthogonal projection for the inner product of $a(\cdot, \cdot)$ of the solution $u$ onto the subspace $V_h$.  

Convergence and a priori estimates

- The following result helps to understand why the variational approximation method is so interesting.

**Lemma 11 (Céa)** Let $u$ (resp. $u_h$) be the solution of problem (179) (resp. (187)). We have the error estimate

$$
\|u_h - u\| \leq \frac{\|A\|}{\alpha} \inf_{v_h \in V_h} \|v_h - u\| ,
$$

(190)

*If the bilinear form $a(\cdot, \cdot)$ is symmetric, then the previous estimate becomes*

$$
\|u_h - u\| \leq \left(\frac{\|A\|}{\alpha}\right)^{1/2} \inf_{v_h \in V_h} \|v_h - u\| .
$$

(191)

- It provides an optimal estimate of the error between the exact solution $u$ of the problem (179) and the approximate solution $u_h$ of the problem (187).
Convergence and a priori estimates

- This lemma shows that the evaluation of the error is equivalent to the evaluation of the quantity $\inf_{v_h \in V_h} \|v_h - u\|$, where $V_h$ is a subspace of $V$.

- It consists in evaluating the distance in $V$ between the solution $u$ of the problem (179) and the subspace $V_h$ of $V$.

- It is useful to obtain error estimates.

**Theorem 25** Suppose the hypothesis on $V$, $a$ and $\ell$ are those of the previous sections. Suppose that there exists a subspace $V \subset V$, dense in $V$ and a linear mapping $r_h : V \to V_h$ such that

$$\lim_{h \to 0} \|r_h v - v\| = 0,$$

for all $v \in V$, then, the approximation method converges, i.e.

$$\lim_{h \to 0} \|u_h - u\| = 0,$$

where $u$ (resp. $u_h$) is the solution of problem (179) (resp. (187)).
Quadrature formulas for triangles

weight are derived from the Lagrange basis polynomials (depend only on the $x_i$ and not on the function $f$)

- Newton-Cotes rules use edge midpoints $x_{ij}$ and center of mass $x_G$
  
  $$x_{ij} = \frac{x_i + x_j}{2} \quad \text{and} \quad x_G = \frac{x_1 + x_2 + x_3}{3}$$

- computation of $\int_T f(x) \, dx$:

  $$\approx |T| f_G \quad \text{exact for } f \in \mathbb{P}_1(T)$$
  $$\approx \frac{|T|}{3} \left(f_1 + f_2 + f_3\right) \quad \text{exact for } f \in \mathbb{P}_1(T)$$
  $$\approx \frac{|T|}{3} \left(f_{12} + f_{13} + f_{23}\right) \quad \text{exact for } f \in \mathbb{P}_2(T)$$
  $$\approx \frac{|T|}{60} \left(3(f_1 + f_2 + f_3) + 8(f_{12} + f_{13} + f_{23}) + 27f_G\right) \quad \text{exact for } f \in \mathbb{P}_3(T)$$