# Galerkin method and analysis of PDEs

S. Repin

Steklov Institute of Mathematics at St. Petersburg

# The Plan

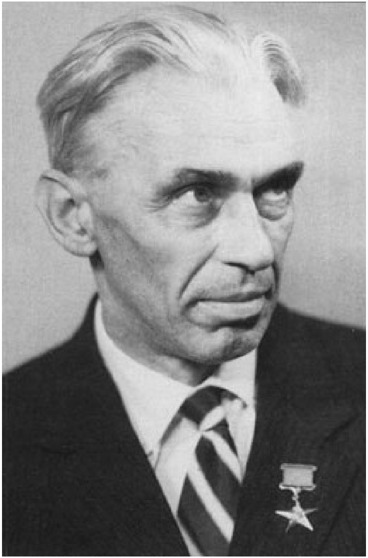
- Short historical comments
- Galerkin method and proving mathematical correctness of BVPs.
- Development of Galerkin type methods for quantitative analysis of PDEs.

## W. Ritz, 1878–1909, born in Switzerland, worked in Zurich and Goettingen

# I.G. Bubnov, 1872–1919, St. Petersburg Marine Academy

B. G. Galerkin, 1871–1945, born in Plotsk, Graduated from St. Petrsburg Technological Institute, USSR Academy of Sciences (1928,1935)

# G.I. Petrov, 1812–1987, MGU, USSR Academy (1953,1958)



# Correctness of boundary value problems for PDEs

A hard mathematical questions of the 19th century: Does the problem

$$\Delta u + f = 0 \qquad \text{in } \Omega \tag{1}$$
$$u = u_0 \qquad \text{on } \Gamma. \tag{2}$$

always uniquely solvable?

Originally, this problem was understood in the classical sense, i.e., find $u \in C^2(\Omega) \cap C(\bar{\Omega})$ such that the boundary condition (2) is satisfied and

$$\frac{\partial^2 u}{\partial^2 x_1} + \frac{\partial^2 u}{\partial^2 x_1} + f = 0$$

at all the points of $\Omega$, $\frac{\partial^2 u}{\partial^2 x_k}$, $k = 1, 2$ mean the classical derivatives.

How should we understand solution of (1)–(2) and when it indeed exists?

This question has been answered only about one hundred years of studies, which completely reconstructed the theory of partial differential equations due to

D. Hilbert, H. Poincaré, S. Sobolev,

R. Courant, O. A. Ladyzhenskaya,

and many others mathematicians who contributed to the conception of a *generalized* or *weak* solution

Galerkin method have stimulated a turn towards this new outlook

In Galerkin method we are aimed to find

$$u_N = u_0 + \sum_{i=1}^{N} \alpha_i w_i$$

(where $\{w_i\}$ is a "proper"collection of functions) such that

$$\int_\Omega (\Delta u_N + f) w_i \, dx = 0 \qquad \forall w_i, \quad i = 1, 2, ...N.$$

In other words: $u_N$ is a function which residual is *orthogonal* to a finite–dimensional space $V_N$ formed by linearly independent trial functions $w_i$.

An obvious idea is to replace $V_N$ by the whole space $V$!

Very soon this idea was genrated and it was understood that "admissible"solutions of a linear PDE $Lu = f$ should be considered among those satisfying

$$\int_{\Omega} u\, L^* w\, dx = \int_{\Omega} fw\, dx$$

for any sufficiently smooth function $w$ with compact support.

N. Wiener, A.A. Fridman, N.E. Kochin 1920-30

For (1)–(2), this idea reads
*find a function that makes the residual orthogonal to all test functions $w$ from a proper functional space $V$*, i.e.,

$$\int_\Omega (\Delta u + f) w \, dx = 0 \qquad \forall w \in V.$$

Integration by parts leads to the *generalized statement* of (1)–(2):
Find $u$ satisfying the boundary conditions with $\nabla u \in L^2(\Omega)$ (in formal notation $u \in \overset{\circ}{H}{}^1(\Omega) + u_0$) such that

$$\int_\Omega \nabla u \cdot \nabla w \, dx = \int_\Omega f w \, dx \qquad \forall w \in V = \overset{\circ}{H}{}^1(\Omega). \tag{3}$$

It is not difficult to show that if (1)–(2) has a classical solution, then it automatically satisfies (3). However, the opposite statement is wrong.

# Roots in physics "Virtual Work Principle"

Conception of "weak" solutions correlates with physical principles that have formed the basement of mechanics

D'Alembert's form of the principle of virtual work, which is used to derive the equations of motion for a mechanical system of rigid bodies. Various forms of this principle have been credited to Johann Bernoulli (1667–1748) and Daniel Bernoulli (1700–1782).
Virtual displacements principle of Lagrange

$$\Pi(u, w) - F(w) = 0 \qquad \text{for all admissible (?) displacements } w$$

$$\int\limits_{\Omega} \sigma : \varepsilon(w)\, dx - \int\limits_{\Omega} f \cdot w\, dx - \int\limits_{\Gamma} g \cdot w\, dx = 0 \quad \forall w \in ?$$

$$\sigma = L\varepsilon(u) \qquad (Hooke's law)$$

## A sequence of Galerkin solutions converges to a (weak) solution

Henceforth, we use notation $u_h$ for the Galerkin solution and $V_{0h} \in V_0$ for the space.

$$\int_\Omega \nabla u_h \cdot \nabla w_h \, dx = \int_\Omega f w_h \, dx \qquad \forall w_h \in V_{0h}$$

Assume that a collection of finite dimensional subspaces $\{V_{0h_k}\}$, $k = 1, 2, 3, ...$ is limit dense in $V_0$,
i.e., for any $w \in V_0$, there exists a sequence $w_h \in V_{0h_k}$ such that

$$w_{h_k} \to w \quad \text{in} \quad V.$$

Galerkin solutions are uniformly bounded: $\|\nabla u_{h_k}\| \leq C\|f\|$, so that there exists a weakly converging subsequence $u_{h_k}$ tending to $u \in V$. Passing to the limit we see that $u$ is a (generalized) solution:

$$\int_\Omega \nabla u \cdot \nabla w \, dx = \int_\Omega f w \, dx$$

This type argumentation was used for justification of weak solutions in various problems of mathematical physics.

E. Hopf 1950-51: proved existence of weak solutions to nonstationary NS equations.

O.A. Ladyzhenskya often used it for proving existence of weak solutions with finite energy norm for various partial differential equations

## Galerkin solutions form the basis of quantitative analysis of PDEs

Galerkin method has generated many important mathematical studies:

- How to find $u_h$? $\Rightarrow$ New methods in linear algebra: DDM, Multigrid, Sparse Matrics methods, Tensor Trains...
- Does the whole sequence $\{u_h\}$ tend to $u$? How to qualify the convergence in terms of $h$?
  $\Rightarrow$ A priori error estimation theory for PDEs
- How to reform Galerkin subspaces ($V_{h_k} \rightarrow V_{h_{k+1}}$)?
  $\Rightarrow$ Mesh adaptive numerical methods, error indicators
- How to evaluate the error (difference between $u$ and $u_h$)?
  $\Rightarrow$ A posteriori estimates for PDEs

# FEM: a special case of Galerkin method

$$\operatorname{div} A \nabla u + f = 0 \qquad \text{in } \Omega\,, \tag{4}$$

$$u = 0 \qquad \text{on } \Gamma\,, \tag{5}$$

$A \in \mathbb{M}^{d\times d}$ is a matrix with bounded entries, $f \in L_2(\Omega)$,

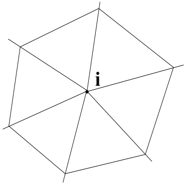
$$c_1^2|\xi|^2 \le A(x)\xi\cdot\xi \le c_2^2|\xi|^2 \qquad \forall \xi \in \mathbb{R}^d\,,\ \text{for a.a. } x \in \Omega\,. \tag{6}$$

**Galerkin solution for $V_h$**

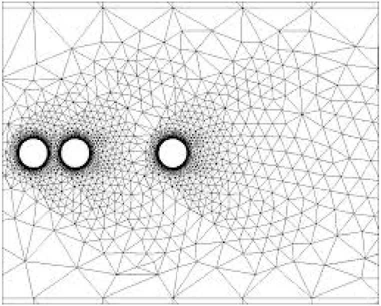
$$\int_\Omega A\nabla u_h \cdot \nabla w_h \, dx = \int_\Omega f w_h \, dx \qquad \forall w_h \in V_h. \tag{7}$$

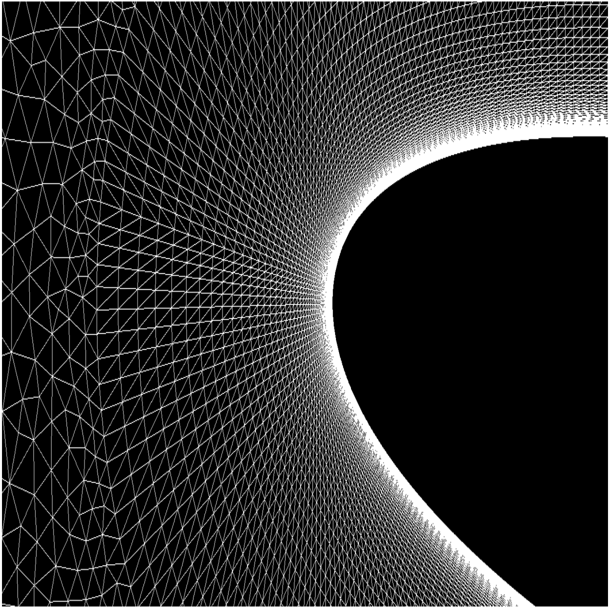
Key issue: how to select $V_h$?
R. Courant, 1943 suggested a way that later generated the Finite Element Method (FEM).

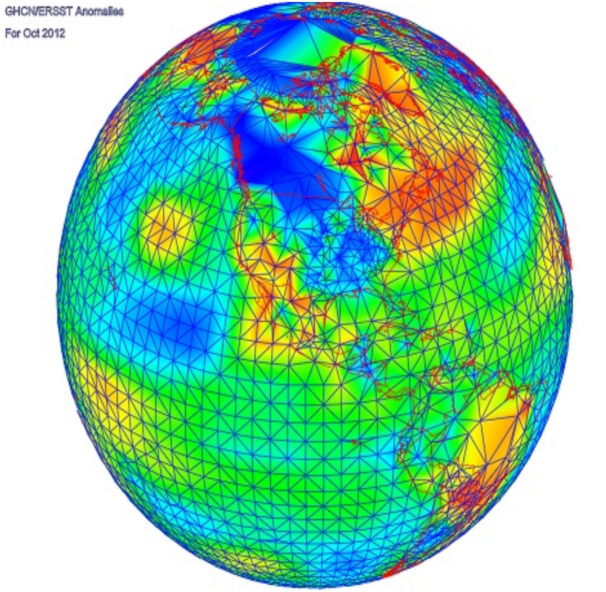
$$\Omega \Rightarrow \cup T_i$$




Nowadays label FEM is used as a common name for Galerkin approximations using trial/basic functions with local supports.






GHCN/ERSST Anomalies
For Oct 2012

Analysis of FEM is based upon two fundamental relations:
**Galerkin orthogonality**
and
**Projection inequality**

# Galerkin orthogonality

$$\int_\Omega A\nabla u \cdot \nabla w_h \, dx = \int_\Omega f w_h \, dx \qquad \forall V_{0h} \in V_{0h}.$$

Hence, the error $e_h = u - u_h$ satisfy

$$\int_\Omega A\nabla e_h \cdot \nabla w_h \, dx = 0 \qquad \forall V_{0h} \in V_{0h}.$$

$e_h$ is orthogonal to the subspace $V_{0h}$ with respect to the energy scalar product $(u, v)_A := \int_\Omega A\nabla u \cdot \nabla v \, dx$.

# Projection inequality

A priori analysis of the accuracy of Galerkin approximations was in the center of attention in 60-70s years
G. Strang and G. Fix, J. Aubin, S. Mikhlin, J. Nitsche, L. Oganesyan, L. Rukhovetz, V. Rivkind, Ph. Ciarlet, J. Cea, C. Johnson, L. Wahlbin and Schatz, C. Johnson, P.-A. Raviart and J.-M. Thomas...

Basic point: Galerkin solution meets the estimate

$$\|\nabla(u - u_h)\| \leq C \inf_{w_h \in V_h} \|\nabla(u - w_h)\|$$

$$\|\nabla(u - u_h)\| \leq Ch^k,$$

Key question: properties of projection operator $\Pi : V \to V_h$.

Kolmogorov's N width of a compact $K \subset V$ (1936):

$$d_N(K) = \inf_{V^{(N)}} \sup_{v \in K} \inf_{v_N \in V^{(N)}} \|\nabla(v - v_N)\|$$

Estimates of $d_N(K)$: S. Stechkin (1954), V. Tikhomirov (1960), K. Babenko

Example: Let $K$ be a set of functions defined on $\Omega = (0,1) \times (0,1)$ and bounded in $W^{2,2}(\Omega)$. Let $V = W^{1,2}(\Omega)$. Then $d_N(K) \sim CN^{-1/2}$.

- Valid only for regular (exact) solutions (counterexamples due to Babuska and Osborn).
- Require certain regularity of meshes.
- Estimates are derived for the whole class of problems and unable to provide an efficient information for a concrete $u_h$.
- Valid only for Galerkin approximations.

Equation (4) originates from two physical relations:

$$\begin{aligned} \operatorname{div} p + f &= 0 && \text{Equilibrium equation,} \\ p &= A \nabla u && \text{Material law} \\ u &= 0 && \text{on } \Gamma \end{aligned}$$

Comment: Now it is commonly accepted that it is better to keep two variables $(p, u)$ instead of one.

We represent the system in a new form using the same idea: "verification via test functions"

Find $(u, p) \in V \times Q$ such that

$$L(u, q) \;\le\; L(u, p) \;\le\; L(w, p) \quad \forall q \in Q\,,\; \forall w \in V \tag{8}$$

where $V = \overset{\circ}{H}{}^1(\Omega)$, $Q := L_2(\Omega, \mathbb{R}^d)$,

$$L(w, q) := \int_\Omega (\nabla w \cdot q - \frac{1}{2} A^{-1} q \cdot q - fw) dx$$

Saddle point satisfies the integral relations

$$A^{-1} p = \nabla u \;\Leftarrow\; \int_\Omega \left(A^{-1} p - \nabla u\right) \cdot q \, dx = 0 \qquad \forall q \in Q\,, \tag{9}$$

$$\mathrm{div} q + f = 0 \;\Leftarrow\; \int_\Omega (p \cdot \nabla w - fw)\, dx = 0 \qquad \forall w \in V\,, \tag{10}$$

Galerkin approximations: $Q_h \subset Q = L^2$, $V_h \subset V$.

$$\int\limits_{\Omega} \left(A^{-1} p_h - \nabla u_h\right) \cdot q_h \, dx = 0 \qquad \forall q_h \in Q_h \,, \tag{11}$$

$$\int\limits_{\Omega} (p_h \cdot \nabla w_h - f w_h) \, dx = 0 \qquad \forall w_h \in V_h \,, \tag{12}$$

This scheme is equivalent to (7).
It is not locally concervative!
It is focused on the relation $p = A \nabla u_h$ instead of $\mathrm{div} p + f = 0$.

How to get a locally concervative scheme?

We can integrate by parts and transform the Lagrangian to

$$\widehat{L}(v,q) = -\int_\Omega (v(\mathrm{div} q) + A^{-1}q\cdot q + fv)dx$$

Now we can define $q \in \widehat{Q}; = H(\Omega, \mathrm{div})$ and $v \in \widehat{V} := L^2(\Omega)$.
New saddle point problem

$$\widehat{L}(\widehat{u},\widehat{q}) \le \widehat{L}(\widehat{u},\widehat{p}) \le \widehat{L}(\widehat{v},\widehat{p}) \quad \forall \widehat{q}\in\widehat{Q}, \quad \widehat{v}\in\widehat{V}$$

Saddle point satisfies

$$\int_\Omega \left(A^{-1}\widehat{p}\cdot\widehat{q} + \widehat{u}\mathrm{div}\widehat{q}\right)\, dx = 0 \qquad \forall \widehat{q}\in\widehat{Q}\,, \tag{13}$$

$$\int_\Omega (\mathrm{div}\widehat{p} + f)\widehat{v}\, dx = 0 \qquad \forall \widehat{v}\in\widehat{V}\,. \tag{14}$$

We can prove that $(\widehat{u}, \widehat{p})$ indeed exists and moreover coincides with $(u, p)$. However, for this purpose we need the so-called Inf-Sup Condition, which in this case has the form

$$\inf_{\widehat{v} \in L^2} \sup_{\widehat{q} \in H(\Omega, \mathrm{div})} \frac{\int\limits_{\Omega} \widehat{v} \,\mathrm{div} \widehat{q} \, dx}{\|\widehat{q}\|_{\mathrm{div}} \|\widehat{v}\|} \geq c > 0. \tag{15}$$

This fact follows from Stability Lemma

**Lemma**

*For any $\widehat{v} \in L_2$ there exists $\widehat{q} \in H(\Omega, \mathrm{div})$ such that*

$$\mathrm{div} \widehat{q} + \widehat{v} = 0, \tag{16}$$

$$\|\widehat{q}\| \leq C \|\widehat{v}\| \tag{17}$$

$\widehat{V}$ contains piecewise constant functions $\widehat{v}_h$

$$\int_\Omega \left(A^{-1}\widehat{p}_h \cdot \widehat{q}_h + \widehat{u}_h \mathrm{div}\widehat{q}_h\right)\, dx = 0 \quad \forall \widehat{q}_h \in \widehat{Q}_h\,,$$

$$\{\!|\, \mathrm{div}\widehat{q} + f \,|\!\}_T = 0 \;\Leftarrow\; \int_\Omega (\mathrm{div}\widehat{p}_h + f)\widehat{v}_h\, dx = 0 \qquad \forall \widehat{v}_h \in \widehat{V}\,.$$

However Discrete Stability Lemma (or InfSup Condition) depends on the pair $\widehat{V}_h, \widehat{Q}_h$.

$$\inf_{\widehat{v}_h \in L^2} \sup_{\widehat{q}_h \in H(\Omega,\mathrm{div})} \frac{\int_\Omega \widehat{v}_h \mathrm{div}\widehat{q}_h\, dx}{\|\widehat{q}_h\|_{\mathrm{div}} \|\widehat{v}_h\|} \geq c_h \geq \widehat{c} > 0. \tag{18}$$

Various InfSup conditions often arise if we wish to guarantee stability of mixed approximations and properly select pairs of spaces $(V_h, Q_h)$

Example: Stokes problem

$$\inf_{\substack{\phi \in L^2 \\ \{| \phi |\}=0}} \sup_{\widehat{\mathbf{v}} \in H^1(\Omega, \mathbb{R}^d)} \frac{\int\limits_\Omega \phi \operatorname{div} \mathbf{v} \, dx}{\|\phi\| \|\nabla \mathbf{v}\|} \geq c > 0. \tag{19}$$

Equivalent form: Aziz–Babuska–Ladyzhenskaya–Solonnikov Lemma

Discrete spaces for the pressure $\phi$ and velocity $\mathbf{v}$ must satisfy the respective InfSup condition with $c_h$ uniformly bounded from below. Otherwise Galerkin approximations will suffer from "locking".

# Why adaptive methods have emerged?

Example: Linear 3D vector problem, $\Omega = (0,1)^3$,

$$N \approx \frac{3}{h^3}$$

$h = 0.01$, Resolving matrix: $3.0\,10^6 \times 3.0\,10^6$ elements.

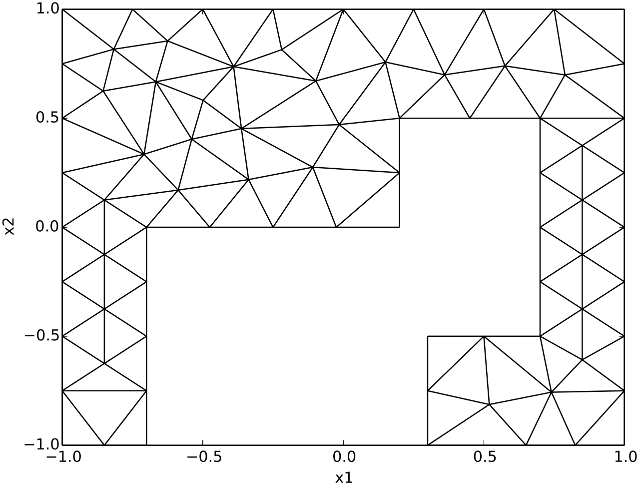
$h = 0.005$, Resolving matrix: $2.4\,10^7 \times 2.4\,10^7$ elements.

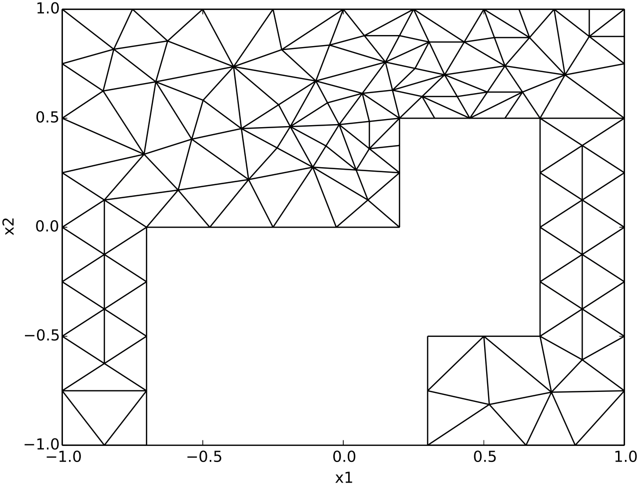
$h = 0.001$, Resolving matrix: $3.0\,10^9 \times 3.0\,10^9$ elements.

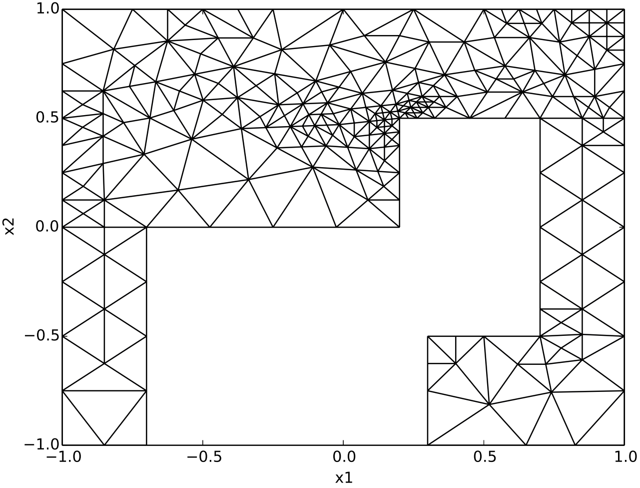
**Theoretical scheme "uniform meshes and $h \to h/2$ refinement"is faced with HUGE technical difficulties!**

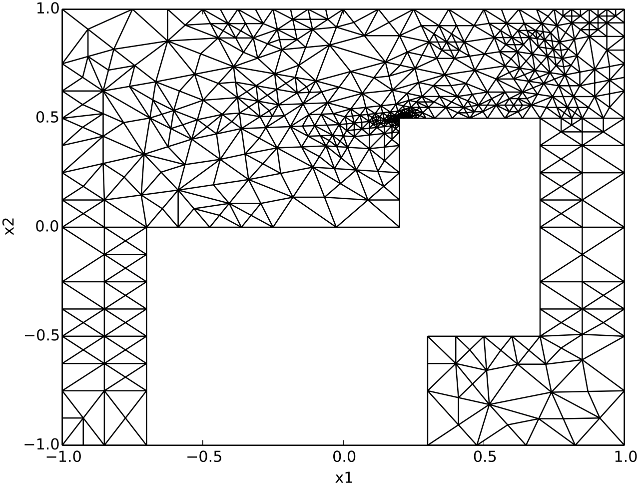
Since late 70s the mainstream direction in numerical analysis of PDEs is
**ADAPTIVITY**

Main idea:
the subspace $V_{k+1}$ must be constructed by analysing approximate solutions computed on $V_k$

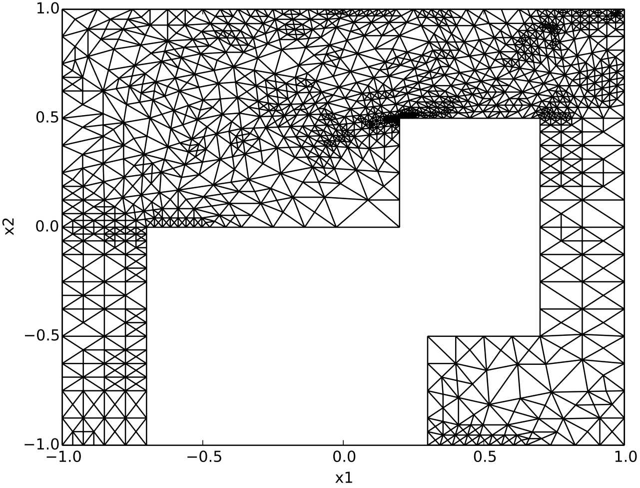

1.0
0.5
0.0
−0.5
−1.0
x2
−1.0
−0.5
0.0
0.5
1.0
x1


1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0
x1
x2

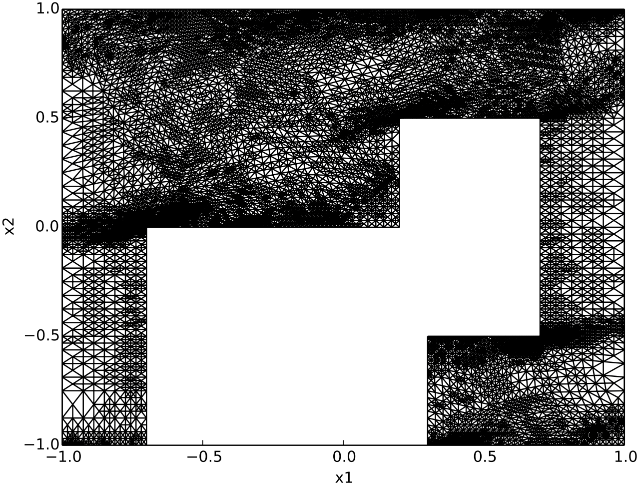

1.0
0.5
0.0
−0.5
−1.0
x2
−1.0
−0.5
0.0
0.5
1.0
x1



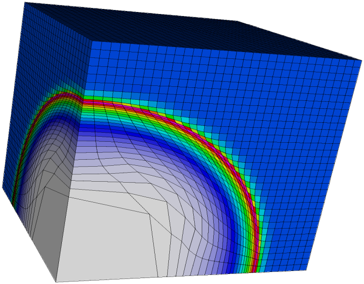
1.0
0.5
0.0
−0.5
−1.0
x2
−1.0
−0.5
0.0
0.5
1.0
x1

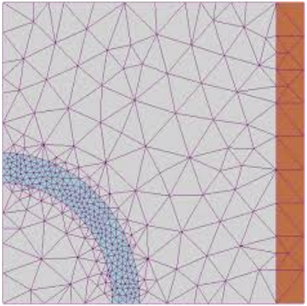


1.0
0.5
0.0
−0.5
−1.0
x2
−1.0
−0.5
0.0
0.5
1.0
x1


1.0
0.5
0.0
−0.5
−1.0
x2
−1.0
−0.5
0.0
0.5
1.0
x1

Adaptation is necessary for approximation of boundary layers and internal free boundaries typical for many nonlinear problems





## Adaptive methods have generated new advanced versions of Galerkin method

Adaptive methods require error indication and mesh reconstruction. Mesh reconstruction must satisfy certain geometrical (topological) restrictions and, therefore, could be very difficult.

Hybrid Methods
Discontinuous Galerkin Methods,
Finite Volume Method
Weak Galerkin Methods

# Discontinuous Galerkin method

Arnold 1982,Babuska 1973,BabuskaZlamal 1973
Consider the simplest elliptic problem

$$p = \nabla u, \tag{20}$$
$$\operatorname{div} p + f = 0. \tag{21}$$

and define the so-called "broken"spaces

$$\widetilde{V} := \{w = \{w^i\},\ w^i \in H^1(\Omega_i), \quad i = 1, 2, ..., N\}$$
$$\widetilde{Q} := \{\, q = \{\, q^i\},\ q^i \in H(\Omega_i, \operatorname{div}),\ q \cdot n \in L^2(\partial\Omega_i)\},$$

For a subdomain $\omega$ we have (due to (20) and (21))

$$p = \nabla u \;\Rightarrow\; \int\limits_{\omega} p \cdot\, q dx = \int\limits_{\omega} -u \operatorname{div} q dx + \int\limits_{\partial\omega} u(\, q \cdot n) ds, \qquad (22)$$

$$\operatorname{div} p + f = 0 \;\Rightarrow\; \int\limits_{\omega} p \cdot \nabla w dx = \int\limits_{\omega} (-\operatorname{div} p) w dx + \int\limits_{\partial\omega} (\, p \cdot n) w ds$$

$$= \int\limits_{\omega} f w dx + \int\limits_{\partial\omega} (\, p \cdot n) w ds. \qquad (23)$$

Let $\Omega$ be decomposed into a collection of subdomains $\Omega_i$.

$$
\begin{aligned}
\int_{\Omega_i} p \cdot q^i dx &= \int_{\Omega_i} -u \operatorname{div} q^i dx + \int_{\partial\Omega_i} u(q^i \cdot n) ds \quad \forall q^i, \\
\int_{\Omega_i} p \cdot \nabla w^i dx &= \int_{\Omega_i} fw dx + \int_{\partial\Omega_i} (p \cdot n) w^i ds \quad \forall w^i.
\end{aligned}
$$

In the DG method $\widetilde{V}$ and $\widetilde{Q}$
are replaced by finite–dimensional spaces $\widetilde{V}_h$ and $\widetilde{Q}_h$
containing locally polynomial but discontinuous functions $q_h^i$ and $w_h^i$.

Then above system is replaced by a close (but different) system

$$\int_{\Omega_i} p_h \cdot q_h^i dx = -\int_{\Omega_i} u_h \mathrm{div}\, q_h^i dx + \int_{\partial\Omega_i} \widetilde{u}( q_h^i \cdot n) ds \qquad \forall q_h^i \quad (24)$$

$$\int_{\Omega_i} p_h \cdot \nabla w_h^i dx = \int_{\Omega_i} fw dx + \int_{\partial\Omega_i} (\widetilde{p} \cdot n) w_h^i ds \qquad \forall w_h^i, \quad (25)$$

where

$$\widetilde{u} = G_1(u_h, p_h) \qquad \text{and} \qquad \widetilde{p} = G_2(u_h, p_h)$$

are *numerical reconstructions*, which are approximations of $u$ and $\nabla u \cdot n$ on the boundary $\partial\Omega_i$. Example: the Bassi–Rebay method.

$$\widetilde{u} = \{\!| u_h |\!\} \quad \text{on } \Gamma_0 \, (\textit{interior}),\ \widetilde{u} = 0 \qquad \text{on } \Gamma,$$
$$\widetilde{p} = \{\!| p_h |\!\} \quad \text{on } \Gamma_0 + \Gamma.$$

Set $w_h^i = 1$ on $\Omega_i$.
Then, from (25) it follows that

$$\int\limits_{\Omega_i} f dx + \int\limits_{\partial\Omega_i} \widetilde{p} \cdot n ds = 0. \tag{26}$$

This fact means that the scheme is *conservative* on any subdomain.

# Weak Galerkin Method

General idea:
Let $\Omega$ be a polygonal domain with boundary $\Gamma$. We define a "weak"function as a pair $v = \{v_0, v_b\}$ such that $v_0 \in L_2(\Omega)$ and $v_b \in H^{1/2}(\partial\Omega)$
A weak gradient of $v$ can be defined as a linear functional in the dual space of $H(\mathrm{div}, \Omega)$ whose action on each $q \in H(\mathrm{div}, \Omega)$ is given by

$$(\nabla_d v, q) := -\int_\Omega v_0 \mathrm{div} q \, dx + \int_\Gamma v_b \, q \cdot n \, ds$$

For any pair $(v_0, v_b)$, the right-hand side defines a bounded linear functional on the normed linear space $H(\mathrm{div}, \Omega)$. Thus, the weak gradient is well defined. With the weak gradient operator $\nabla_d$, the trial and test functions can be allowed to take separate values/definitions on the interior of each element T and its boundary. Consequently, we are left with a greater option in applying the Galerkin to partial differential equations.

- **Generation I:**
  Special spaces of test functions $\Rightarrow$ geometrical flexibility, suitable resolving matrixes.
- **Generation II:**
  Mixed formulations $\Rightarrow$ locally conservative solutions
- **Generation III:**
  Adaptivity $\Rightarrow$ Weak Galerkin approximations

# Main questions/problems in study now

**Problem I.**
We have applied a stable, locally conservative, and adaptive Galerkin method and finally stopped with some $V_h$.
**What we have computed?**

**Problem II.**
**Which components of a (weak) solution are indeed quantitatively/numerically observable ?**

Thank you for attention!