LECTURES ON A POSTERIORI ERROR CONTROL

History. Error indicators for finite element methods.
Functional a posteriori error estimates.
A posteriori estimates for the Stokes problem.
A posteriori estimates for the linear elasticity problem.
A posteriori estimates for mixed methods.
Evaluation of errors arising due to data indeterminacy.
A posteriori estimates for iteration methods.
Functional a posteriori estimates for variational inequalities.

S. Repin,
V.A. Steklov Institute of Mathematics in St.-Petersburg

Linz, December 2005
1 Lecture 1. INTRODUCTION. ERROR ANALYSIS IN THE MATHEMATICAL MODELING
   - Errors in mathematical modeling
   - Mathematical background
   - A priori estimates

2 Lecture 2. A CONCISE OVERVIEW OF A POSTERIORI ERROR ESTIMATION METHODS FOR APPROXIMATIONS OF DIFFERENTIAL EQUATIONS.
   - First approaches
   - Residual method
   - A posteriori methods based on post-processing
   - A posteriori methods using adjoint problems

3 Lecture 3. FUNCTIONAL A POSTERIORI ESTIMATES. FIRST EXAMPLES.
   - Functional A Posteriori Estimates. The concept
   - Derivation by the variational method
   - Derivation by the method of integral identities
   - Properties of functional a posteriori estimates
   - How to use them in practice?
4 Lecture 4. AN INTRODUCTION TO DUALITY THEORY
- A class of convex variational problems
- Dual (polar) functionals
- Examples
- Subdifferentials
- Compound functionals
- Uniformly convex functionals

5 Lecture 5. FUNCTIONAL A POSTERIORI ESTIMATES. GENERAL APPROACH.
- Two-sided a posteriori estimates for linear elliptic problems
- Properties: computability, efficiency, reliability
- Relationships with other methods
- Residual based estimates

6 Lecture 6. FUNCTIONAL A POSTERIORI ESTIMATES. LINEAR ELLIPTIC PROBLEMS.
- Diffusion equation
- Linear elasticity
- Equations of the 4th order
Lecture 7. FUNCTIONAL A POSTERIORI ESTIMATES. STOKES PROBLEM.

- Stokes problem
- Inf-Sup condition
- Existence of a saddle point
- Estimates of the distance to the set of solenoidal fields
- Comments on the value of the LBB-constant
- Functional a posteriori estimates for the Stokes problem
- Estimates for problems with condition $\text{div} u = \phi$.
- Problems for almost incompressible fluids
- Generalizations to problems where a solution is seeking in a subspace

Lecture 8. ESTIMATION OF INDETERMINACY ERRORS.

- Errors arising due to data indeterminacy. Examples
- General concept
- Upper bound of the error
- Lower bound of the error

Lecture 9. A POSTERIORI ESTIMATES FOR MIXED METHODS.

- Mixed approximations. A glance from the minimax theory
- A priori error analysis for the dual mixed method
- A posteriori error estimates for the primal mixed method

S. Repin

RICAM, Special Radon Semester, Linz, 2005.
Lecture 10. A POSTERIORI ERROR ESTIMATES FOR ITERATION METHODS

- Banach fixed point theorem
- A posteriori estimates
- Advanced two-sided a posteriori estimates
- Applications to problems with bounded linear operators
- Iteration methods in linear algebra
- Positivity methods and a posteriori error bounds
- Applications to integral equations
- Applications to ordinary differential equations

Lecture 11. A POSTERIORI ESTIMATES FOR VARIATIONAL INEQUALITIES

- Variational inequalities
- Obstacle problem. Introduction
- Deviation estimate for variational inequalities
- Problems with two obstacles
- Perturbed problem
- A posteriori estimates for problems with two obstacles
- Numerical examples
The elasto-plastic torsion problem
Preface

This lecture course was prepared for the Special Radon Semester organized in October–December 2005 by J. Radon Institute of Computational and Applied Mathematics (RICAM) in Linz, Austria. The main purpose of the course is to present (at least for certain classes of partial differential equations) a mathematically justified and practically efficient answer to the question:

*How to verify the accuracy of approximate solutions computed by various numerical methods?*

During the last decade, this question has been intensively investigated by the functional methods of the theory of partial differential equations. As a result a new (functional) approach to the a posteriori error control of differential equations has been formed. In the present course of lectures, I tried to present the main ideas and results of this approach in the most transparent form and discuss it using several classical problems (diffusion problem, linear elasticity, Stokes problem) as basic examples.
The material is based on earlier lectures on a posteriori estimates and adaptive methods (University of Houston (2002), USA; Summer Schools of the University of Jyväskylä, Finland (2003, 2005); St.-Petersburg Polytechnical University). Also, I used some publications appeared in 2000-2004. However, in many parts the course is quite new and reflects the latest achievements in the area. A list of the literature is given at the end of the text, but certain key publications are also cited in the respective places related to the topic discussed.

I am grateful to RICAM and especially to Prof. U. Langer for the kind support. Also, I thank Prof. D. Braess, Prof. R. Lazarov, Dr. J. Valdman, and Dr. S. Tomar for the interest and discussions.

Sergey Repin

Linz, December 2005
Lecture 1.
INTRODUCTION. ERROR ANALYSIS IN THE MATHEMATICAL MODELING
Lecture plan

- Errors arising in mathematical modeling;
- Basic mathematical knowledge
  - Notation
  - Functional spaces and inequalities;
  - Generalized solutions.
- A priori error estimates for elliptic type PDE’s
We begin with two assertions that present a motivation of this lecture course.

I. In the vast majority of cases, exact solutions of differential equations are unknown. We have no other way to use differential equations in the mathematical modeling, but to compute their approximate solutions and analyze them.

II. Approximate solutions contain errors of various nature.

From I and II, it follows that

III. Error analysis of the approximate solutions to differential equations is one of the key questions in the Mathematical Modeling.
Errors in mathematical modeling

\( \varepsilon_1 \) – error of a mathematical model used

\( \varepsilon_2 \) – approximation error arising when a differential model is replaced by a discrete one;

\( \varepsilon_3 \) – numerical errors arising when solving a discrete problem.
MODELING ERROR

Let $U$ be a physical value that characterizes some process and $u$ be a respective value obtained from the mathematical model. Then the quantity

$$
\varepsilon_1 = |U - u|
$$

is an error of the mathematical model.

Mathematical model always presents an ”abridged” version of a physical object. Therefore, $\varepsilon_1 > 0$. 
TYPICAL SOURCES OF MODELING ERRORS

(a) "Second order" phenomena are neglected in a mathematical model.

(b) Problem data are defined with an uncertainty.

(c) Dimension reduction is used to simplify a model.
Let $u_h$ be a solution on a mesh of the size $h$. Then, $u_h$ encompasses the approximation error

$$\varepsilon_2 = |u - u_h|.$$ 

Classical error control theory is mainly focused on approximation errors.
Finite–dimensional problems are also solved approximately, so that instead of $u_h$ we obtain $u_h^\varepsilon$. The quantity

$$\varepsilon_3 = |u_h - u_h^\varepsilon|$$

shows an error of the numerical algorithm performed with a concrete computer. This error includes

- roundoff errors,
- errors arising in iteration processes and in numerical integration,
- errors caused by possible defects in computer codes.
Roundoff errors

Numbers in a computer are presented in a floating point format:

\[ x = \pm \left( \frac{i_1}{q} + \frac{i_2}{q^2} + \ldots + \frac{i_k}{q^k} \right) q^\ell, \quad i_s < q. \]

These numbers form the set \( R_{q\ell k} \subset \mathbb{R} \).

\( q \) is the base of the representation,
\( \ell \in [\ell_1, \ell_2] \) is the power.

\( R_{q\ell k} \) is not closed with respect to the operations \( +, -, \ast \)!
The set $R_{q\ell k} \times R_{q\ell k}$
Example

\[ k = 3, \quad a = \left( \frac{1}{2} + 0 + 0 \right) \times 2^5, \quad b = \left( \frac{1}{2} + 0 + 0 \right) \times 2^1 \]
\[ b = \left( 0 + \frac{1}{2} + 0 \right) \times 2^2 = \left( 0 + 0 + \frac{1}{2} \right) \times 2^3 = (0 + 0 + 0) \times 2^4 \]

\[ a + b = a!!! \]

Definition. The smallest floating point number which being added to 1 gives a quantity different from 1 is called the machine accuracy.
Numerical integration

\[ \int_a^b f(x) \, dx \approx \sum_{i=1}^{n} c_i f(x_i) h = \sum_{i=1}^{n/2} c_i f(x_i) h + c_{n/2+1} f(x_{n/2+1}) h + \ldots \]
Errors in computer simulation

\[ U \xrightarrow{\text{Physical object/process}} \varepsilon_1 \xrightarrow{\text{Error of a model}} u \xrightarrow{\text{Differential model } Au = f} \varepsilon_2 \xrightarrow{\text{Approximation error}} u_h \xrightarrow{\text{Discrete model } A^h u_h = f_h} \varepsilon_3 \xrightarrow{\text{Computational error}} u_h^\varepsilon \xrightarrow{\text{Numerical solution } A^h u_h^\varepsilon = f_h + \varepsilon.} \]
Two principal relations

I. Computations on the basis of a reliable (certified) model. Here $\varepsilon_1$ is assumed to be small and $u_h^\varepsilon$ gives a desired information on $U$.

$$\|U - u_h^\varepsilon\| \leq \varepsilon_1 + \varepsilon_2 + \varepsilon_3.$$  \hspace{1cm} (1.1)

II. Verification of a mathematical model. Here physical data $U$ and numerical data $u_h^\varepsilon$ are compared to judge on the quality of a mathematical model

$$\|\varepsilon_1\| \leq \|U - u_h^\varepsilon\| + \varepsilon_2 + \varepsilon_3.$$  \hspace{1cm} (1.2)
Thus, two major problems of mathematical modeling, namely,

- reliable computer simulation,
- verification of mathematical models by comparing physical and mathematical experiments,

require efficient methods able to provide COMPUTABLE AND REALISTIC estimates of $\varepsilon_2 + \varepsilon_3$. 
What is \( u \) and what is \( \| \cdot \| \)?

If we start a more precise investigation, then it is necessary to answer the question

**What is a solution to a boundary-value problem?**

Example.

\[
\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + f = 0, \quad u = u_0 \text{ on } \partial \Omega.
\]

Does such a function \( u \) exists and unique? It is not a trivial question, so that about one hundred years passed before mathematicians have found an appropriate concept for PDE’s.
Without proper understanding of a mathematical model no real modeling can be performed. Indeed,

If we are not sure that a solution \( u \) exists then what we try to approximate numerically?

If we do not know to which class of functions \( u \) belongs to, then we cannot properly define the measure for the accuracy of computed approximations.

Thus, we need to recall a

CONCISE MATHEMATICAL BACKGROUND
Vectors and tensors

\( \mathbb{R}^n \) contains real \( n \)-vectors. \( \mathbb{M}^{n\times m} \) contains \( n \times m \) matrices and \( \mathbb{M}_s^{n\times n} \) contains \( n \times n \) symmetric matrices (tensors) with real entries.

\[
a \cdot b = \sum_{i=1}^{n} a_i b_i \in \mathbb{R}, \quad a, b \in \mathbb{R}^n \quad \text{(scalar product of vectors)},
\]

\[
a \otimes b = \{a_i b_j\} \in \mathbb{M}^{n\times n} \quad \text{(tensor product of vectors)},
\]

\[
\sigma : \varepsilon = \sum_{i,j=1}^{n} \sigma_{ij} \varepsilon_{ij} \in \mathbb{R}, \quad \sigma, \varepsilon \in \mathbb{M}^{n\times n} \quad \text{(scalar product of tensors)}.
\]

\[
|a| := \sqrt{a \cdot a}, \quad |\sigma| := \sqrt{\sigma : \sigma},
\]

Unit matrix is denoted by \( \mathbb{I} \). If \( \tau \in \mathbb{M}^{n\times n} \), then \( \tau^D = \tau - \frac{1}{n} \mathbb{I} \) is the deviator of \( \tau \).
Spaces of functions

Let $\Omega$ be an open bounded domain in $\mathbb{R}^n$ with Lipschitz continuous boundary.

- $C^k(\Omega)$ – $k$ times continuously differentiable functions.
- $C^k_0(\Omega)$ – $k$ times continuously differentiable functions vanishing at the boundary $\partial \Omega$.
- $C_0^\infty(\Omega)$ – $k$ smooth functions with compact supports in $\Omega$.
- $L^p(\Omega)$ – summable functions with finite norm

\[ \| g \|_{p,\Omega} = \| g \|_p = \left( \int_{\Omega} |g|^p \right)^{1/p} . \]

For $L^2(\Omega)$ the norm is denoted by $\| \cdot \|$. 

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
If $g$ is a vector (tensor)-valued function, then the respective spaces are denoted by $C^k(\Omega, \mathbb{R}^n)$ ($C^k(\Omega, \mathbb{M}^{n \times n})$), $L^p(\Omega, \mathbb{R}^n)$ ($L^p(\Omega, \mathbb{M}^{n \times n})$) with similar norms.

We say that $g$ is **locally integrable** in $\Omega$ and write $f \in L^{1,\text{loc}}(\Omega)$, if $g \in L^1(\omega)$ for any $\omega \subset \subset \Omega$. Similarly, one can define the space $L^{p,\text{loc}}(\Omega)$ that consists of functions locally integrable with degree $p \geq 1$. 
Generalized derivatives

Let \( f, g \in L^{1,\text{loc}}(\Omega) \) and

\[
\int_{\Omega} g \varphi \, dx = - \int_{\Omega} f \frac{\partial \varphi}{\partial x_i} \, dx, \quad \forall \varphi \in \mathcal{C}^1(\Omega).
\]

Then \( g \) is called a \textit{generalized derivative} (in the sense of Sobolev) of \( f \) with respect to \( x_i \) and we write

\[
g = \frac{\partial f}{\partial x_i}.
\]
Higher order generalized derivatives

If \( f, g \in L^{1,\text{loc}}(\Omega) \) and

\[
\int_{\Omega} g \varphi \, dx = \int_{\Omega} f \frac{\partial^2 \varphi}{\partial x_i \partial x_j} \, dx, \quad \forall \varphi \in C^2(\Omega),
\]

then \( g \) is a generalized derivative of \( f \) with respect to \( x_i \) and \( x_j \). For generalized derivatives we keep the classical notation and write

\[
g = \frac{\partial^2 f}{\partial x_i \partial x_j} = f_{,ij}.
\]

If \( f \) is differentiable in the classical sense, then its generalized derivatives coincide with the classical ones!
To extend this definition further, we use the multi-index notation and write $D^{\alpha}f$ in place of $\partial^k f / \partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \ldots \partial x_n^{\alpha_n}$.

**Definition**

Let $f, g \in L^{1,\text{loc}}(\Omega)$ and

$$\int_{\Omega} g \varphi \, dx = (-1)^{|\alpha|} \int_{\Omega} f D^{\alpha} \varphi \, dx, \quad \forall \varphi \in C^k(\Omega).$$

Then, $g$ is called a **generalized derivative** of $f$ of degree

$$|\alpha| := \alpha_1 + \alpha_2 + \ldots + \alpha_n$$

and we write

$$g = D^{\alpha}f.$$
Sobolev spaces

The spaces of functions that have integrable generalized derivatives up to a certain order are called Sobolev spaces.

Definition

Let \( f \in W^{1,p}(\Omega) \) if \( f \in L^p \) and all the generalized derivatives of \( f \) of the first order are integrable with power \( p \), i.e.,

\[
f_{,i} = \frac{\partial f}{\partial x_i} \in L^p(\Omega).
\]

The norm in \( W^{1,p} \) is defined as follows:

\[
\| f \|_{1,p,\Omega} := \left( \int_{\Omega} \left( |f|^p + \sum_{i=1}^{n} |f_{,i}|^p \right) dx \right)^{1/p}.
\]
The other Sobolev spaces are defined quite similarly: \( f \in W^{k,p}(\Omega) \) if all generalized derivatives up to the order \( k \) are integrable with power \( p \) and the quantity

\[
\|f\|_{k,p,\Omega} := \left( \int_{\Omega} \left( \sum_{|\alpha| \leq k} |D^\alpha f|^p \right) dx \right)^{1/p}
\]

is finite. For the Sobolev spaces \( W^{k,2}(\Omega) \) we also use a simplified notation \( H^k(\Omega) \).

Sobolev spaces of vector- and tensor-valued functions are introduced by obvious extensions of the above definitions. We denote them by \( W^{k,p}(\Omega, \mathbb{R}^n) \) and \( W^{k,p}(\Omega, \mathbb{M}^{n \times n}) \), respectively.
Embedding Theorems

Relationships between the Sobolev spaces and $L^p(\Omega)$ and $C^k(\Omega)$ are given by Embedding Theorems.

If $p, q \geq 1$, $\ell > 0$ and $\ell + \frac{n}{q} \geq \frac{n}{p}$, then $W^{\ell,p}(\Omega)$ is continuously embedded in $L^q(\Omega)$. Moreover, if $\ell + \frac{n}{q} > \frac{n}{p}$, then the embedding operator is compact.

If $\ell - k > \frac{n}{p}$, then $W^{\ell,p}(\Omega)$ is compactly embedded in $C^k(\Omega)$. 

S. Repin  
RICAM, Special Radon Semester, Linz, 2005. 
LECTURES ON A POSTERIORI ERROR CONTROL
Traces

The functions in Sobolev spaces have counterparts on $\partial \Omega$ called **traces**. Thus, there exist some bounded operators mapping the functions defined in $\Omega$ to functions defined on the boundary, e.g.,

$$\gamma : H^1(\Omega) \rightarrow L^2(\partial \Omega)$$

is called the **trace operator** if it satisfies the following conditions:

$$\gamma v = v |_{\partial \Omega}, \quad \forall v \in C^1(\Omega),$$

$$\|\gamma v\|_{2, \partial \Omega} \leq c \|v\|_{1,2,\Omega},$$

where $c$ is a positive constant independent of $v$. From these relations, we observe that such a trace is a natural generalization of the trace defined for a continuous function.
It was established that $\gamma v$ forms a subset of $L^2(\partial \Omega)$, which is the space $H^{1/2}(\partial \Omega)$. The functions from other Sobolev spaces also are known to have traces in Sobolev spaces with fractional indices.

Henceforth, we understand the boundary values of functions in the sense of traces, so that

$$u = \psi \quad \text{on} \quad \partial \Omega$$

means that the trace $\gamma u$ of a function $u$ defined in $\Omega$ coincides with a given function $\psi$ defined on $\partial \Omega$.

All the spaces of functions that have zero traces on the boundary are marked by the symbol $\circ$ (e.g., $W^{1,p}(\Omega)$ and $H^1(\Omega)$).
In the lectures we will use the following inequalities 1. **Friederichs-Steklov inequality.**

\[ \|w\| \leq C_\Omega \|\nabla w\|, \quad \forall w \in H^1(\Omega), \quad (1.3) \]

2. **Poincaré inequality.**

\[ \|w\| \leq \tilde{C}_\Omega \|\nabla w\|, \quad \forall w \in \tilde{H}^1(\Omega), \quad (1.4) \]

where \( \tilde{H}^1(\Omega) \) is a subset of \( H^1 \) of functions with zero mean.

3. **Korn’s inequality.**

\[ \int_{\Omega} \left( |v|^2 + |\varepsilon(v)|^2 \right) dx \geq \mu_\Omega \|v\|^2_{1,2,\Omega}, \quad \forall v \in H^1(\Omega, \mathbb{R}^n), \quad (1.5) \]
**Definition**

Linear functionals defined on the functions of the space $\dot{C}^\infty(\Omega)$ are called **distributions**. They form the space $\mathcal{D}'(\Omega)$.

Value of a **distribution** $g$ on a function $\varphi$ is $\langle g, \varphi \rangle$. Distributions possess an important property:

they have derivatives of any order

Let $g \in \mathcal{D}'(\Omega)$, then the quantity $-\langle g, \frac{\partial \varphi}{\partial x_i} \rangle$ is another linear functional on $\mathcal{D}(\Omega)$. It is viewed as a generalized partial derivative of $g$ taken over the $i$-th variable.
Derivatives of $L^q$–functions

Any function $g$ from the space $L^q(\Omega)$ ($q \geq 1$) defines a certain distribution as

$$\langle g, \varphi \rangle = \int_\Omega g \varphi \, dx$$

and, therefore, has generalized derivatives of any order. The sets of distributions, which are derivatives of $q$-integrable functions, are called Sobolev spaces with negative indices.
Definition

The space $W^{-\ell,q}(\Omega)$ is the space of distributions $g \in \mathcal{D}'(\Omega)$ such that

$$g = \sum_{|\alpha| \leq \ell} D^\alpha g_\alpha,$$

where $g_\alpha \in L^q(\Omega)$. 

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
Spaces $W^{-1,p}(\Omega)$

$W^{-1,p}(\Omega)$ contains distributions that can be viewed as generalized derivatives of $L^q$-functions. The functional

$$\left\langle \frac{\partial f}{\partial x_i}, \varphi \right\rangle := -\int_{\Omega} f \frac{\partial \varphi}{\partial x_i} \, dx \quad f \in L^q(\Omega)$$

is linear and continuous not only for $\varphi \in \mathcal{C}^\infty(\Omega)$ but, also, for $\varphi \in \overset{\circ}{W}^{1,p}(\Omega)$, where $1/p + 1/q = 1$ (density property). Hence, first generalized derivatives of $f$ lie in the space dual to $\overset{\circ}{W}^{1,p}(\Omega)$ denoted by $W^{-1,p}(\Omega)$.

For $\overset{\circ}{W}^{1,2}(\Omega) = \overset{\circ}{H}^1(\Omega)$, the respective dual space is denoted by $H^{-1}(\Omega)$. 
Norms in "negative spaces"

For $g \in H^{-1}(\Omega)$ we may introduce two equivalent "negative norms".

$$\|g\|_{-1, \Omega} := \sup_{\varphi \in H^1(\Omega)} \frac{|\langle g, \varphi \rangle|}{\|\varphi\|_{1,2, \Omega}} < +\infty$$

$$|g| := \sup_{\varphi \in H^1(\Omega)} \frac{|\langle g, \varphi \rangle|}{\|\nabla \varphi\|_{\Omega}} < +\infty$$

From the definitions, it follows that

$$\langle g, \varphi \rangle \leq \|g\|_{-1, \Omega} \|\varphi\|_{1,2, \Omega}$$

$$\langle g, \varphi \rangle \leq |g| \|\nabla \varphi\|_{\Omega}$$
Generalized solutions

The concept of generalized solutions to PDE’s came from Petrov-Bubnov-Galerkin method.

$$\int_{\Omega} (\Delta u + f)w \, dx = 0 \quad \forall w$$

Integration by parts leads to the so-called generalized formulation of the problem: find $u \in H^1(\Omega) + u_0$ such that

$$\int_{\Omega} \nabla u \cdot \nabla w \, dx = \int_{\Omega} f w \, dx \quad \forall w \in H^1(\Omega)$$

This idea admits wide extensions.
References


Definition

A symmetric form $B : V \times V \rightarrow \mathbb{R}$, where $V$ is a Hilbert space, called $V$ — elliptic if $\exists c_1 > 0, c_2 > 0$ such that

$$B(u, u) \geq c_1 \|u\|^2, \quad \forall u \in V$$

$$|B(u, v)| \leq c_2 \|u\| \|v\|, \quad \forall u, v \in V$$

General formulation for linear PDE’s is: for a certain linear continuous functional $f$ (from the space $V^*$ topologically dual to $V$) find $u$ such that

$$B(u, w) = \langle f, w \rangle \quad w \in V.$$
Existence of a solution

Usually, existence is proved by the Lax-Milgram Lemma. For a bilinear form $B$ there exists a linear bounded operator $A \in \mathcal{L}(V, V)$ such that

$$B(u, v) = (Au, v), \quad \forall u, v \in V$$

It has an inverse $A^{-1} \in \mathcal{L}(V, V)$, such that

$$\|A\| \leq c_2, \quad \|A^{-1}\| \leq \frac{1}{c_1}$$

We will follow another *modus operandi*. 
Variational approach

Lemma

If $J : K \rightarrow \mathbb{R}$ is convex, continuous and coercive, i.e.,

$$J(w) \rightarrow +\infty \quad \text{as} \quad \|w\|_V \rightarrow +\infty$$

and $K$ is a convex closed subset of a reflexive space $V$, then the problem

$$\inf_{w \in K} J(w)$$

has a minimizer $u$. If $J$ is strictly convex, then the minimizer is unique.

Coercivity

Take $J(w) = \frac{1}{2} B(w, w) - \langle f, w \rangle$ and let $K$ be a certain subspace. Then

$$\frac{1}{2} B(w, w) \geq c_1 \|w\|_V^2, \quad \|f\|_* \|w\|_V \leq \|f\|_* \|w\|_V.$$ 

We see, that

$$J(w) \geq c_1 \|w\|_V^2 - \|f\|_* \|w\|_V \rightarrow +\infty \quad \text{as} \quad \|w\|_V \rightarrow +\infty$$

Since $J$ is strictly convex and continuous we conclude that a minimizer exists and unique.
Useful algebraic relation

First we present the algebraic identity

\[
\frac{1}{2} B(u - v, u - v) = \frac{1}{2} B(v, v) - \langle f, v \rangle + \langle f, u \rangle - \frac{1}{2} B(u, u) - B(u, v - u) + \langle f, v - u \rangle = J(v) - J(u) - B(u, v - u) + \langle f, v - u \rangle
\]

From this identity we derive two important results:

- (a) Minimizer \( u \) satisfies \( B(u, w) = \langle f, w \rangle \);
- (b) Error is subject to the difference of functionals.
Let us show (a), i.e., that from (1.6) it follows the identity

\[ B(u, v - u) = < f, v - u > \quad \forall v \in K, \]

which is \( B(u, w) = < f, w > \) if set \( w = v - u \). Indeed, assume the opposite, i.e. \( \exists \bar{v} \in K \) such that

\[ B(u, \bar{v} - u) - < f, \bar{v} - u > = \delta > 0 \quad (\bar{v} \neq u!) \]

Set \( \tilde{v} := u + \alpha(\bar{v} - u), \alpha \in \mathbb{R} \). Then \( \tilde{v} - u = \alpha(\bar{v} - u) \) and

\[
\frac{1}{2} B(u - \tilde{v}, u - \tilde{v}) + B(u, \tilde{v} - u) + < f, \tilde{v} - u > =
\]

\[ = \frac{\alpha^2}{2} B(\bar{v} - u, \bar{v} - u) + \alpha \delta = J(\tilde{v}) - J(u) \geq 0 \]

However, for arbitrary \( \alpha \) such an inequality cannot be true. Denote \( a = B(\bar{v} - u, \bar{v} - u) \). Then in the left–hand side we have a function \( 1/2 \alpha^2 a^2 + \alpha \delta \), which always attains negative values for certain \( \alpha \). For example, set \( \alpha = -\delta/a^2 \). Then, the left–hand side is equal to \( -\frac{1}{2} \delta^2 / a^2 < 0 \) and we arrive at a contradiction.
Now, we show (b). From

\[
\frac{1}{2} B(u - v, u - v) = J(v) - J(u) - B(u, v - u) + \langle f, v - u \rangle
\]

we obtain the error estimate:

\[
\frac{1}{2} B(u - v, u - v) = J(v) - J(u). \quad (1.7)
\]


which immediately gives the **projection estimate**
Projection estimate

Let $u_h$ be a minimizer of $J$ on $K_h \subset K$. Then

$$\frac{1}{2} B(u - u_h, u - u_h) = J(u_h) - J(u) \leq J(v_h) - J(u) =$$

$$= \frac{1}{2} B(u - v_h, u - v_h) \quad \forall v_h \in K_h.$$

and we observe that

$$B(u - u_h, u - u_h) = \inf_{v_h \in K_h} B(u - v_h, u - v_h) \quad (1.8)$$

Projection type estimates serve a basis for deriving a priori convergence estimates.
Interpolation in Sobolev spaces

Two key points: PROJECTION ESTIMATE and INTERPOLATION IN SOBOLEV SPACES.
Interpolation theory investigates the difference between a function in a Sobolev space and its piecewise polynomial interpolant. Basic estimate on a simplex $T_h$ is

$$|v - \Pi_h v|_{m,t,T_h} \leq C(m,n,t) \left( \frac{h}{\rho} \right)^m h^{2-m} \|v\|_{2,t,T_h},$$

and on the whole domain

$$|v - \Pi_h v|_{m,t,\Omega_h} \leq Ch^{2-m} \|v\|_{2,t,\Omega_h}.$$

Here $h$ is the element size and $\rho$ is the inscribed ball diameter.
Asymptotic convergence estimates

Typical case is $m = 1$ and $t = 2$. Since

$$B(u - u_h, u - u_h) \leq B(u - \Pi_h u, u - \Pi_h u) \leq c_2 \|u - \Pi_h u\|^2$$

for

$$B(w, w) = \int_{\Omega} \nabla w \cdot \nabla w \, dx$$

we find that

$$\|\nabla (u - u_h)\| \leq Ch\|u\|_{2,2,\Omega}.$$

provided that

- Exact solution is $H^2$ – regular;
- $u_h$ is the Galerkin approximation;
- Elements do not "degenerate" in the refinement process.
A priori convergence estimates cannot guarantee that the error monotonically decreases as $h \to 0$. Besides, in practice we are interested in the error of a concrete approximation on a particular mesh. Asymptotic estimates can hardly serve these purposes because, in general the constant $C$ in such an estimate is either unknown or highly overestimated. Therefore, a priori convergence estimates have mainly a theoretical value: they show that an approximation method is correct ”in principle”.

For these reasons, starting from late 70th a quite different approach to error control is rapidly developing. Nowadays it has already formed a new direction: **A Posteriori Error Control for PDE’s**
Lecture 2.
A CONCISE OVERVIEW OF A POSTERIORI ERROR ESTIMATION METHODS FOR APPROXIMATIONS OF DIFFERENTIAL EQUATIONS.
Lecture plan

- Heuristic Runge’s rule;
- Prager and Synge estimate. Estimate of Mikhlin;
- Estimates using negative norm of the equation residual;
  - Basic idea;
  - Estimates in 1D case;
  - Estimates in 2D case;
  - Comments;
- Methods based on post-processing;
- Methods using adjoint problems;
Runge’s rule

At the end of 19th century a heuristic error control method was suggested by C. Runge who investigated numerical integration methods for ordinary differential equations.

Carle Runge
Heuristic rule of C. Runge

If the difference between two approximate solutions computed on a coarse mesh $\mathcal{T}_h$ with mesh size $h$ and refined mesh $\mathcal{T}_{h_{\text{ref}}}$ with mesh size $h_{\text{ref}}$ (e.g., $h_{\text{ref}} = h/2$) has become small, then both $u_{h_{\text{ref}}}$ and $u_h$ are probably close to the exact solution.

In other words, this rule can be formulated as follows:

If $[u_h - u_{h_{\text{ref}}}]$ is small then $u_{h_{\text{ref}}}$ is close to $u$

where $[\cdot]$ is a certain functional or mesh-dependent norm.

Also, the quantity $[u_h - u_{h_{\text{ref}}}]$ can be viewed (in terms of modern terminology) as a certain \textit{a posteriori error indicator}. 
Runge’s heuristic rule is simple and was easily accepted by numerical analysts.

However, if we do not properly define the quantity \[ \cdot \], for which \[ u_h - u_{h_{\text{ref}}} \] is small, then the such a principle may be not true.

One can present numerous examples where two subsequent elements of an approximation sequence are close to each other, but far from a certain joint limit. For example, such cases often arise in the minimization (maximization) of functionals with "saturation" type behavior or with a "sharp–well" structure. Also, the rule may lead to a wrong presentation if, e.g., the refinement has not been properly done, so that new trial functions were added only in subdomains were an approximation is almost coincide with the true solution. Then two subsequent approximations may be very close, but at the same time not close to the exact solution.
Also, in practice, we need to now precisely what the word ”close” means, i.e. we need to have a more concrete presentation on the error. For example, it would be useful to establish the following rule:

\[
\text{If } [u_h - u_{h,\text{ref}}] \leq \varepsilon \text{ then } \|u_h - u\| \leq \delta(\varepsilon),
\]

where the function \(\delta(\varepsilon)\) is known and computable.

In subsequent lectures we will see that for a wide class of boundary–value problems it is indeed possible to derive such type generalizations of the Runge’s rule.
Prager and Synge estimates


W. Prager and J. L. Synge
Prager and Synge derived an estimate on the basis of purely geometrical grounds. In modern terms, there result for the problem

\[ \Delta u + f = 0, \quad \text{in } \Omega, \]
\[ u = 0, \quad \text{on } \partial \Omega \]

reads as follows:

\[ \| \nabla (u - v) \|_2^2 + \| \nabla u - \tau \|_2^2 = \| \nabla v - \tau \|_2^2, \]

where \( \tau \) is a function satisfying the equation \( \text{div} \tau + f = 0 \). We can easily prove it by the orthogonality relation

\[ \int_{\Omega} \nabla (u - v) \cdot (\nabla u - \tau) \, dx = 0 \quad (\text{div}(\nabla u - \tau) = 0!). \]
Estimate of Mikhlin

A similar estimate was derived by variational arguments (see Lecture 1). It is as follows:

\[ \frac{1}{2} \left\| \nabla (u - v) \right\|^2 \leq J(v) - \inf J, \]

where

\[ J(v) := \frac{1}{2} \left\| \nabla v \right\|^2 - (f, v), \quad \inf J := \inf_{v \in H_1(\Omega)} J(v). \]
Dual problem

Since

\[ \inf J = \sup_{\tau \in Q_f} \left\{ -\frac{1}{2} \| \tau \|^2 \right\}, \]

where

\[ Q_f := \left\{ \tau \in L_2(\Omega, \mathbb{R}^d) \mid \int_{\Omega} \tau \cdot \nabla \mathbf{w} \, dx = \int_{\Omega} f \mathbf{w} \, dx \quad \forall \mathbf{w} \in \mathring{H}^1 \right\}, \]

we find that

\[ \frac{1}{2} \| \nabla (u - v) \|^2 \leq J(v) + \frac{1}{2} \| \tau \|^2, \quad \forall \tau \in Q_f. \]
Since

\[ J(v) + \frac{1}{2} \| \tau \|^2 = \frac{1}{2} \| \nabla v \|^2 - \int_{\Omega} fv \, dx + \frac{1}{2} \| \tau \|^2 = \]

\[ = \frac{1}{2} \| \nabla v \|^2 - \int_{\Omega} \tau \cdot \nabla v \, dx + \frac{1}{2} \| \tau \|^2 = \]

\[ = \frac{1}{2} \| \nabla v - \tau \|^2 \]

we arrive at the estimate

\[ \frac{1}{2} \| \nabla (u - v) \|^2 \leq \frac{1}{2} \| \nabla v - \tau \|^2, \quad \forall \tau \in Q_f. \] 

(2.1)
Difficulties

Estimates of Prager and Synge and of Mikhlin are valid for any \( v \in H_1(\Omega) \), so that, formally, that they can be applied to any conforming approximation of the problem. However, from the practical viewpoint these estimates have an essential drawback:

\textbf{they use a function \( \tau \) in the set \( Q_f \) defined by the differential relation,}

which may be difficult to satisfy exactly. Probably by this reason further development of a posteriori error estimates for Finite Element Methods (especially in 80’-90’) was mainly based on different grounds.
Errors and Residuals: first glance

If an analyst is not sure in an approximate solution, then the very first idea that comes to his mind is to substitute it into the equation considered, i.e. to look at the equation residual.

We begin by recalling basic relations between residuals and errors that hold for systems of linear simultaneous equations. Let $A \in M^{n \times n}$, $\det A \neq 0$, consider the system

$$Au + f = 0.$$ 

For any $v$ we have the simplest residual type estimate

$$A(v - u) = Av + f; \quad \Rightarrow \quad \|e\| \leq \|A^{-1}\|\|r\|.$$ 

where $e = v - u$ and $r = Av + f$. 

S. Repin

RICAM, Special Radon Semester, Linz, 2005.
Two–sided estimates

Define the quantities

\[ \lambda_{\text{min}} = \min_{y \in \mathbb{R}^n, y \neq 0} \frac{\|Ay\|}{\|y\|} \quad \text{and} \quad \lambda_{\text{max}} = \max_{y \in \mathbb{R}^n, y \neq 0} \frac{\|Ay\|}{\|y\|}. \]

Since \( Ae = r \), we see that

\[ \lambda_{\text{min}} \leq \frac{\|Ae\|}{\|e\|} = \frac{\|r\|}{\|e\|} \leq \lambda_{\text{max}} \Rightarrow \lambda_{\text{max}}^{-1} \|r\| \leq \|e\| \leq \lambda_{\text{min}}^{-1} \|r\|. \]

Since \( u \) is a solution, we have

\[ \lambda_{\text{min}} \leq \frac{\|Au\|}{\|u\|} = \frac{\|f\|}{\|u\|} \leq \lambda_{\text{max}} \Rightarrow \lambda_{\text{max}}^{-1} \|f\| \leq \|u\| \leq \lambda_{\text{min}}^{-1} \|f\|. \]

Thus,

\[ \frac{\lambda_{\text{min}}}{\lambda_{\text{max}}} \frac{\|r\|}{\|f\|} \leq \frac{\|e\|}{\|u\|} \leq \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}} \frac{\|r\|}{\|f\|}. \]
Since

\[ \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}} = \text{Cond } A, \]

we arrive at the basic relation where the matrix condition number serves as an important factor

\[ (\text{Cond } A)^{-1} \frac{\|r\|}{\|f\|} \leq \frac{\|e\|}{\|u\|} \leq \text{Cond } A \frac{\|r\|}{\|f\|}. \tag{2.2} \]

Thus, the relative error is controlled by the relative value of the residual. However, the bounds deteriorates when the conditional number is large.
In principle, the above consideration can extended to a wider set of linear problems, where

$$\mathcal{A} \in \mathcal{L}(X, Y)$$

is a coercive linear operator acting from a Banach space $X$ to another space $Y$ and $f$ is a given element of $Y$.

However, if $\mathcal{A}$ is related to a boundary-value problem, then one should properly define the spaces $X$ and $Y$ and find a practically meaningful analog of the estimate (2.2).
Elliptic equations

Let $\mathcal{A} : X \rightarrow Y$ be a linear elliptic operator. Consider the boundary-value problem

$$\mathcal{A}u + f = 0 \quad \text{in } \Omega, \quad u = u_0 \quad \text{on } \partial\Omega.$$  

Assume that $v \in X$ is an approximation of $u$. Then, we should measure the error in $X$ and the residual in $Y$, so that the principal form of the estimate is

$$\|v - u\|_X \leq C\|\mathcal{A}v + f\|_Y,$$  

(2.3)

where the constant $C$ is independent of $v$. The key question is as follows:

**Which spaces $X$ and $Y$ should we choose for a particular boundary-value problem?**
Consider the problem

$$\Delta u + f = 0 \quad \text{in} \Omega, \quad u = 0 \quad \text{on} \partial \Omega,$$

with $f \in L^2(\Omega)$. The generalized solution satisfies the relation

$$\int_{\Omega} \nabla u \cdot \nabla w \, dx = \int_{\Omega} fw \, dx \quad \forall w \in V_0 := H^1(\Omega),$$

which implies the energy estimate

$$\|\nabla u\|_{2,\Omega} \leq C_\Omega \|f\|_{2,\Omega}.$$

Here $C_\Omega$ is a constant in the Friederichs-Steklov inequality. Assume that an approximation $v \in V_0$ and $\Delta v \in L^2(\Omega)$. Then,

$$\int_{\Omega} \nabla (u - v) \cdot \nabla w \, dx = \int_{\Omega} (f + \Delta v)w \, dx, \quad \forall w \in V_0.$$
Setting $w = u - v$, we obtain the estimate

$$\|\nabla (u - v)\|_{2,\Omega} \leq C_\Omega \| f + \Delta v \|_{2,\Omega},$$

(2.4)

whose right-hand side of (2.4) is formed by the $L^2$-norm of the residual. However, usually a sequence of approximations $\{v_k\}$ converges to $u$ only in the energy space, i.e.,

$$\{v_k\} \rightarrow u \quad \text{in } H^1(\Omega),$$

so that $\|\Delta v_k + f\|$ may not converge to zero!

This means that the consistency (the key property of any practically meaningful estimate) is lost.
Which norm of the residual leads to a consistent estimate of the error in the energy norm?

To find it, we should consider \( \Delta \) not as \( H^2 \rightarrow L^2 \) mapping, but as \( H^1 \rightarrow H^{-1} \) mapping. For this purpose we use the integral identity

\[
\int_{\Omega} \nabla u \cdot \nabla w \, dx = \langle f, w \rangle, \quad \forall \ w \in V_0 := \overset{\circ}{H}^1(\Omega).
\]

Here, \( \nabla u \in L^2 \), so that it has derivatives in \( H^{-1} \) and we consider the above as equivalence of two distributions on all trial functions \( w \in V_0 \).

By \( \langle f, w \rangle \leq \| f \|_{2,\Omega} \| \nabla w \|_{2,\Omega} \), we obtain another "energy estimate"

\[
\| \nabla u \|_{2,\Omega} \leq \| f \|_{2,\Omega}.
\]
Consistent residual estimate

Let $v \in V_0$ be an approximation of $u$. We have

$$
\int_{\Omega} \nabla(u - v) \cdot \nabla w \, dx = \int_{\Omega} (fw - \nabla v \cdot \nabla w) \, dx = \langle \Delta v + f, w \rangle, \quad f + \Delta v \in H^{-1}(\Omega).
$$

By setting $w = v - u$, we obtain

$$
\|\nabla(u - v)\|_{2, \Omega} \leq \left\| f + \Delta v \right\|.
$$

(2.5)

where

$$
\left\| f + \Delta v \right\| = \sup_{\varphi \in H^1(\Omega)} \frac{|\langle f + \Delta v, \varphi \rangle |}{\|\nabla \varphi\|} = \sup_{\varphi \in H^1(\Omega)} \frac{\left| \int_{\Omega} \nabla(u - v) \cdot \nabla \varphi \right|}{\|\nabla \varphi\|} \leq \sup_{\varphi \in H^1(\Omega)} \frac{\|\nabla(u - v)\|}{\|\nabla \varphi\|} \leq \|\nabla(u - v)\|.
$$
Thus, for the problem considered

\[ \| \nabla (u - v) \|_{2, \Omega} = \int f + \Delta v \] \tag{2.6} 

From (2.6), it readily follows that

\[ \int f + \Delta v_k \to 0 \quad \text{as} \quad \{v_k\} \to u \text{ in } H^1. \]

We observe that the estimate (2.6) is consistent.
Diffusion equation

Similar estimates can be derived for

$$\mathcal{A}u + f = 0, \quad \text{in } \Omega, \quad u = 0 \text{ on } \partial \Omega,$$

where

$$\mathcal{A}u = \text{div } A \nabla u := \sum_{i,j=1}^{d} \frac{\partial}{\partial x_i} \left( a_{ij}(x) \frac{\partial u}{\partial x_j} \right),$$

$$a_{ij}(x) = a_{ji}(x) \in L^\infty(\Omega),$$

$$\lambda_{\min} |\eta|^2 \leq a_{ij}(x) \eta_i \eta_j \leq \lambda_{\max} |\eta|^2, \quad \forall \eta \in \mathbb{R}^n, \ x \in \Omega,$$

$$\lambda_{\max} \geq \lambda_{\min} \geq 0.$$
Let $v \in V_0$ be an approximation of $u$. Then,

\[ \int_{\Omega} A \nabla (u - v) \cdot \nabla w \, dx = \int_{\Omega} (fw - A \nabla v \cdot \nabla w) \, dx, \quad \forall w \in V_0. \]

Again, the right-hand side of this relation is a bounded linear functional on $V_0$, i.e.,

\[ f + \text{div} (A \nabla v) \in H^{-1}. \]

Hence, we have the relation

\[ \int_{\Omega} A \nabla (u - v) \cdot \nabla w \, dx = \langle f + \text{div} (A \nabla v), w \rangle, \quad \forall w \in V_0. \]

Setting $w = u - v$, we derive the estimate

\[ \| \nabla (u - v) \|_{2, \Omega} \leq \lambda_{\min}^{-1} \int f + \text{div} (A \nabla v) \, dx. \quad (2.7) \]
Next,

\[ \| f + \text{div} (A \nabla v) \| = \sup_{\varphi \in H^1(\Omega)} \frac{| \langle f + \text{div} (A \nabla v), \varphi \rangle |}{\| \nabla \varphi \|_{2,\Omega}} \]

\[ = \sup_{\varphi \in H^1(\Omega)} \frac{| \int_{\Omega} A \nabla (u - v) \cdot \nabla \varphi \, dx |}{\| \nabla \varphi \|_{2,\Omega}} \leq \lambda_{\max} \| \nabla (u - v) \|_{2,\Omega}. \] (2.8)

Combining (2.7) and (2.8) we obtain

\[ \lambda_{\max}^{-1} \| R(v) \| \leq \| \nabla (u - v) \|_{2,\Omega} \leq \lambda_{\min}^{-1} \| R(v) \|, \] (2.9)

where \( R(v) = f + \text{div} (A \nabla v) \in H^{-1}(\Omega) \). We see that upper and lower bounds of the error can be evaluated in terms of the negative norm of \( R(v) \).
Main goal

We observe that to find guaranteed bounds of the error reliable estimates of \( |R(v)| \) are required.

In essence, a posteriori error estimates derived in 70-90’ for Finite Element Methods (FEM) offer several approaches to the evaluation of \( |R(v)| \).

We consider them starting with the so–called explicit residual method where such estimates are obtained with help of two key points:

- Galerkin orthogonality property;
- \( H^1 \to V_h \) interpolation estimates by Clément.
Explicit residual method in 1D case

Take the simplest model

\[(\alpha u')' + f = 0, \quad u(0) = u(1).\]

Let \(I := (0, 1), f \in L^2(I), \alpha(x) \in C(\bar{I}) \geq \alpha_0 > 0\). Divide \(I\) into a number of subintervals \(I_i = (x_i, x_{i+1})\), where \(x_0 = 0, x_{N+1} = 1\), and

\[|x_{i+1} - x_i| = h_i.\]  
Assume that \(v \in H^1(I)\) and it is smooth on any interval \(I_i\). 

---

S. Repin  
RICAM, Special Radon Semester, Linz, 2005.  

LECTURES ON A POSTERIORI ERROR CONTROL
In this case,

\[ | R(v) | = \sup_{w \in V_0(I), \ w \neq 0} \frac{\int_0^1 (-\alpha v' w' + f w) \, dx}{\| w' \|_{2,1}} = \]

\[ = \sup_{w \in H^1(I); \ w \neq 0} \frac{\sum_{i=0}^N \int_{l_i} (-\alpha v' w' + f w) \, dx}{\| w' \|_{2,1}} = \]

\[ = \sup_{w \in V_0(I), \ w \neq 0} \frac{\sum_{i=0}^N \int_{l_i} r_i(v) w \, dx + \sum_{i=1}^N \alpha(x_i) w(x_i) j(v'(x_i))}{\| w' \|_{2,1}}, \]

where \( j(\phi(x)) := \phi(x + 0) - \phi(x - 0) \) is the "jump–function" and \( r_i(v) = (\alpha v')' + f \) is the residual on \( l_i \).

For arbitrary \( v \) we can hardly get an upper bound for this supremum.
Use Galerkin orthogonality

Assume that $v = u_h$, i.e., it is the \textit{Galerkin approximation} obtained on a finite–dimensional subspace $V_{0h}$ formed by piecewise polynomial continuous functions. Since

$$
\int_I \alpha u_h' w_h' \, dx - \int_I f w_h \, dx = 0 \quad \forall w_h \in V_{0h}.
$$

we may add the left–hand side with any $w_h$ to the numerator what gives

$$
\| R(u_h) \| = \sup_{w \in V_0(I)} \frac{\int_0^1 (-\alpha u_h'(w - \pi_h w)' + f(w - \pi_h w)) \, dx}{\| w' \|_{2,I}},
$$

where $\pi_h : V_0 \to V_{0h}$ is the interpolation operator defined by the conditions $\pi_h v \in V_{0h}$, $\pi_h v(0) = \pi_h v(1) = 0$ and

$$
\pi_h v(x_i) = v(x_i), \quad \forall x_i, \ i = 1, 2, \ldots, N.
$$
Integrating by parts

Now, we have

\[ \| R(u_h) \| = \sup_{w \in V_0(\Omega)} \left\{ \sum_{i=0}^{N} \int_{I_i} r_i(u_h)(w - \pi_h w) \, dx \right\} \]

\[ \frac{\| w' \|_{2,1}}{2,1} + \sum_{i=1}^{N} \alpha(x_i)(w(x_i) - \pi_h w(x_i))j(u'_h(x_i)) \right\} \]

Since \( w(x_i) - \pi_h w(x_i) = 0 \), the second sum vanishes. For first one we have

\[ \sum_{i=0}^{N} \int_{I_i} r_i(u_h)(w - \pi_h w) \, dx \leq \sum_{i=0}^{N} \| r_i(u_h) \|_{2,1_i} \| w - \pi_h w \|_{2,1_i}. \]
Since for $w \in H^1(I_i)$

$$\|w - \pi_h w\|_{2,I_i} \leq c_i \|w'\|_{2,I_i},$$

we obtain for the numerator of the above quotient

$$\sum_{i=0}^{N} \int_{I_i} r_{i}(u_h)(w - \pi_h w) \, dx \leq \sum_{i=0}^{N} c_i \|r_{i}(u_h)\|_{2,I_i} \|w'\|_{2,I_i} \leq \left( \sum_{i=0}^{N} c_i^2 \|r_{i}(u_h)\|_{2,I_i}^2 \right)^{1/2} \|w'\|_{2,I_i},$$

which implies the desired upper bound

$$\left| \mathbf{R}(u_h) \right| \leq \left( \sum_{i=0}^{N} c_i^2 \|r_{i}(u_h)\|_{2,I_i}^2 \right)^{1/2} \quad (2.10)$$

This bound is the sum of local residuals $r_{i}(u_h)$ with weights given by the interpolation constants $c_i$. 

S. Repin

LEcTUrES ON A POSTERIORI ERROR CONTROL

RICAM, Special Radon Semester, Linz, 2005.
Interpolation constants

For piecewise affine approximations, the interpolation constants $c_i$ are easy to find. Indeed, let $\gamma_i$ be a constant that satisfies the condition

$$\inf_{w \in H^1(I_i)} \frac{\|w\|_{2,I_i}^2}{\|w - \pi_h w\|_{2,I_i}^2} \geq \gamma_i.$$ 

Then, for all $w \in H^1(I_i)$, we have

$$\|w - \pi_h w\|_{2,I_i} \leq \gamma_i^{-1/2} \|w\|_{2,I_i}$$

and one can set $c_i = \gamma_{I_i}^{-1/2}$.

Let us estimate $\gamma_{I_i}$. 

S. Repin
Note that

$$\int_{x_i}^{x_{i+1}} |w'|^2 \, dx = \int_{x_i}^{x_{i+1}} |(w - \pi_h w)' + (\pi_h w)'|^2 \, dx,$$

where $(\pi_h w)'$ is constant on $(x_i, x_{i+1})$. Therefore,

$$\int_{x_i}^{x_{i+1}} (w - \pi_h w)'(\pi_h w)' \, dx = 0$$

and

$$\int_{x_i}^{x_{i+1}} |w'|^2 \, dx = \int_{x_i}^{x_{i+1}} |(w - \pi_h w)'|^2 \, dx + \int_{x_i}^{x_{i+1}} |(\pi_h w)'|^2 \, dx \geq \int_{x_i}^{x_{i+1}} |(w - \pi_h w)'|^2 \, dx.$$
Interpolation constants in 1D problem

Thus, we have

\[
\inf_{w \in H^1(I_i)} \frac{\int_{x_i}^{x_i+1} |w'|^2 \, dx}{\int_{x_i}^{x_i+1} |w - \pi_h w|^2 \, dx} \geq \inf_{w \in H^1(I_i)} \frac{\int_{x_i}^{x_i+1} |(w - \pi_h w)'|^2 \, dx}{\int_{x_i}^{x_i+1} |w - \pi_h w|^2 \, dx} \geq \inf_{\eta \in H^1(I_i)} \frac{\int_{x_i}^{x_i+1} |\eta'|^2 \, dx}{\int_{x_i}^{x_i+1} |\eta|^2 \, dx} = \frac{\pi^2}{h_i^2},
\]

so that \( \gamma_i = \pi^2 / h_i^2 \) and \( c_i = h_i / \pi \).

Remark. To prove the very last relation we note that

\[
\inf_{\eta \in H^1((0,h))} \frac{\int_0^h |\eta'|^2 \, dx}{\int_0^h |\eta|^2 \, dx} = \frac{\pi^2}{h^2},
\]

is attained on the eigenfunction \( \sin \frac{\pi}{h} x \), of the problem \( \phi'' + \lambda \phi = 0 \) on \( (0, h) \).
Residual method in 2D case

Let $\Omega$ be represented as a union $\mathcal{T}_h$ of simplexes $T_i$. For the sake of simplicity, assume that $\bar{\Omega} = \bigcup_{i=1}^N T_i$ and $V_{0h}$ consists of piecewise affine continuous functions. Then the Galerkin approximation $u_h$ satisfies the relation

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

where

$$V_{0h} = \{ w_h \in V_0 \mid w_h \in P^1(T_i), \ T_i \in \mathcal{F}_h \}.$$
In this case, negative norm of the residual is

\[ \| R(u_h) \| = \sup_{w \in V_0} \int_{\Omega} (fw - A \nabla u_h \cdot \nabla w) \, dx \left/ \| \nabla w \|_{2,\Omega} \right. \].

Let \( \pi : H^1 \rightarrow V_{0h} \) be a continuous interpolation operator. Then, for the Galerkin approximation

\[ \| R(u_h) \| = \sup_{w \in V_0} \int_{\Omega} (f(w - \pi_h w) - A \nabla u_h \cdot \nabla (w - \pi_h w)) \, dx \left/ \| \nabla w \|_{2,\Omega} \right. \].

For finite element approximations such a type projection operators has been constructed. One of the most known was suggested in Ph. Clément. Approximations by finite element functions using local regularization, \textit{RAIRO Anal. Numér.}, 9(1975).

and is often called the \textbf{Clement’s interpolation operator}. Its properties play an important role in the a posteriori error estimation method considered.
Clement’s Interpolation operator

Let $E_{ij}$ denote the common edge of the simplexes $T_i$ and $T_j$. If $s$ is an inner node of the triangulation $F_h$, then $\omega_s$ denotes the set of all simplexes having this node. For any $s$, we find a polynomial $p_s(x) \in P^1(\omega_s)$ such that

$$\int_{\omega_s} (v - p_s)q\,dx = 0 \quad \forall q \in P^1(\omega_s).$$

Now, the interpolation operator $\pi_h$ is defined by setting

$$\pi_h v(x_s) = p(x_s), \quad \forall x_s \in \Omega,$$
$$\pi_h v(x_s) = 0, \quad \forall x_s \in \partial \Omega.$$

It is a linear and continuous mapping of $H^1(\Omega)$ to the space of piecewise affine continuous functions.
Interpolation estimates in 2D

Moreover, it is subject to the relations

\[ \| v - \pi_h v \|_{2,T_i} \leq c_i^T \text{diam} (T_i) \| v \|_{1,2,\omega_N(T_i)}, \]  
\[ \| v - \pi_h v \|_{2,E_{ij}} \leq c_{ij}^E \| E_{ij} \|^{1/2} \| v \|_{1,2,\omega_E(T_i)}, \]  

where \( \omega_N(T_i) \) is the union of all simplexes having at least one common node with \( T_i \) and \( \omega_E(T_i) \) is the union of all simplexes having a common edge with \( T_i \).

Interpolation constants \( c_i^T \) and \( c_{ij}^E \) are LOCAL and depend on the shape of patches \( \omega_N(T_i) \) and \( \omega_E(T_i) \).
Quotient relations for the constants

Evaluation of $c^T_i$ and $c^E_{ij}$ requires finding exact lower bounds of the following variational problems:

$$\gamma^T_i := \inf_{w \in V_0} \frac{\|w\|_{1,2,\omega_N(T_i)}}{\|w - \pi_h w\|_{2,T_i}} \text{diam}(T_i)$$

and

$$\gamma^E_{ij} := \inf_{w \in V_0} \frac{\|w\|_{1,2,\omega_E(T_i)}}{\|w - \pi_h w\|_{2,E_{ij}}} |E_{ij}|^{1/2}.$$

Certainly, we can replace $V_0$ be $H^1(\omega_N(T_i))$ and $H^1(\omega_E(T_i))$, respectively, but, anyway finding the constants amounts solving functional eigenvalue type problems!
Let $\sigma_h = A \nabla u_h$. Then,

$$
\| R(u_h) \| = \sup_{w \in V_0} \frac{\int_\Omega (f(w - \pi_h w) - \sigma_h \cdot \nabla (w - \pi_h w)) \, dx}{\| \nabla w \|_{2, \Omega}}.
$$

If $\nu_{ij}$ is the unit outward normal to $E_{ij}$, then

$$
\int_{T_i} \sigma_h \cdot \nabla (w - \pi_h w) \, dx =
\sum_{E_{ij} \subset \partial T_i} \int_{E_{ij}} (\sigma_h \cdot \nu)(w - \pi_h w) \, ds - \int_{T_i} \text{div} \, \sigma_h (w - \pi_h w) \, dx,
$$

Since on the boundary $w - \pi_h w = 0$, we obtain

$$
\| R(u_h) \| = \sup_{w \in V_0} \left\{ \frac{\sum_{i=1}^N \int_{T_i} (\text{div} \, \sigma_h + f)(w - \pi_h w) \, dx}{\| \nabla w \|_{2, \Omega}} + \frac{\sum_{i=1}^N \sum_{j \geq i} \int_{E_{ij}} j(\sigma_h \cdot \nu_{ij})(w - \pi_h w) \, ds}{\| \nabla w \|_{2, \Omega}} \right\}.
$$
First term in $\text{sup}$

\[
\int_{T_i} (\text{div}\sigma_h + f)(w - \pi_h w) \, dx \leq \|\text{div}\sigma_h + f\|_{2,T_i} \|w - \pi_h w\|_{2,T_i} \\
\leq c_i^T \|\text{div}\sigma_h + f\|_{2,T_i} \text{diam}\,(T_i) \|w\|_{1,2,\omega_N(T_i)},
\]

Then, the first sum is estimated as follows:

\[
\sum_{i=1}^{N} \int_{T_i} (\text{div}\sigma_h + f)(w - \pi_h w) \, dx \leq \\
\leq d_1 \left( \sum_{i=1}^{N} (c_i^T)^2 \text{diam}\,(T_i)^2 \|\text{div}\sigma_h + f\|_{2,T_i}^2 \right)^{1/2} \|w\|_{1,2,\Omega},
\]

where the constant $d_1$ depends on the maximal number of elements in the set $\omega_N(T_i)$.
Second term in $\sup$

For the second one, we have

$$\sum_{i=1}^{N} \sum_{j>i}^{N} \int_{E_{ij}} j(\sigma_h \cdot \nu_{ij})(w - \pi_h w) \, dx \leq$$

$$\leq \sum_{i=1}^{N} \sum_{j>i}^{N} \|j(\sigma_h \cdot \nu_{ij})\|_{2,E_{ij}} c_{ij}^E |E_{ij}|^{1/2} \|w\|_{1,2,\omega_E(T_i)} \leq$$

$$\leq d_2 \left( \sum_{i=1}^{N} \sum_{j>i}^{N} (c_{ij}^E)^2 |E_{ij}| \|j(\sigma_h \cdot \nu_{ij})\|^2_{2,E_{ij}} \right)^{1/2} \|w\|_{1,2,\Omega},$$

where $d_2$ depends on the maximal number of elements in the set $\omega_E(T_i)$. 

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
Residual type error estimate

By the above estimates we obtain

\[ \| R(u_h) \| \leq C_0 \left( \sum_{i=1}^{N} \left( c_i^T \right)^2 \text{diam} (T_i)^2 \| \text{div} \sigma_h + f \|_{2,T_i}^2 \right)^{1/2} + \]

\[ + \left( \sum_{i=1}^{N} \sum_{j>i}^{N} \left( c_{ij}^E \right)^2 |E_{ij}| \| j(\sigma_h \cdot \nu_{ij}) \|_{2,E_{ij}}^2 \right)^{1/2} \]. \quad (2.13)\]

Here \( C_0 = C_0(d_1, d_2) \). We observe that the right-hand side is the sum of local quantities (usually denoted by \( \eta(T_i) \)) multiplied by constants depending on properties of the chosen splitting \( \mathcal{F}_h \).
Error indicator for quasi-uniform meshes

For quasi–uniform meshes all generic constants \( c_i^T \) have approximately the same value and can be replaced by a single constant \( c_1 \). If the constants \( c_{ij}^E \) are also estimated by a single constant \( c_2 \), then we have

\[
\| R(u_h) \| \leq C \left( \sum_{i=1}^{N} \eta^2(T_i) \right)^{1/2},
\]

(2.14)

where \( C = C(c_1, c_2, C_0) \) and

\[
\eta^2(T_i) = c_1^2 \text{diam}(T_i)^2 \| \text{div} \sigma_h + f \|_{2,T_i}^2 + \frac{c_2^2}{2} \sum_{E_{ij} \subset \partial T_i} |E_{ij}| \| j(\sigma_h \cdot \nu_{ij}) \|_{2,E_{ij}}^2.
\]

The multiplier \( 1/2 \) arises, because any interior edge is common for two elements.
General form of the residual type a posteriori error estimates is as follows:

$$\| u - u_h \| \leq M(u_k, c_1, c_2, \ldots c_N, D),$$

where $D$ is the data set, $u_h$ is the Galerkin approximation, and $c_i, i = 1, 2, \ldots N$ are the interpolation constants. The constants depend on the mesh and properties of the special type interpolation operator. The number $N$ depends on the dimension of $V_h$ and may be rather large. If the constants are not sharply defined, then this functional is not more than a certain error indicator. However, in many cases it successfully works and was used in numerous researches.
Comment 2

It is worth noting that for nonlinear problems the dependence between the error and the respective residual is much more complicated. A simple example below shows that the value of the residual may fail to control the distance to the exact solution.
It is commonly accepted that this approach brings its origin from the papers
A posteriori methods based on post–processing

Post–processing of approximate solutions is a numerical procedure intended to modify already computed solution in such a way that the post–processed function would fit some a priori known properties much better than the original one.
Let $e$ denotes the error of an approximate solution $v \in V$ and

$$\mathcal{E}(v) : V \to \mathbb{R}_+$$

denotes the value of an error estimator computed on $v$.

**Definition**

The estimator is said to be equivalent to the error for the approximations $v$ from a certain subset $\tilde{V}$ if

$$c_1 \mathcal{E}(v) \leq \|e\| \leq c_2 \mathcal{E}(v) \quad \forall v \in \tilde{V}$$
Definition

The ratio

\[ i_{\text{eff}} := 1 + \frac{\mathcal{E}(v) - \|e\|}{\|e\|} \]

is called the \textit{effectivity index} of the estimator \( \mathcal{E} \).

Ideal estimator has \( i_{\text{eff}} = 1 \). However, in real life situations it is hardly possible, so that values \( i_{\text{eff}} \) in the diapason from 1 to 2-3 are considered as quite good.
In FEM methods with mesh size $h$ one other term is often used:

**Definition**

The estimator $E$ is called **asymptotically equivalent to the error** if for a sequence of approximate solutions $\{u_h\}$ obtained on consequently refined meshes there holds the relation

$$\inf_{h \to 0} \frac{E(u_h)}{\|u - u_h\|} = 1$$

It is clear that an estimator may be asymptotically exact for one sequence of approximate solutions (e.g. computed on regular meshes) and not exact for another one.
Typically, the function $T u_h$ (where $T$ is a certain linear operator, e.g., $\nabla$) lies in a space $U$ that is wider than the space $\bar{U}$ that contains $T u$. If we have a computationally inexpensive continuous mapping $G$ such that $G(T v_h) \in U$, $\forall v_h \in V_h$. then, probably, the function $G(T u_h)$ is much closer to $T u$ than $T u_h$. 
These arguments form the basis of various **post-processing algorithms** that change a computed solution in accordance with some a priori knowledge of properties of the exact solution. If the error caused by violations of a priori regularity properties is dominant and the post-processing operator $G$ is properly constructed, then

$$\|GTu_h - Tu\| \ll \|Tu_h - Tu\|.$$  

In this case, the explicitly computable norm $\|GTu_h - Tu_h\|$ can be used to evaluate upper and lower bounds of the error. Indeed, assume that there is a positive number $\alpha < 1$ such that for the mapping $T$ the estimate

$$\|GTu_h - Tu\| \leq \alpha \|Tu_h - Tu\|.$$
Two–sided estimate

Then, for \( e = u_h - u \) we have

\[
(1 - \alpha) \| T e \| = (1 - \alpha) \| T u_h - T u \| \leq \nonumber \\
\leq \| T u_h - T u \| - \| G T u_h - T u \| \leq \nonumber \\
\leq \| G T u_h - T u_h \| \leq \nonumber \\
\leq \| G T u_h - T u \| + \| T u_h - T u \| \leq \nonumber \\
\leq (1 + \alpha) \| T u_h - T u \| = (1 + \alpha) \| T e \|. 
\]

Thus, if \( \alpha \ll 1 \), then

\[
\| T u_h - T u \| \simeq \| G T u_h - T u_h \|. 
\]

and the right-hand can be used as an error indicator.
Post-processing by averaging

Post-processing operators are often constructed by averaging $\mathbf{T}\mathbf{u}_h$ on finite element patches or on the entire domain.

**Integral averaging on patches**

If $\mathbf{T}\mathbf{u}_h \in L^2$, then post-processing operators are obtained by various averaging procedures. Let $\Omega_i$ be a **patch** of $M_i$ elements, i.e.,

$$\bar{\Omega}_i = \bigcup T_{ij}, \quad j = 1, 2, \ldots M_i.$$ 

Let $P^k(\Omega_i, \mathbb{R}^n)$ be a subspace of $\mathbf{U}$ that consists of vector-valued polynomial functions of degrees less than or equal to $k$. Define $\mathbf{g}_i \in P^k(\Omega_i, \mathbb{R}^n)$ as the minimizer of the problem:

$$\inf_{\mathbf{g} \in P^k(\Omega_i, \mathbb{R}^n)} \int_{\Omega_i} |\mathbf{g} - \mathbf{T}\mathbf{u}_h|^2 \, dx.$$ 

S. Repin, RICAM, Special Radon Semester, Linz, 2005.
The minimizer $g_i$ is used to define the values of an averaged function at some points (nodes). Further, these values are utilized by a prolongation procedure that defines an averaged function

$$G_T u_h : \Omega \rightarrow \mathbb{R}.$$ 

Consider the simplest case. Let $T$ be the operator $\nabla$ and $u_h$ be a piecewise affine continuous function. Then,

$$\nabla u_h \in P^0(T_{ij}, \mathbb{R}^n) \text{ on each } T_{ij} \subset \Omega_i.$$ 

We denote the values of $\nabla u_h$ on $T_{ij}$ by $(\nabla u_h)_{ij}$. 
Set \( k = 0 \) and find \( g_i \in P^0 \) such that

\[
\int_{\Omega_i} |g_i - \nabla u_h|^2 \, dx = \inf_{g \in P^0(\Omega_i)} \int_{\Omega_i} |g - \nabla u_h|^2 \, dx = \\
= \inf_{g \in P^0(\Omega_i)} \left\{ |g|^2 |\Omega_i| - 2g \cdot \sum_{j=1}^{M_i} (\nabla u_h)_{ij} |T_{ij}| + \sum_{j=1}^{M_i} |(\nabla u_h)_{ij}|^2 |T_{ij}| \right\}.
\]

It is easy to see that \( g_i \) is given by a weighted sum of \( (\nabla u_h)_{ij} \), namely,

\[
g_i = \sum_{j=1}^{M_i} \frac{|T_{ij}|}{|\Omega_i|} (\nabla u_h)_{ij}.
\]

Set \( \mathcal{G}(\nabla u_h)(x_i) = g_i \).
Repeat this procedure for all nodes and define the vector-valued function \( G \nabla (u_h) \) by the piecewise affine prolongation of these values. For regular meshes with equal \(|T_{ij}|\), we have

\[
g_i = \sum_{j=1}^{M_i} \frac{1}{M_i} (\nabla u_h)_{ij}.
\]

Various averaging formulas of this type are represented in the form

\[
g_i = \sum_{j=1}^{M_i} \lambda_{ij} (\nabla u_h)_{ij}, \quad \sum_{j=1}^{M_i} \lambda_{ij} = 1,
\]

where \( \lambda_{ij} \) are the weight factors. For internal nodes, they may be taken, e.g., as follows

\[
\lambda_{ij} = \frac{|\gamma_{ij}|}{2\pi}, \quad |\gamma_{ij}| \text{ is the angle.}
\]
However, if a node **belongs to the boundary**, then it is better to choose **special weights**. Their values depend on the mesh and on the type of the boundary. Concerning this point see

Discrete averaging on patches

Consider the problem

$$\inf_{g \in \mathbb{P}^k(\Omega_i)} \sum_{s=1}^{m_i} |g(x_s) - T u_h(x_s)|^2,$$

where the points $x_s$ are specially selected in $\Omega_i$. Usually, the points $x_s$ are the so-called superconvergent points.

Let $g_i \in \mathbb{P}^k(\Omega_i)$ be the minimizer of this problem.

If $k = 0$, and $T = \nabla$ then

$$g_i = \frac{1}{m_i} \sum_{s=1}^{m_i} \nabla u_h(x_s).$$
Global averaging

Global averaging makes the post-processing not on patches, but on the whole domain.

Assume that $T u_h \in L^2$ and find $\bar{g}_h \in V_h(\Omega) \subset \overline{U}$ such that

$$\| \bar{g}_h - T u_h \|_\Omega^2 = \inf_{g_h \in V_h(\Omega)} \| g_h - T u_h \|_\Omega^2.$$ 

The function $\bar{g}_h$ can be viewed as $G T u_h$. Very often $\bar{g}_h$ is a better image of $T u$ than the functions obtained by local procedures.
Remark

Moreover, mathematical justifications of the methods based on global averaging procedures can be performed under weaker assumptions what makes them applicable to a wider class of problems see, e.g.,

Justifications of the method. Superconvergence

Let \( u_h \) be a Galerkin approximation of \( u \) computed on \( V_h \). For piecewise affine approximations of the diffusion problem we have the estimate

\[
\| \nabla (u - u_h) \|_{2, \Omega} \leq c_1 h, \quad \| u - u_h \|_{2, \Omega} \leq c_2 h^2
\]

However, it was discovered see, e.g., L. A. Oganesjan and L. A. Ruchovec. Z. Vychisl. Mat. i Mat. Fiz., 9(1969); M. Zlámal. Lecture Notes. Springer, 1977; L. B. Wahlbin. Lecture Notes. Springer, 1969 that in certain cases this rate may be higher. For example it may happen that

\[
|u(x_s) - u_h(x_s)| \leq Ch^{2+\sigma} \quad \sigma > 0
\]

at a superconvergent point \( x_s \).
Certainly, existence and location of superconvergent points strongly depends on the structure of $\mathcal{T}_h$.

For the paradigm of the diffusion problem we say that an operator $G$ possesses a superconvergence property in $\omega \subset \Omega$ if

$$\| \nabla u - G \nabla u_h \|_{2,\omega} \leq c_2 h^{1+\sigma},$$

where the constant $c_2$ may depend on higher norms of $u$ and the structure of $\mathcal{T}_h$. 
For the diffusion problem estimates of such a type can be found, e.g., in

I. Hlaváček and M. Křížek. On a superconvergence finite element
scheme for elliptic systems. I. Dirichlet boundary conditions. *Aplikace

M. Křížek and P. Neittaanmäki. Superconvergence phenomenon in the
By exploiting the superconvergence properties, e.g.,

$$\|\nabla u - G\nabla u_h\|_{2,\omega} \leq c_2 h^{1+\sigma},$$

while

$$\|\nabla u - \nabla u_h\|_{2,\omega} \leq c_2 h,$$

one can usually construct a simple post-processing operator $G$ satisfying the condition

$$\|G\nabla u_h - \nabla u\| \leq \alpha \|\nabla u_h - \nabla u\|.$$

where the value of $\alpha$ decreases as $h$ tends to zero.
Since
\[ \| G \nabla u_h - \nabla u_h \| \leq \| \nabla u_h - \nabla u \| + \| G \nabla u_h - \nabla u \|, \]
\[ \| G \nabla u_h - \nabla u_h \| \geq \| \nabla u_h - \nabla u \| - \| G \nabla u_h - \nabla u \|. \]

where the first term in the right–hand side is of the order $h$ and the second one is of $h^{1+\delta}$. We see that
\[ \| G \nabla u_h - \nabla u_h \| \sim h \]

Therefore, we observe that in the decomposition
\[ \| \nabla (u_h - u) \| \leq \| \nabla u_h - G \nabla u_h \| + \| G \nabla u_h - \nabla u \| \]
asymptotically dominates the second directly computable term.
Thus, we obtain a simple error indicator:

$$\| \nabla (u_h - u) \| \approx \| \nabla u_h - G \nabla u_h \|. $$

Note that

$$i_{eff} = \frac{\| \nabla (u_h - u) \|}{\| \nabla u_h - G \nabla u_h \|} \approx 1 + ch^\delta$$

so that this error indicator is asymptotically exact provided that $u_h$ is a Galerkin approximation, $u$ is sufficiently regular and $h$ is small enough. Such type error indicators (often called ZZ indicators by the names of Zienkiewicz and Zhu) are widely used as cheap error indicators in engineering computations.
Some references


Post-processing by equilibration

For a solution of the diffusion problem we know that

$$\text{div} \sigma + f = 0,$$

where $\sigma = A \nabla u$. This suggests an idea to construct an operator $G$ such that

$$\text{div}(G(A \nabla u_h)) + f = 0.$$  

If $G$ possesses additional properties (linearity, boundedness), then we may hope that the function $G A \nabla u_h$ is closer to $\sigma$ than $A \nabla u_h$ and use the quantity $\| A \nabla u_h - G A \nabla u_h \|$ as an error indicator.
This idea can be applied to an important class of problems

\[ \Lambda^* Tu + f = 0, \quad Tu = \mathcal{A} \Lambda u, \]  

(2.15)

where \( \mathcal{A} \) is a positive definite operator, \( \Lambda \) is a linear continuous operator, and \( \Lambda^* \) is the adjoint operator.

In continuum mechanics, equations of the type (2.15) are referred to as the **equilibrium equations**. Therefore, it is natural to call an operator \( G \) an **equilibration** operator.

If the equilibration has been performed exactly then it is not difficult to get an upper error bound. However, in general, this task is either cannot be fulfilled or lead to complicated and expensive procedures. Known methods are usually end with approximately equilibrated fluxes.
Goal–oriented error estimates

Global error estimates give a general idea on the quality of an approximate solution and stopping criteria. However, often it is useful to estimate the errors in terms of specially selected linear functionals $\ell_s, s = 1, 2, \ldots, M$, e.g.,

$$< \ell, v - u > = \int_{\Omega} \varphi_0 (v - u) \, dx,$$

where $\varphi$ is a locally supported function. Since

$$| < \ell, u - u_h > | \leq \| \ell \| \| u - u_h \|_V,$$

we can obtain such an estimate throughout the global a posteriori estimate. However, in many cases, such a method will strongly overestimate the quantity.
Adjoint problem

A posteriori estimates of the errors evaluated in terms of linear functionals are derived by attracting the **adjoint** boundary-value problem whose right-hand side is formed by the functional $\ell$. Let us represent this idea in the simplest form. Consider a system

$$Au = f,$$

where $A$ is a positive definite matrix and $f$ is a given vector. Let $v$ be an approximate solution. Define $u_\ell$ by the relation

$$A^*u_\ell = \ell,$$

where $A^*$ is the matrix adjoint to $A$. Then,

$$\ell \cdot (u - v) = A^*u_\ell \cdot u - \ell \cdot v = f \cdot u_\ell - \ell \cdot v = (f - Av) \cdot u_\ell$$
Certainly, the above consideration holds in a more general (operator) sense, so that for a pair of operators $A$ and $A^*$ we have

\[ <\ell, u - v> = <f - Av, u_\ell> . \]  

(2.16)

and find the error with respect to a linear functional by the product of the residual and the exact solution of the adjoint problem:

\[ A^*u_\ell = \ell. \]

Practical application of this principle depends on the ability to find either $u_\ell$ or its sharp approximation.
Consider again the diffusion problem. Now, it is convenient to denote the solution of the original problem by \( u_f \), i.e.

\[
\int_{\Omega} A \nabla u_f \cdot \nabla w \, dx = \int_{\Omega} f w \, dx, \quad \forall w \in V_0(\Omega).
\]

Since in our case \( A = A^* \), the adjoint problem is to find \( u_\ell \in V_0(\Omega) \) such that

\[
\int_{\Omega} A \nabla u_\ell \cdot \nabla w \, dx = \int_{\Omega} \ell w \, dx, \quad \forall w \in V_0(\Omega).
\]
Let $\Omega$ be divided into a number of elements $T_i$, $i = 1, 2, \ldots, N$. Given approximations on the elements, we define a finite-dimensional subspace $V_{0h} \in V_0(\Omega)$ and the Galerkin approximations $u_{fh}$ and $u_{\ell h}$:

$$\int_{\Omega} A \nabla u_{fh} \cdot \nabla w_h \, dx = \int_{\Omega} f w_h \, dx, \quad \forall w_h \in V_{0h},$$

$$\int_{\Omega} A \nabla u_{\ell h} \cdot \nabla w_h \, dx = \int_{\Omega} \ell w_h \, dx, \quad \forall w_h \in V_{0h}.$$

Since

$$\int_{\Omega} \ell (u_f - u_{fh}) \, dx = \int_{\Omega} A \nabla u_{\ell} \cdot \nabla (u_f - u_{fh}) \, dx$$

and

$$\int_{\Omega} A \nabla u_{\ell h} \cdot \nabla (u_f - u_{fh}) \, dx = 0,$$
We arrive at the relation

$$\int_{\Omega} \ell(u_f - u_{fh}) \, dx = \int_{\Omega} A \nabla (u_\ell - u_{\ell h}) \cdot \nabla (u_f - u_{fh}) \, dx \quad (2.17)$$

whose right-hand side is expressed in the form

$$\sum_{i=1}^{N} \int_{T_i} A \nabla (u_f - u_{fh}) \cdot \nabla (u_\ell - u_{\ell h}) \, dx =$$

$$\sum_{i=1}^{N} \left\{ - \int_{T_i} \text{div} \left( A \nabla (u_f - u_{fh}) \right) (u_\ell - u_{\ell h}) \, dx +$$

$$+ \frac{1}{2} \int_{\partial T_i} j (\nu_i \cdot A \nabla (u_f - u_{fh})) (u_\ell - u_{\ell h}) \, ds \right\}.$$

This relation implies the estimate
\[
\int_\Omega \ell (u_f - u_{fh}) \, dx = \sum_{i=1}^{N} \left\{ \| \text{div} A \nabla (u_f - u_{fh}) \|_{2,T_i} \| u_\ell - u_{\ell h} \|_{2,T_i} + \\
+ \frac{1}{2} \| j (\mathbf{\nu}_i \cdot A \nabla (u_f - u_{fh})) \|_{2,\partial T_i} \| u_\ell - u_{\ell h} \|_{2,\partial T_i} \right\} = \\
= \sum_{i=1}^{N} \left\{ \| f + \text{div} A \nabla u_{fh} \|_{2,T_i} \| u_\ell - u_{\ell h} \|_{2,T_i} + \\
+ \frac{1}{2} \| j (\mathbf{\nu}_i \cdot A \nabla u_{fh}) \|_{2,\partial T_i} \| u_\ell - u_{\ell h} \|_{2,\partial T_i} \right\}.
\]

Here, the principal terms are the same as in the explicit residual method, but the weights are given by the norms of \( u_\ell - u_{\ell h} \).
Assume that $u_\ell \in H^2$ and $u_{\ell h}$ is constructed by piecewise affine continuous approximations. Then the norms $\| u_\ell - u_{\ell h} \|_{T_i}$ and $\| u_\ell - u_{\ell h} \|_{2,\partial T_i}$ are estimated by the quantities $h^{\alpha} | u_\ell |_{2,2,T_i}$ with $\alpha = 1$ and $1/2$ and the multipliers $\hat{c}_i$ and $\hat{c}_{ij}$, respectively.

In this case, we obtain an estimate with constants defined by the standard

$$H^2 \rightarrow V_{0h}$$

interpolation operator whose evaluation is much simpler than that of the constants arising in the

$$H^1 \rightarrow V_{0h}$$

interpolation.
A posteriori estimates in $L^2$–norm

In principle, this technology can be exploited to evaluate estimates in $L^2$–norm. Indeed,

$$\| u_f - u_{fh} \| = \sup_{\ell \in L^2} \frac{(\ell, u_f - u_{fh})}{\| \ell \|} = \sup_{\ell \in L^2} \frac{(A \nabla u_{\ell}, \nabla (u_f - u_{fh}))}{\| \ell \|} =$$

$$= \sup_{\ell \in L^2} \frac{(A \nabla (u_{\ell} - \pi_h(u_{\ell})), \nabla (u_f - u_{fh}))}{\| \ell \|} =$$

$$= \sup_{\ell \in L^2} \frac{(\nabla (u_{\ell} - \pi_h(u_{\ell})), A \nabla (u_f - u_{fh}))}{\| \ell \|} =$$

$$= \sup_{\ell \in L^2} \frac{\sum_{i=1}^{N} \left\{ \int_{T_i} \nabla (u_{\ell} - \pi_h(u_{\ell})), A \nabla (u_f - u_{fh}) \, dx \right\}}{\| \ell \|}$$

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
Integrating by parts, we obtain

\[
\sum_{i=1}^{N} \left\{ \| f + \text{div} A \nabla u_{fh} \|_{T_i} \| u_{\ell} - \pi_h(u_{\ell}) \|_{T_i} + \frac{1}{2} \| j(\nu_i \cdot A \nabla u_{fh}) \|_{\partial T_i} \| u_{\ell} - \pi_h(u_{\ell}) \|_{\partial T_i} \right\} \| \ell \|
\]

If for any \( \ell \in L^2 \) the adjoint problem has a regular solution (e.g., \( u_{\ell} \in H^2 \)), so that we could combine the standard interpolation estimate for the interpolant of \( u_{\ell} \) with the regularity estimate for the PDE (e.g., \( \| u_{\ell} \| \leq C_1 \| \ell \| \)), then we obtain

\[
\| u_{\ell} - \pi_h(u_{\ell}) \|_{T_i} \leq C_1 h^{\alpha_1} \| \ell \|, \quad \| u_{\ell} - \pi_h(u_{\ell}) \|_{\partial T_i} \leq C_1 h^{\alpha_2} \| \ell \|
\]

with certain \( \alpha_k \).

Under the above conditions \( \| \ell \| \) is reduced and we arrive at the estimate, in which the element residuals and interelement jumps are weighted with factors \( C_1 h^{\alpha_1} \) and \( C_2 h^{\alpha_2} \).
Methods using adjoint problems has been investigated in the works of R. Becker, C. Johnson, R. Rannacher and other scientists. A more detailed exposition of these works can be found in
We end this lecture with one comment concerning the terminology. In the existing literature devoted to a posteriori error analysis one can find often find terms like "duality approach to a posteriori error estimation" or "dual-based error estimates". However, the essence that is behind this terminology may be quite different because the word "duality" is used in at least 3 different meanings:

(a) **Duality in the sense of functional spaces.** We have seen that if for the equation $Lu = f$ errors are measured in the original (energy) norm then a consistent upper bound is given by the residual in the norm of the space topologically dual to a subspace of the energy space (e.g., $H^{-1}$).

(b) **Duality in the sense of using the Adjoint Problem.**

(c) **Duality in the sense of the Theory of the Calculus of Variations.**
In the next lecture we will proceed to the detailed exposition of the approach (c).
Lecture 3.
FUNCTIONAL A POSTERIORI ESTIMATES. FIRST EXAMPLES.
In the lecture, we derive Functional A Posteriori Estimate for the problem

$$\Delta u + f = 0, \Omega \quad u = 0 \partial \Omega$$

and discuss its meaning, principal features and practical implementation.
Lecture plan

- 1. Functional a posteriori estimates
- 2. How to derive them? Paradigm of a simple elliptic problem
- 3. How to use them in practice?
- 4. Examples.
Functional A Posteriori Estimates

Functional A Posteriori Estimate is a **computable majorant** of the difference between exact solution \( u \) and any conforming approximation \( v \) having the general form:

\[
\Phi(u - v) \leq M(D, v) \quad \forall v \in V!
\]  

(3.1)

where \( D \) is the data set (coefficients, domain, parameters, etc.),
\( \Phi : V \to \mathbb{R}_+ \) is a given functional.

\( M \) must be computable and continuous in the sense that

\[
M(D, v) \to 0, \quad \text{if } v \to u
\]
Types of $\Phi$

- Energy norm $\Phi(u - v) = \|u - v\|_\Omega$
- Local norm $\Phi(u - v) = \|u - v\|_\omega$
- Goal-oriented quantity $\Phi(u - v) = (\ell, u - v)$
METHODS OF THE DERIVATION.

These estimates are derived by purely functional methods using the analysis of variational problems or integral identities.

Variational method 96’-97’


Nonvariational method 2000’


Complete list of publications on the matter can be found in the References appended to the Lectures.
Functional a posteriori estimate gives complete solution of the error control problem from the viewpoint of the MATHEMATICAL THEORY of PDE’s.

Variational Method

Let $u$ be a (generalized) solution of the problem

$$\Delta u + f = 0, \quad \Omega, \quad u = 0 \partial \Omega.$$ 

As we have seen in Lecture 1, this problem is equivalent to the following variational problem:

**Problem $P$.** Find $u \in V_0 := H^1(\Omega)$ such that

$$J(u) = \inf_{v \in V_0} J(v),$$

where

$$J(v) = \frac{1}{2} \|\nabla v\|^2 - (f, v).$$

By the reasons that we discussed earlier this problem has a unique solution.
Lagrangian

Note that

\[ J(v) = \sup_{y \in Y} L(\nabla v, y), \quad L(\nabla v, y) = \int_{\Omega} \left( \nabla v \cdot y - \frac{1}{2} |y|^2 - f v \right) dx \]

where \( Y = L^2(\Omega, \mathbb{R}^n) \). Indeed, the value of the above supremum cannot exceed the one we obtain if for almost all \( x \in \Omega \) solve the pointwise problems

\[ \sup_{y(x)} (\nabla v)(x) \cdot y(x) - \frac{1}{2} |y(x)|^2 \quad x \in \Omega \]

whose upper bound is attained if set \( y(x) = (\nabla v)(x) \). Since \( \nabla v \in Y \), we observe that the respective maximizer belongs to \( Y \) and, therefore

\[ \sup_{y \in Y} L(\nabla v, y) = L(\nabla v, \nabla v) = J(v). \]
Minimax Formulations

Then, the original problem comes in the minimax form:

\[(\mathcal{P}) \quad \inf_{v \in V_0} \sup_{y \in Y} L(\nabla v, y)\]

If the order of inf and sup is changed, then we arrive at the so-called dual problem

\[(\mathcal{P}^*) \quad \sup_{y \in Y} \inf_{v \in V_0} L(\nabla v, y)\]

Note that

\[
\inf_{v \in V_0} \int_{\Omega} \left( \nabla v \cdot y - \frac{1}{2} |y|^2 - fv \right) dx = -\frac{1}{2} \|y\|^2 + \inf_{v \in V_0} \int_{\Omega} (\nabla v \cdot y - fv) dx =
\]

\[
= \begin{cases} 
-\frac{1}{2} \|y\|^2 & \text{if } y \in Q_f := \{y \in Y \mid \text{div} y + f = 0\} \\
-\infty & \text{if } y \notin Q_f
\end{cases}
\]
Thus, we observe that the dual problem has the form: find $p \in Q_f$ such that

$$-I^*(p) = \sup_{y \in Q_f} -I^*(y)$$

where

$$I^*(q) = \frac{1}{2} \|q\|^2$$

How are these two problems related?

First, we establish one relation that holds regardless of the structure of the Lagrangian.
Lemma

Let $L(x, y)$ be a functional defined on the elements of two nonempty sets $X$ and $Y$. Then

$$\sup_{y \in Y} \inf_{x \in X} L(x, y) \leq \inf_{x \in X} \sup_{y \in Y} L(x, y). \quad (3.2)$$

Proof

It is easy to see that

$$L(x, y) \geq \inf_{\xi \in X} L(\xi, y), \quad \forall x \in X, \ y \in Y.$$ 

Taking the supremum over $y \in Y$, we obtain
proof

\[ \sup_{y \in Y} L(x, y) \geq \sup_{y \in Y} \inf_{\xi \in X} L(\xi, y), \quad \forall x \in X. \]

The left-hand side depends on \( x \), while the right-hand side is a number. Thus, we may take infimum over \( x \in X \) and obtain the inequality

\[ \inf \sup_{x \in X} L(x, y) \geq \sup \inf_{y \in Y} L(\xi, y). \]

Therefore, we always have

\[ \sup \mathcal{P}^* \leq \inf \mathcal{P} \]
Duality relations

However, in our case we have a stronger relation, namely

$$\sup \mathcal{P}^* = \inf \mathcal{P}$$

To prove this fact, we note that

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} fv \, dx \quad \forall v \in V_0.$$ 

Therefore $p = \nabla u \in Q_f$ and

$$-l^*(p) = -\frac{1}{2} \| \nabla u \|^2 = \int_{\Omega} \left( \frac{1}{2} |\nabla u|^2 - fu \right) \, dx = J(u).$$
Let us use the Mikhlin’s estimate established in Lecture 2:
\[
\frac{1}{2} \| \nabla (u - v) \|^2 \leq J(v) - J(u).
\]

Since \( J(u) = -l^*(p) \), we have
\[
\frac{1}{2} \| \nabla (u - v) \|^2 \leq J(v) + l^*(p) \leq J(v) + l^*(q) \quad \forall q \in Q_f.
\]

Reform this estimate by using the fact that \( q \in Q_f \).
\[
J(v) + l^*(q) = \frac{1}{2} \| \nabla v \|^2 - (f, v) + \frac{1}{2} \| q \|^2
= \frac{1}{2} \| \nabla v \|^2 + \frac{1}{2} \| q \|^2 - (\nabla v, q) = \frac{1}{2} \| \nabla v - q \|^2.
\]
Now, we have

$$\|\nabla (v - u)\| \leq \|\nabla v - q\| \quad \forall q \in Q_f.$$ 

Take arbitrary $y \in L^2(\Omega)$. Then,

$$\|\nabla (v - u)\| \leq \|\nabla v - y\| + \inf_{q \in Q_f} \|y - q\|.$$ 

**How to estimate the above infimum?**

Various methods give one and the same answer:

$$\inf_{q \in Q_f} \|y - q\| \leq \|\text{div}y + f\| \quad y \in L^2(\Omega), \quad (3.3)$$

$$\inf_{q \in Q_f} \|y - q\| \leq C_\Omega \|\text{div}y + f\| \quad y \in H(\Omega, \text{div}), \quad (3.4)$$
Proof

To prove these estimates we consider an auxiliary problem

$$\Delta \eta + f + \text{div} y = 0 \quad \Omega \quad \eta = 0 \quad \partial \Omega.$$  

$$\int_{\Omega} \nabla \eta \cdot \nabla w \, dx = \int_{\Omega} (f + \text{div} y) w \, dx = \int_{\Omega} (fw - y \cdot \nabla w) \, dx$$  

$$\bar{q}$$

$$\int_{\Omega} \left( \nabla \eta + y \right) \cdot \nabla w \, dx = \int_{\Omega} fw \, dx \quad \forall w \in V_0$$

Thus, $\bar{q} \in Q_f$ !!!
Since $\eta$ is a solution of the boundary–value problem with right–hand side $\text{div } y + f \in H^{-1}$, we have

$$\|\nabla \eta\| \leq \| \text{div } y + f \|,$$

Then

$$\inf_{q \in Q} \| y - q \| \leq \| y - \bar{q} \| = \| \nabla \eta \| \leq \| \text{div } y + f \|.$$ 

Here

$$\| \text{div } y + f \| = \sup_{w \in V_0} \frac{\int_{\Omega} (y \cdot \nabla w - fw) dx}{\| \nabla w \|}.$$
If $y$ has a square summable divergence, then we have

$$\|\text{div}y + f\| = \sup_{w \in V_0} \frac{\int_\Omega (\text{div}y + f) w dx}{\|\nabla w\|} \leq C_\Omega \|\text{div}y + f\|,$$

where $C_\Omega$ is the constant in the Friederichs–Steklov inequality for the domain $\Omega$. Thus, by taking the flux vector–valued function in the space that contains the flux of the true solution we make a "noncomputable" negative norm computable.
Thus, for any \( y \in H(\Omega, \text{div}) \) we obtain

\[
\|\nabla (v - u)\| \leq \|\nabla v - y\| + \inf_{q \in Q_f} \|y - q\| \leq \|\nabla v - y\| + C_\Omega \|\text{div}y + f\|.
\]

Above presented \textit{modus operandi} can be viewed as a simplest version of the variational approach to the derivation of Functional Error Majorants.
Method of integral identities. First glance.

For many problems the variational techniques cannot be applied because they have no variational formulation.

Nonvariational method in the simplest case

Let us expose its simplest version adapted to our model problem. We have seen that

\[ \| \nabla (u - v) \| \leq \| \Delta v + f \| \]

Instead of the estimation of the negative norm by Galerkin orthogonality and special interpolation estimates we suggest another method of finding an upper bound that is based on the functional relation

\[ \int_{\Omega} (\text{div}yw + \nabla w \cdot y) \, dx = 0 \quad \forall w \in V_0 \]
We have

\[
\left[ \Delta v + f \right] = \sup_{w \in V_0} \frac{\int_{\Omega} (\nabla v \cdot \nabla w - fw)}{||\nabla w||} = \\
\sup_{w \in V_0} \frac{\int_{\Omega} (\nabla v \cdot \nabla w - fw - (\text{div} y \cdot w + \nabla w \cdot y))}{||\nabla w||} = \\
\sup_{w \in V_0} \frac{\int_{\Omega} ((\nabla v - y) \cdot \nabla w - (f + \text{div} y)w)dx}{||\nabla w||} = \\
\sup_{w \in V_0} \frac{||\nabla v - y|| ||\nabla w|| + ||f + \text{div} y|| ||w||}{||\nabla w||} \leq \\
||\nabla v - y|| + C_{\Omega} ||f + \text{div} y||.
\]
Functional error estimate. Meaning and properties

For the problem

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

we have obtained the estimate

\[ \| \nabla (u - v) \| \leq \| \nabla v - y \| + C_\Omega \| \text{div } y + f \| \]  \hspace{1cm} (3.5)

The estimate is valid for any \( v \in V_0 \) and \( y \in H(\Omega, \text{div}) \).

The estimate is valid for any \( v \in V_0 \) and \( y \in H(\Omega, \text{div}) \).

Two terms in the right–hand side have a clear sense: they present measures of the errors in two basic relations

\[ p = \nabla u, \quad \text{div } p + f = 0 \quad \text{in } \Omega \]

that jointly form the equation.
The estimate is sharp

If set \( v = 0 \) and \( y = 0 \), we obtain the energy estimate for the generalized solution

\[
\| \nabla u \| \leq C_\Omega \| f \|
\]

Therefore, no constant less than \( C_\Omega \) can be stated in the second term. If set \( y = \nabla u \), than the inequality holds as the equality.

Thus, we see that the estimate (3.5) is sharp in the sense that the multipliers of both terms cannot be taken smaller and in the set of admissible \( y \) there exists a function that makes the inequality hold as equality.
The estimate as a quadratic functional

By means of the algebraic Young’s inequality

\[ 2ab \leq \beta a^2 + \frac{1}{\beta} b^2, \quad \beta > 0 \]

we rewrite this estimate in the form

\[
\| \nabla (u - v) \|^2 \leq (1 + \beta) \| \nabla v - y \|^2 + \frac{1 + \beta}{\beta} C_\Omega^2 \| \text{div} y + f \|^2
\]  

(3.6)

For any \( \beta \) the right–hand side is a quadratic functional. This property makes it possible to apply well known methods for the minimization with respect to \( y \).
Denote the right–hand side of (3.6) by $M_\oplus$, i.e.,

$$M_\oplus(v, y, \beta, C_\Omega, f) := (1 + \beta) \| \nabla v - y \|^2 + \frac{1 + \beta}{\beta} C_\Omega^2 \| \text{div} y + f \|^2.$$ 

This functional provides an upper bound for the norm of the deviation of $v$ from $u$. Therefore, it is natural to call it the Deviation Majorant.
BVP $\Delta u + f = 0$ has another variational formulation

$$\inf_{v \in V_0, \ \beta > 0, \ y \in H(\Omega, \text{div})} M_\oplus(v, y, \beta, C_\Omega, f)$$

- Minimum of this functional is zero;
- it is attained if and only if $v = u$ and $y = A \nabla u$;
- $M_\oplus$ contains only one global constant $C_\Omega$, which is problem independent;
In principle, one can select certain sequences of subspaces \( \{V_{hk}\} \in V_0 \) and \( \{Y_{hk}\} \in H(\Omega, \text{div}) \) and minimize the Error Majorant with respect to these subspaces

\[
\inf_{v \in V_{hk}}, \beta > 0, y \in Y_{hk} \quad M(\{v, y, \beta, C_\Omega, f\})
\]

If the subspaces are limit dense, then we would obtain a sequence of approximate solutions \((v_k, y_k)\) and the sequence of numbers

\[
\gamma_k := \inf_{\beta > 0} M(\{v_k, y_k, \beta, C_\Omega, f\}) \rightarrow 0
\]
How to use the Majorants in practice?

Consider **CONFORMING FEM APPROXIMATIONS**.

We have 3 basic ways to use the deviation estimate:
(a) Use flux averaging on the mesh $\mathcal{T}_h$;
(b) Use flux averaging on the refined mesh $h_{ref}$;
(c) Minimization with respect to $y$. 
(a) Use recovered gradients

Let $u_h \in V_h$, then

$$p_h := \nabla u_h \in L_2(\Omega, \mathbb{R}^d), \quad p_h \notin H(\Omega, \text{div}).$$

Use an averaging operator $G_h : L_2(\Omega, \mathbb{R}^d) \to H(\Omega, \text{div})$ and have a directly computable estimate

$$\|\nabla(u - u_h)\| \leq \|\nabla u_h - G_h p_h\| + C_\Omega \|\text{div}G_h p_h + f\|$$
(b) Use recovered gradients from $\mathcal{T}_{h_{\text{ref}}}$

Let $u_1, u_2, ..., u_k, ...$ be a sequence of approximations on meshes $\mathcal{T}_{h_k}$. Compute $p_k := \nabla u_k$, average it by $G_k$ and for $u_{k-1}$ use the estimate

$$\|u - u_{k-1}\| \leq \|\nabla u_{k-1} - G_k p_k\| + C_\Omega \|\text{div} G_k p_k + f\|$$

This estimate gives a quantitative form of the Runge’s rule.
(c) Minimize $\mathcal{M}_\oplus$ with respect to $y$.

Select a certain subspace $Y_\tau$ in $H(\Omega, \text{div})$. Generally, $Y_\tau$ may be constructed on another mesh $T_\tau$ and with help of different trial functions. Then

$$\| \nabla (u - u_h) \| \leq \inf_{y_h \in Y_h} \left\{ \| \nabla u_h - y_h \| + C_\Omega \| \text{div} y_h + f \| \right\}$$

The wider $Y_h \subset H(\Omega, \text{div})$ the sharper is the upper bound.
Quadratic type functional

From the technical point of view it is better to square both parts of the estimate and apply minimization to a quadratic functional, namely

\[ \| \nabla (u - u_h) \|^2 \leq \inf_{y_h \in Y_h} \left\{ (1 + \beta) \| \nabla u_h - y_h \| + 
\right. \]
\[ \left. + C_\Omega \left( 1 + \frac{1}{\beta} \right) \| \text{div} y_h + f \|^2 \right\} \]

Here, the positive parameter $\beta$ can be also used to minimize the right-hand side.
Before going to more complicated problems where Deviation Majorants are derived by a more sophisticated theory, we observe several simple examples that nevertheless reflect key points of the above method.
Simple 1-D problem

\[(\alpha(x)u')' = f(x),\]
\[u(a) = 0, \quad u(b) = u_b.\]

It is equivalent to the variational problem

\[J(v) = \int_a^b \left( \frac{1}{2} \alpha(x) |v'|^2 + f(x)v \right) \, dx.\]

Assume that the coefficient \(\alpha\) belongs to \(L^\infty\) and bounded from below by a positive constant. Now

\[V_0 + u_0 = \{v \in H^1(a, b) \mid v(a) = 0, v(b) = u_b\}.\]
Deviation Majorant

\[ \mathcal{M}_\oplus(v, \beta, y) = (1 + \beta) \left( \int_a^b |\alpha v' - y|^2 \, dx + \frac{C^2_{(a,b)}}{\beta} \int_a^b |y' - f|^2 \right) \, dx. \]

In this simple model, \( u \) can be presented in the form

\[
    u(x) = \int_a^x \frac{1}{\alpha(t)} \int_a^t f(z) \, dz \, dt + \frac{x}{b} \left( u_b - \int_a^b \frac{1}{\alpha(t)} \int_a^t f(z) \, dz \, dt \right).
\]

what gives an opportunity to verify how error estimation methods work.
Approximations

Let \( \mathbf{V}_h \) be made of piecewise-\( P^1 \) continuous functions on uniform splittings of the interval and consider approximations of the following types:

- Galerkin approximations;
- Approximations very close to Galerkin (sharp);
- Approximations which are "good" but not Galerkin;
- Coarse (rough) approximations.

Our aim is to show that the Deviation Majorant can be effectively used as an error estimation instrument in all the above cases.
Computation of the Majorant

To find a sharp upper bound, we minimize $\mathcal{M}_\oplus$ with respect to $y$ and $\beta$ starting from the function $y_0 = G(v')$, where $G$ is a simple averaging operator, e.g., defined by the relations

$$G(v')(x_i) = \frac{1}{2}(v'(x_i - 0) + v'(x_i + 0)),$$

By the quantity

$$\inf_{\beta > 0} \mathcal{M}_\oplus(v, \beta, y_0),$$

we obtain a coarse upper bound of the error. It is further improved by minimizing $\mathcal{M}_\oplus$ with respect to $y$. 
Example 1

Let $\alpha(x) = 1$, $f(x) = c$, $a = 0$, $b = 1$, $u_b = 1$, e.g., we consider the problem

$$u'' = 2, \quad u(0) = 0, \quad u(1) = 1.$$  

In this case, $C_{(a,b)} = 1/\pi$ and

$$u = \frac{c}{2} x^2 + (1 - \frac{c}{2})x, \quad u' = cx + 1 - \frac{c}{2}.$$  

Take a rough approximation $v = x$. Then

$$\|u - v\|^2 = \int_0^1 c^2(x - 0.5)^2 dx = \frac{c^2}{12} \approx 0.083c^2.$$
Exact solution and an approximation.
Various $y$ give different upper bounds

(a) Take $y = v' = 1$, then the first term in

$$\mathcal{M}_\oplus(v, \beta, y) = (1 + \beta) \left( \int_0^1 |v' - y|^2 \, dx + \frac{1}{\pi^2 \beta} \int_0^1 |y' - f|^2 \right) \, dx.$$ 

vanishes and we have $\mathcal{M}_\oplus \to c^2 / \pi^2 \approx 0.101c^2$; as $\beta \to +\infty$. We see that this upper bound overestimates true error. Note that in this case, all sensible averagings of $v' = 1$ give exactly the same function: $G(1) = 1$! Therefore,

$$G(v') - v' \equiv 0$$

and formally ZZ indicator ”does not see the error”.
For the choice $y = v'$ the Majorant give a certain upper bound of the error (which is not so bad), but the integrand cannot indicate the distribution of local errors. Indeed, we have

$$M_\oplus = \frac{1}{\pi^2} \int_0^1 c^2 \, dx.$$ 

However, the integrand of the Majorant is a constant function, but the error is distributed in accordance with a parabolic law:

$$(u - v)' = c(x - 0.5)^2.$$
(b). Take $y = cx + 1 - c/2$. Then, $y' = c$ and the second term of the majorant vanishes. We have (for $\beta = 0$)

$$M_\oplus = \int_0^1 c^2(x - 1/2)^2 \, dx = c^2/12.$$ 

We observe that both the global error and the error distribution are exactly reproduced. In real life computations such an ”ideal” function $y$ may be unattainable. However, the minimization makes the Majorant close to the sharp value. In this elementary example, we have minimized the Majorant on using piecewise affine approximations of $y$ on 20 subintervals. The elementwise error distribution obtained as the result of this procedure is exposed on the next picture.
True errors and those computed by the Majorant.

S. Repin

LECTURES ON A POSTERIORI ERROR CONTROL
To give further illustrations, we consider the functions

\[ u_\delta = u + \delta \phi, \]

where \( \delta \) is a number and \( \phi \) is a certain function (perturbation).

**Approximate solutions (whose errors are measured)** are piecewise affine continuous interpolants of \( u_\delta \) defined on a uniform mesh with 20 subintervals.

We take \( \phi = x \sin(\pi x) \) and \( \delta = 0.1, 0.01, 0.001, \) and 0.

<table>
<thead>
<tr>
<th>( \delta )</th>
<th>( e )</th>
<th>( 2M_{\oplus} )</th>
<th>( 2M_{\ominus} )</th>
<th>( i_{\text{eff}} )</th>
<th>( i_{\text{esh}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.019692</td>
<td>0.019743</td>
<td>0.019683</td>
<td>1.003</td>
<td>1.018</td>
</tr>
<tr>
<td>0.01</td>
<td>0.001022</td>
<td>0.001025</td>
<td>0.001013</td>
<td>1.003</td>
<td>1.011</td>
</tr>
<tr>
<td>0.001</td>
<td>0.000835</td>
<td>0.000839</td>
<td>0.000827</td>
<td>1.005</td>
<td>1.002</td>
</tr>
<tr>
<td>0</td>
<td>0.000833</td>
<td>0.000836</td>
<td>0.000825</td>
<td>1.004</td>
<td>1.002</td>
</tr>
</tbody>
</table>

In this experiment the Majorant was computed for \( \frac{1}{2} \| e \|^2 \).
Error estimation for $\delta = 0.1$
Functions $u$, $v$ and $i_{eff}$ for $\delta = 0.1$
Error estimation for $\delta = 0.01$

A more precise approximation.
Functions $e(y)$, $\beta$ and $i_{\text{eff}}$ for $\delta = 0.1$
Error estimation for $\delta = 0.01$

A more precise approximation.

![Graph showing error estimation for DEM and GA]
Functions $e(y)$, $\beta$ and $i_{\text{eff}}$ for $\delta = 0.01$
Error estimation for $\delta = 0.001$

Sharp approximation.

Error bars for DEM and GA.
Functions $e(y)$, $\beta$ and $i_{\text{eff}}$ for $\delta = 0.001$
Error estimation for $\delta = 0$

Interpolant of the exact solution.

![Graph showing error estimation for DEM and GA methods.](graph.png)
Functions $e(y)$, $\beta$ and $i_{eff}$ for $\delta = 0$
Functional a posteriori error estimates were derived by the methods of duality theory in convex analysis in 1996. These results are published in [7,8]. In [5,9,23,24], they were applied to certain linear and nonlinear variational problems with convex functionals. First consequent study of their computational properties was presented in [10]. Later a detailed investigation of the practical aspects was done in [1,3,18]. General a posteriori estimates for the class of convex functionals are derived and discussed in [3,5,11,12]. A posteriori estimates for a class of nonconvex problems are can be found in [6]. A posteriori estimates which take into account errors in main boundary conditions were derived in [20], there readers can also find a method of the derivation of the estimates based upon the orthogonal decomposition of the space $L^2$ (Helmgholts decomposition). In, [17,21,22] a posteriori estimates were derived for modeling errors in dimension reduction models. Estimates for the Stokes problem will be further discussed in this lecture course (see [15,16]). In [19], functional type a posteriori estimates were obtained for the Reissner-Mindlin plate.


(POMI), 288(2002), 178-203.


Lecture 4.
AN INTRODUCTION TO DUALITY THEORY.
Lecture goal

In subsequent lectures we will present the general theory of a posteriori error control for convex variational problems. In the framework of this theory we are able to derive computable upper bounds for the errors for problems of the type

$$\inf_{v \in V} J(v, \Lambda v), \quad J(v, \Lambda v) := G(\Lambda v) + F(v),$$

where $\Lambda : V \to Y$ is a linear continuous operator from a Banach space $V$ to another Banach space $Y$ and $J : Y \to \mathbb{R}$ and $F : V \to \mathbb{R}$ are convex l.s.c. functionals.
In particular, if
\[ \Lambda v = \nabla v, \quad G(y) = (Ay, y), \quad F(v) = (f, v), \]
then we arrive to the variational formulation of the problem
\[ \text{div } A \nabla u + f = 0 \]
with certain boundary conditions.
Many other problems have the above form, were

- **G** is the **energy functional** whose form is dictated by
  the dissipative properties of a media.
- **F** is the functional associated with **external forces**.
Many problems in continuum mechanics encompassed in the general scheme are: linear elasticity, biharmonic problems, Kirhghoff and Mindlin plates, deformation theory of elastoplasticity, Stokes problem. Also, this scheme is applicable to the p-Laplace equation and certain nonlinear models in the theory of viscous fluids.

In such models the structure of the "energy functional" $G$ plays crucial role in all the parts of the mathematical analysis: existence and differentiability properties of minimizers and estimates of deviations from the minimizers.
To understand the basic principles of the functional approach to the derivation of a posteriori bounds of the approximation errors we need to make a concise overview of some parts of the duality theory in the calculus of variations.
Lecture plan

- Dual and bidual functionals;
- Compound functionals;
- Uniformly convex functionals.
Dual (polar) functionals

Hereafter $V^*$ contains all linear continuous functionals defined on $V$. The elements of $V^*$ are marked by stars, $\langle v^*, v \rangle$ is called the **duality pairing** of the spaces $V$ and $V^*$. Let $J : V \rightarrow \mathbb{R}$, then $J^*$ defined by the relation

$$J^*(v^*) = \sup_{v \in V} \{\langle v^*, v \rangle - J(v)\}$$

is called **dual** to $J$.

If $J$ is a smooth function that increases at infinity faster than any linear function, then $J^*$ is the Legendre transform of $J$. The above general definition comes from Young and Fenchel. The functional $J^*$ is also called **polar** to $J$. 

S. Repin

*RICAM, Special Radon Semester, Linz, 2005.*
Bipolar functionals

The functional

\[ J^{**}(v) = \sup_{v^* \in V^*} \{ \langle v^*, v \rangle - J^*(v^*) \} \]

is called the \textbf{bidual} to \( J \) (or \textbf{bipolar}).

Straightforwardly from the definition, it follows that \( J^* \) and \( J^{**} \) are convex functionals (they are defined as upper bounds of affine functionals). Formally, one can also define

\[ J^{***}(v^*) := \sup_{v \in V} \{ \langle v^*, v \rangle - J^{**}(v) \}. \]

However, this definition brings nothing new. It is proved that

\[ J^{***}(v^*) = J^*(v^*), \quad \forall v^* \in V^*. \]
Mutually dual functionals

Let $J : V \to \overline{\mathbb{R}} := \{\mathbb{R}, -\infty, +\infty\}$ and $G^* : V^* \to \overline{\mathbb{R}}$ be two functionals defined on a Banach space $V$ and its dual space $V^*$, respectively. These two functionals are called mutually dual if

$$(G^*)^* = J \quad \text{and} \quad J^* = G^*.$$
Examples

To illustrate the definitions of conjugate functionals, we present below several examples for functionals defined on the Euclidean space $\mathbb{E}^d$. In this case, $\mathbf{V}$ and $\mathbf{V}^*$ are isometrically isomorphic. Their elements are $d$-dimensional vectors denoted by $\xi$ and $\xi^*$, respectively, so that

$$\langle \xi^*, \xi \rangle = \xi^* \cdot \xi = \xi_i^* \xi_i.$$

These examples have a practical meaning because for a wide class of integral type functionals (in the mechanics they are the energy functionals) finding the dual energy functional is reduced to finding dual to its integrand!
In other words, if the ”primal energy functional” has the form

$$G(v) := \int_{\Omega} g(\Lambda v) dx$$

where \( g \) is the ”internal energy” or ”dissipative potential”, then
the so–called ”complementary energy” is given by the integral functional

$$G^*(y^*) := \int_{\Omega} g^*(y^*) dx,$$

where \( g^* \) is conjugate to \( g \) in the algebraic sense.
Example 1 (Diffusion problems)

Let $A = \{a_{ij}\}$ be a real, positive definite matrix and

$$g(\xi) = \frac{1}{2} A \xi \cdot \xi = \frac{1}{2} a_{ij} \xi_i \xi_j.$$ 

Then

$$g^*(\xi^*) = \sup_{\xi \in E^d} \left\{ \xi^* \cdot \xi - \frac{1}{2} A \xi \cdot \xi \right\}.$$ 

This supremum is attained on an element $\xi_0$ such that

$$\xi^* = A \xi_0 \implies \xi_0 = A^{-1} \xi^*.$$ 

Therefore, we have a pair of mutually conjugate functionals

$$g(\xi) = \frac{1}{2} A \xi \cdot \xi \quad \text{and} \quad g^*(\xi^*) = \frac{1}{2} A^{-1} \xi^* \cdot \xi^*.$$
In diffusion type boundary–value problems we arrive at the functional (with $y = \nabla v$)

$$\frac{1}{2} \int_{\Omega} Ay \cdot y \, dx \quad y \in L^2(\Omega, \mathbb{R}^n),$$

which is mutually dual to

$$\frac{1}{2} \int_{\Omega} A^{-1} y^* \cdot y^* \, dx \quad y^* \in L^2(\Omega, \mathbb{R}^n)$$
Example 2 (Linear elasticity)

Let $\mathbf{L} = \{ L_{ijkl} \}$ be a real, positive definite tensor of the 4-th order and $\mathbf{\tau}$ be a tensor of the second order ($d \times d$–matrix). Then,

$$g(\varepsilon) = \frac{1}{2} \mathbf{L} \varepsilon : \varepsilon = \frac{1}{2} L_{ijkl} \varepsilon_{ij} \varepsilon_{km}.$$ 

Then

$$g^*(\varepsilon^*) = \sup_{\varepsilon \in \mathbb{M}^{d \times d}} \left\{ \varepsilon^* : \varepsilon - \frac{1}{2} \mathbf{A} \varepsilon : \varepsilon \right\}.$$ 

This supremum is attained on an element $\varepsilon_0$ such that

$$\mathbf{\tau}^* = \mathbf{L} \varepsilon_0 \quad \Rightarrow \quad \varepsilon_0 = \mathbf{L}^{-1} \varepsilon^*.$$ 

Therefore, we have a pair of mutually dual functionals

$$g(\varepsilon) = \frac{1}{2} \mathbf{L} \varepsilon : \varepsilon \quad \text{and} \quad g^*(\varepsilon^*) = \frac{1}{2} \mathbf{L}^{-1} \varepsilon^* : \varepsilon^*. $$
In linear elasticity problems we arrive at the energy functional in terms of strains $\varepsilon(v) = \frac{1}{2}(\nabla v + (\nabla v)^T)$

$$\frac{1}{2} \int_\Omega \mathbb{L} \varepsilon : \varepsilon \, dx \quad \varepsilon \in L^2(\Omega, \mathbb{M}^{n \times n}),$$

which is mutually dual to the ”complementary energy” functional written in terms of stresses $\varepsilon^*(x) \Rightarrow \tau(x)$

$$\frac{1}{2} \int_\Omega \mathbb{L}^{-1} \tau : \tau \, dx \quad \tau \in L^2(\Omega, \mathbb{M}^{n \times n})$$
Example 3 (Nonlinear elasticity, p-Laplacian)

Consider the functional

\[ g(\xi) = \frac{1}{p} |\xi|^p, \]

where \( p > 1 \) and \( |\xi| = (\xi \cdot \xi)^{1/2} \). It is easy to verify that the quantity \( \xi^* \cdot \xi - \frac{1}{p} |\xi|^p \) attains a supremum if \( \xi = \xi_0 \), where \( \xi_0 \) satisfies the relation

\[ \xi^* - |\xi_0|^{p-2} \xi_0 = 0, \]

which yields \( |\xi^*| = |\xi_0|^{p-1} \) and \( \xi^* \cdot \xi_0 = |\xi_0|^p \). Therefore,

\[ g^*(\xi^*) = \xi^* \cdot \xi_0 - \frac{1}{p} |\xi_0|^p = \left( 1 - \frac{1}{p} \right) |\xi_0|^p = \frac{1}{p^*} |\xi^*|^{p^*}, \]

where \( p^* = \frac{p}{p-1} \).
Thus, we obtain another pair of mutually conjugate functionals

\[ g(\xi) = \frac{1}{p}|\xi|^p \quad \text{and} \quad g^*(\xi^*) = \frac{1}{p^*}|\xi^*|^{p^*}, \]

where \( \frac{1}{p} + \frac{1}{p^*} = 1 \).

Remark

This relation admits generalizations. Namely, let \( \varphi : \mathbb{R} \rightarrow \mathbb{R} \) be a proper convex function that is, in addition, odd and let \( \varphi^* : \mathbb{R} \rightarrow \mathbb{R} \) be its conjugate. Then

\[ (\varphi(\|\mathbf{u}\|\mathbf{v}))^* = \varphi^*(\|\mathbf{u}^*\|\mathbf{v}^*). \]
In certain nonlinear boundary–value problems we arrive at the functional (with $y = \nabla v$ or $y = \varepsilon(v)$)

$$\frac{1}{p} \int_{\Omega} |y|^p \, dx \quad y \in L^p(\Omega, \mathbb{R}^n[M^{n \times n}]),$$

which is mutually dual to

$$\frac{1}{p^*} \int_{\Omega} |y^*|^{p^*} \, dx \quad y^* \in L^{p^*}(\Omega, \mathbb{R}^n[M^{n \times n}]).$$
Example 4 (Action of external forces)

Let $g(\xi)$ be a linear functional, i.e.,

$$g(\xi) = \ell \cdot \xi, \quad \ell \in \mathbb{E}^d.$$  

It is easy to see that

$$g^*(\xi^*) = \sup_{\xi \in \mathbb{E}^d} \{\xi^* \cdot \xi - \ell \cdot \xi\} = \begin{cases} 0 & \xi^* = \ell, \\ +\infty & \xi^* \neq \ell. \end{cases}$$

Denote by $\mathcal{X}_{\{\ell\}}$ the characteristic functional of the set $\{\ell\} \subset \mathbb{E}^d$. Then, another pair of mutually conjugate functionals is as follows:

$$g(\xi) = \ell \cdot \xi \quad \text{and} \quad g^*(\xi^*) = \mathcal{X}_{\{\ell\}}(\xi^*).$$
Thus, for the functional $G : L^2 \to \mathbb{R}$

$$G(v) := \int_{\Omega} fv \, dx, \quad f \in L^2(\Omega)$$

the respective dual functional is $G^* : L^2 \to \mathbb{R}$

$$G^*(v^*) = 0 \text{ if } v^* = f, \quad G^*(v^*) = +\infty \text{ in other cases.}$$
Example 5 (Variational inequalities, friction)

Let \( g(\xi) = |\xi| \). Then

\[
\sup_{\xi} \{ \xi^* \cdot \xi - |\xi| \}
\]

may be finite or infinite depending on the value of \( |\xi^*| \). If \( |\xi^*| > 1 \), then, obviously, it is infinite. If \( |\xi^*| \leq 1 \), then, on the one hand,

\[
\sup_{\xi} \{ \xi^* \cdot \xi - |\xi| \} \leq \sup_{\xi} \{ 1|\xi| - |\xi| \} = 0.
\]

On the other hand, \( \sup_{\xi} \{ \xi^* \cdot \xi - |\xi| \} \geq \xi^* \cdot 0 - 0 = 0 \). This means that \( g^*(\xi^*) = 0 \) if \( |\xi^*| \leq 1 \) and, thus,

\[
g(\xi) = |\xi|, \quad g^*(\xi^*) = \mathcal{X}_{B^*(0,1)}(\xi^*), \quad \text{where } B^*(0,1) = \{ \xi^* \in E^d \mid |\xi^*| \leq 1 \}
\]
Thus, for the functional $G : L^1 \rightarrow \mathbb{R}$

$$G(v) := \int_{\Omega} |v| \, dx,$$

the respective dual functional is $G^* : L^\infty \rightarrow \mathbb{R}$

$$G^*(v^*) = 0 \text{ if } |v^*(x)| \leq 1 \text{ a.e. in } \Omega, \quad G^*(v^*) = +\infty \text{ in other cases.}$$
Example 6 (Variational inequalities, perfect plasticity)

Let $K$ be a convex closed set in $E^d$ and

$$g(\xi) = \mathcal{X}_K(\xi).$$

The respective conjugate functional is defined as follows:

$$g^*(\xi^*) = \sup_{\xi \in E^d} \{\xi^* \cdot \xi - \mathcal{X}_K(\xi)\} = \sup_{\xi \in K} \xi^* \cdot \xi.$$

This function is called the support function of $K$ and is denoted by $\mathcal{X}_K^*(\xi^*)$. For example, if $K = B(0, 1)$, then

$$\sup_{\xi \in K} \xi^* \cdot \xi = |\xi^*| \Rightarrow \mathcal{X}_{B(0,1)}^*(\xi^*) = |\xi^*|.$$
Example 7 (Elasto-plasticity)

Let us find conjugate for the functional

\[ g^*(\xi^*) = \frac{k}{2}|\xi^*|^2 + \mathcal{X}_{B^*(0,\lambda)}(\xi^*), \quad k > 0, \; \lambda > 0. \]

In this case,

\[ g(\xi) = \sup_{\xi^* \in B^*(0,\lambda)} \{ \xi^* \cdot \xi - \frac{k}{2}|\xi^*|^2 \}. \]

If \( \xi_0^* \) meets the relation \( \xi = k\xi_0^* \) and satisfies the condition \( |\xi_0^*| \leq \lambda \), then it is the required maximizer. For such a \( \xi_0^* \) we have

\[ \xi \cdot \xi_0^* - \frac{k}{2}|\xi_0^*|^2 = \frac{1}{k}|\xi_0^*|^2 - \frac{1}{2k}|\xi_0^*|^2 = \frac{1}{2k}|\xi|^2. \]
If $|\xi_0^*| > \lambda$, then the maximizer $\xi_m^*$ meets the conditions

$$|\xi_m^*| = \lambda, \quad \xi \cdot \xi_m^* \geq \xi^* \cdot \xi^*, \quad \forall \xi^* \in B^*(0, \lambda),$$

which mean that $\xi_m^* = \lambda \frac{\xi}{|\xi|}$ and, consequently,

$$\xi \cdot \xi_m^* - \frac{k}{2} |\xi_m^*|^2 = \lambda |\xi| - \frac{k}{2} \lambda^2.$$

Thus, we obtain

$$g(\xi) = \begin{cases} 
\frac{1}{2k} |\xi|^2 & \text{if } |\xi| \leq k\lambda, \\
\lambda |\xi| - \frac{k}{2} \lambda^2 & \text{if } |\xi| > k\lambda.
\end{cases}$$
In the theory of perfect elasto–plasticity stresses are subject to the condition $\tau \in K = \text{plastic yield set}$ and the stress energy functional is defined and finite only on such $\tau$:

$$G^*(\tau) = \frac{1}{2} \int_\Omega \mathbb{L}^{-1} \tau : \tau \, dx \quad \text{for} \quad \tau \in K.$$ 

The respective dual functional (for strains) is given by a linear growth functional

$$G(\varepsilon) = \int_\Omega g(\varepsilon) \, dx,$$

where $g$ is a linear growth functional of the type given on the previous page.
Example 8 (Minimal surfaces, capillary problems)

Consider the functional \( g(\xi) = \sqrt{1 + |\xi|^2} \), arising in some variational problems having a geometrical meaning (e.g., for the nonparametric minimal surface problem). If \(|\xi^*| > 1\), then the value of

\[
\sup_{\xi \in \mathbb{E}^d} \left\{ \xi^* \cdot \xi - \sqrt{1 + |\xi|^2} \right\}
\]

is infinite. If \(|\xi^*| \leq 1\), then the maximizer \( \xi_0 \) satisfies the condition

\[
\xi^* - \frac{\xi_0}{\sqrt{1 + |\xi_0|^2}} = 0, \quad \Rightarrow \quad |\xi_0|^2 = \frac{|\xi^*|^2}{1 - |\xi^*|^2}.
\]

Therefore, we obtain

\[
g^*(\xi^*) = \begin{cases} 
-\sqrt{1 - |\xi^*|^2} & \text{if } |\xi^*| \leq 1, \\
+\infty & \text{if } |\xi^*| > 1.
\end{cases}
\]
Energy functional for the minimal surface problem (with $y = \nabla v$)

\[ \int_{\Omega} \sqrt{1 + |y|^2} \, dx \quad y \in L^1(\Omega, \mathbb{R}^2), \]

which is mutually dual to

\[ - \int_{\Omega} \sqrt{1 - |y^*|^2} \, dx \quad |y^*| \leq 1. \]
Properties of dual functionals

Property 1

If \( J : V \to \mathbb{R} \) and \( G : V \to \mathbb{R} \) are such that
\[
J(v) \geq G(v), \quad \forall v \in V,
\]
then
\[
J^*(v^*) \leq G^*(v^*), \quad \forall v^* \in V^*.
\]

Proof. We have
\[
J^*(v^*) = \sup_{v \in V} \{ \langle v^*, v \rangle - J(v) \} \leq \sup_{v \in V} \{ \langle v^*, v \rangle - G(v) \} = G^*(v^*).
\]

S. Repin

LECTURES ON A POSTERIORI ERROR CONTROL
Property 2

For any $\lambda > 0$,

$$(\lambda J)^*(v^*) = \lambda J^* \left( \frac{v^*}{\lambda} \right).$$

**Proof.** This property is justified by direct calculations:

$$(\lambda J)^*(v^*) = \sup_{v \in V} \{ \langle v^*, v \rangle - \lambda J(v) \} =$$

$$= \lambda \sup_{v \in V} \left\{ \left\langle \frac{v^*}{\lambda}, v \right\rangle - J(v) \right\} = \lambda J^* \left( \frac{v^*}{\lambda} \right).$$
Property 3

Let $J : \mathbf{V} \to \overline{\mathbb{R}}$ and $J_\alpha(v) = J(v) + \alpha$, where $\alpha \in \mathbb{R}$. Then

$$J^*_\alpha(v^*) = J^*(v^*) - \alpha.$$ 

**Proof.** It follows from the obvious relation

$$\sup_{v \in \mathbf{V}} \{\langle v^*, v \rangle - J(v) - \alpha\} = \sup_{v \in \mathbf{V}} \{\langle v^*, v \rangle - J(v)\} - \alpha.$$
Property 4

Let $v_0 \in V$ and $G(v) = J(v - v_0)$. Then

$$G^*(v^*) = J^*(v^*) + \langle v^*, v_0 \rangle.$$

**Proof.** Since

$$\sup_{v \in V} \{\langle v^*, v \rangle - J(v - v_0)\} = \sup_{w \in V} \{\langle v^*, w + v_0 \rangle - J(w)\}$$

$$= \sup_{w \in V} \{\langle v^*, w \rangle - J(w)\} + \langle v^*, v_0 \rangle = J^*(v^*) + \langle v^*, v_0 \rangle,$$

we arrive at the required relation.
Property 5

If \( G(v) = \min_{i=1,\ldots,N} \{ J_i(v) \} \), then \( G^*(v^*) = \max_{i=1,\ldots,N} \{ J_i^*(v^*) \} \).

**Proof.** We have

\[
G^*(v^*) = \sup_{v \in V} \{ \langle v^*, v \rangle - \min_{i=1,\ldots,N} \{ J_i(v) \} \}
\]

\[
= \sup_{v \in V} \{ \langle v^*, v \rangle + \max_{i=1,\ldots,N} \{ -J_i(v) \} \}
\]

\[
= \sup_{v \in V} \max_{i=1,\ldots,N} \{ \langle v^*, v \rangle - J_i(v) \}
\]

\[
= \max_{i=1,\ldots,N} \sup_{v \in V} \{ \langle v^*, v \rangle - J_i(v) \} = \max_{i=1,\ldots,N} \{ J_i^*(v^*) \}.
\]
Property 6

If $G(v) = \max_{i=1,...,N} \{J_i(v)\}$, then $G^*(v^*) \leq \min_{i=1,...,N} \{J_i^*(v^*)\}$.

Proof. By definition, we have

$$G^*(v^*) = \sup_{v \in V} \{\langle v^*, v \rangle - \max_{i=1,...,N} \{J_i(v)\} \}$$

$$= \sup_{v \in V} \{\langle v^*, v \rangle + \min_{i=1,...,N} \{-J_i(v)\} \}$$

$$= \sup_{v \in V} \min_{i=1,...,N} \{\langle v^*, v \rangle - J_i(v) \}.$$ 

Now we apply $\sup \inf \leq \inf \sup$ relation to $\langle v^*, v \rangle - J_i(v)$. Then,

$$G^*(v^*) \leq \min_{i=1,...,N} \sup_{v \in V} \{\langle v^*, v \rangle - J_i(v) \} = \min_{i=1,...,N} \{J_i^*(v^*)\}.$$
The functional $J: V \rightarrow \mathbb{R}$ is called subdifferentiable at $v_0$ if there exists an affine minorant $\ell \in \text{AM}(J)$ such that $J(v_0) = \ell(v_0)$. A minorant with this property is called the exact minorant at $v_0$.

Obviously, any affine minorant exact for $J$ at $v_0$ has the form

$$\ell(v) = \langle v^*, v - v_0 \rangle + J(v_0), \quad \ell(v) \leq J(v), \quad \forall v \in V.$$ 

The element $v^*$ is called a subgradient of $J$ at $v_0$. 
The set of all subgradients of $J$ at $v_0$ forms a **subdifferential**, which is usually denoted by $\partial J(v_0)$. It may be empty or contain one element or infinitely many elements.

An important property of convex functionals follows directly from the above definition. For a convex functional $J$ at a point $v_0$ where it is finite, the exact affine minorant is evidently **exist**!

In other words, there is at least one element $v^* \in \partial J(v_0)$ that ”creates” an affine minorant such that

$$\langle v^*, v \rangle - \alpha \leq J(v), \quad \forall v \in V,$$

$$\langle v^*, v_0 \rangle - \alpha = J(v_0).$$

By subtracting, we obtain

$$J(v) - J(v_0) \geq \langle v^*, v - v_0 \rangle.$$

The inequality (4.1) presents the **basic incremental relation for convex functionals**.
Compound functionals

Let $J$ and $J^*$ be a pair of mutually dual convex functionals.

The functional $D_J : V \times V^* \rightarrow \mathbb{R}$ of the form

$$D_J(v, v^*) := J(v) + J^*(v^*) - \langle v^*, v \rangle.$$ 

is called it the **compound functional** associated with these pair of functionals.

We will see that compound functionals play an important role in the a posteriori analysis of linear and nonlinear variational problems.
Compound functionals are always nonnegative. Indeed,

\[ J^*(v^*) = \sup_{v \in V} (\langle v^*, v \rangle - J(v)) \geq \langle v^*, v \rangle - J(v) \quad \forall v \in V \]

and

\[ J^*(v^*) + J(v) - \langle v^*, v \rangle \geq 0 \quad \forall v, v^* \]
Compound functionals may vanish only on special sets, where \( \mathbf{v} \) and \( \mathbf{v}^* \) satisfy certain relations.

**Theorem**

Let \( \mathbf{J} \) be a proper convex functional and \( \mathbf{J}^* \) be its polar. Then, the following two statements are equivalent:

\[
\mathbf{J}(\mathbf{v}) + \mathbf{J}^*(\mathbf{v}^*) - \langle \mathbf{v}^*, \mathbf{v} \rangle = 0,
\]

(4.1)

\[
\mathbf{v}^* \in \partial \mathbf{J}(\mathbf{v}) \text{ and } \mathbf{v} \in \partial \mathbf{J}^*(\mathbf{v}^*).
\]

(4.2)

Relations (4.2) are also called **duality relations** for the pair \( (\mathbf{v}, \mathbf{v}^*) \).
Proof.

Assume that $v^* \in \partial J(v)$, i.e,

$$J(w) \geq J(v) + \langle v^*, w - v \rangle, \quad \forall w \in V.$$ 

Hence,

$$\langle v^*, v \rangle - J(v) \geq \langle v^*, w \rangle - J(w), \quad \forall w \in V$$

and, consequently,

$$\langle v^*, v \rangle - J(v) \geq \sup_{w \in V} \{ \langle v^*, w \rangle - J(w) \} = J^*(v^*),$$

what leads to the conclusion that $J^*(v^*) + J(w) - \langle v^*, w \rangle \leq 0$. But the left–hand side is nonnegative, so that we obtain $D_J(v^*, v) = 0$. 

S. Repin

*RICAM, Special Radon Semester, Linz, 2005.*

**LECTURES ON A POSTERIORI ERROR CONTROL**
Assume that $v \in \partial J^*(v^*)$. Then

$$J^*(w^*) \geq J^*(v^*) + \langle w^* - v^*, v \rangle,$$

and we continue similarly to the previous case:

$$\langle v^*, v \rangle - J^*(v^*) \geq \langle w^*, v \rangle - J^*(w^*), \quad \forall w^* \in V^*,$$

$$\langle v^*, v \rangle - J^*(v^*) \geq J^{**}(v) = J(v).$$

Thus, we again arrive at the conclusion that it can only be if $D_J(v^*, v) = 0$. 

S. Repin  
RICAM, Special Radon Semester, Linz, 2005.  
LECITURES ON A POSTERIORI ERROR CONTROL
Assume that $D_J(v^*, v) = 0$. Since

$$J^*(v^*) = \sup_{w \in V} \{\langle v^*, w \rangle - J(w)\},$$

we obtain

$$0 = J(v) + J^*(v^*) - \langle v^*, v \rangle \geq J(v) - J(w) - \langle v^*, v - w \rangle, \quad \forall w \in V.$$  

Rewrite this inequality in a more familiar form:

$$J(w) - J(v) \geq \langle v^*, w - v \rangle, \quad \forall w \in V,$$

which means that $J(v) + \langle v^*, v - w \rangle$ is an exact affine minorant of $J$ (at $v$) and, consequently, $v^* \in \partial J(v)$. The proof of the fact that $v^* \in \partial J^*(v^*)$ is quite analogous.
Properties of compound functionals

First, we note that, $D_G(y, y^*)$ is \textit{convex} with respect to $y$ and $y^*$, but, in general, $D_G(y, y^*)$ is a \textit{nonconvex} functional on $Y \times Y^*$. This fact is easily observed in the simplest case $Y = \mathbb{R}$ if set

$$G(y) = \frac{1}{\alpha} |y|^{\alpha} \quad \text{and} \quad G^*(y) = \frac{1}{\alpha^*} |y|^{\alpha^*}.$$

Only for $\alpha = 2$ we have a convex functional

$$D_G(y, y^*) = \frac{1}{2} |y|^2 + \frac{1}{2} |y^*|^2 - yy^* = \frac{1}{2} (y - y^*)^2.$$

For other $\alpha \in (1, +\infty)$ $D_G$ is nonconvex on $\mathbb{R} \times \mathbb{R}$. 
Example 1: \( D(\xi_1, \xi_2) = \frac{1}{3}|\xi_1|^3 + \frac{2}{3}|\xi_2|^{3/2} - \xi_1 \xi_2 \)

Compound functional on \( \mathbb{R} \times \mathbb{R} \) and its level lines
Example 2: $D(\xi_1, \xi_2) = \frac{5}{6} |\xi_1|^{6/5} + \frac{1}{6} |\xi_2|^6 - \xi_1 \xi_2$

**Compound functional on** $\mathbb{R} \times \mathbb{R}$ **and its level lines**
However, they have an important property, which is to some extent similar to convexity.

**Theorem**

For any $y_1, y_2 \in Y$ and $y_1^*, y_2^* \in Y^*$,

$$D_G\left(\frac{y_1 + y_2}{2}, \frac{y_1^* + y_2^*}{2}\right) \leq \frac{1}{4} \left( D_G(y_1, y_1^*) + D_G(y_1, y_2^*) + \right.$$  

$$+ D_G(y_2, y_1^*) + D_G(y_2, y_2^*) \right)$$
Proof

From the definition it follows that

\[
D_G \left( y, \frac{y_1^* + y_2^*}{2} \right) = G(y) + G^* \left( \frac{y_1^* + y_2^*}{2} \right) - \langle \frac{y_1^* + y_2^*}{2}, y \rangle \\
\leq \frac{1}{2} \left( D_G(y, y_1^*) + D_G(y, y_2^*) \right)
\]

and

\[
D_G \left( \frac{y_1 + y_2}{2}, y^* \right) = G \left( \frac{y_1 + y_2}{2} \right) + G^* (y^*) - \langle y^*, \frac{y_1 + y_2}{2} \rangle \\
\leq \frac{1}{2} \left( D_G(y_1, y^*) + D_G(y_2, y^*) \right).
\]

Therefore,

\[
D_G \left( \frac{y_1 + y_2}{2}, \frac{y_1^* + y_2^*}{2} \right) \leq \frac{1}{2} \left( D_G(y_1, \frac{y_1^* + y_2^*}{2}) + D_G(y_2, \frac{y_1^* + y_2^*}{2}) \right).
\]

and we arrive at the required estimate.
Important property

If $G$ and $G^*$ are Gateaux differentiable, then

$$\langle y^* - G'(y), G^*(y^*) - y \rangle \geq D_G(y, y^*).$$

Note, that from this relation we conclude that $D_J$ vanishes if the duality relations are satisfied.
Uniformly convex functionals

Let a proper l.s.c. functional $\mathcal{Y} : Y \to \overline{\mathbb{R}}$ be subject to the conditions

$$
\mathcal{Y}(y) \geq 0, \quad \forall y \in Y,
\quad \mathcal{Y}(y) = 0 \iff y = 0_Y.
$$

**Definition**

A convex functional $J : Y \to \overline{\mathbb{R}}$ is called uniformly convex in $\mathcal{B}(0_Y, \delta)$ if there exists a functional $\mathcal{Y}_\delta$ such that $\mathcal{Y}_\delta \not\equiv 0$ and for all $y_1, y_2 \in \mathcal{B}(0_Y, \delta)$ the following inequality holds:

$$
J \left( \frac{y_1 + y_2}{2} \right) + \mathcal{Y}_\delta(y_1 - y_2) \leq \frac{1}{2} (J(y_1) + J(y_2)).
$$

The functional $\mathcal{Y}_\delta$ enforces standard convexity inequality. For this reason, it is called a **forcing** functional.
It is clear that any uniformly convex functional is convex in $B(0_Y, \delta)$. Now we establish two important inequalities that hold for uniformly convex functionals.

**Theorem**

If $J : Y \rightarrow \overline{\mathbb{R}}$ is uniformly convex in $B(0_Y, \delta)$ and Gâteaux differentiable in $B(0_Y, \delta)$, then for any $y, z \in B(0_Y, \delta)$ the following relations hold:

$$J(z) \geq J(y) + \langle J'(y), z - y \rangle + 2\Upsilon_\delta(z - y)$$

and

$$\langle J'(z) - J'(y), z - y \rangle \geq 2\Upsilon_\delta(z - y) + 2\Upsilon_\delta(y - z).$$
Proof.

We have

\[ \mathcal{R}_\delta(z - y) \leq \frac{1}{2} J(z) + \frac{1}{2} J(y) - J\left(\frac{z+y}{2}\right) \].

Since \( J \) is convex and differentiable

\[ J\left(\frac{z+y}{2}\right) = J\left(y + \frac{z-y}{2}\right) \geq J(y) + \langle J'(y), \frac{z-y}{2} \rangle, \]

and, therefore,

\[ 2\mathcal{R}_\delta(z - y) \leq J(z) - J(y) - \langle J'(y), z - y \rangle \].

We can rewrite it replacing \( z \) by \( y \)

\[ 2\mathcal{R}_\delta(y - z) \leq J(y) - J(z) + \langle J'(z), z - y \rangle \]

and obtain the second inequality. \( \square \)
Deviations from the minimizer

**Theorem**

Let a functional $J$ be uniformly convex in $B(0_Y, \delta)$ and $y_m \in B(0_Y, \delta)$ be the minimizer of $J$.

\[
\gamma_{\delta}(z - y_m) \leq \frac{1}{2} (J(z) - J(y_m)), \quad \forall z \in B(0_Y, \delta).
\] (4.4)

**Proof.**

Since $J\left(\frac{y_m + z}{2}\right) \geq J(y_m)$, we obtain

\[
\gamma_{\delta}(z - y_m) \leq \frac{1}{2} J(y_m) + \frac{1}{2} J(z) - J \left(\frac{y_m + z}{2}\right) \leq \frac{1}{2} (J(z) - J(y_m)).
\]
Estimate (4.4) is the first step in deriving a posteriori error estimates of the functional type by means of the variational techniques. It shows that deviations from the minimizer (measured in terms of the functional $\mathcal{Y}_\delta$) are controlled by the difference of the functionals.
Corollary 1

Rewrite (4.3) in the form

\[ \gamma_\delta (z - y_m) + J \left( \frac{y_m + z}{2} \right) - J(y_m) \leq \frac{1}{2} (J(z) - J(y_m)) . \]

By virtue of (4.4), we have

\[ J \left( \frac{y_m + z}{2} \right) - J(y_m) \geq 2 \gamma_\delta \left( \frac{z - y_m}{2} \right) \]

and, therefore, we arrive at the strengthened estimate

\[ \gamma_\delta (z - y_m) + 2 \gamma_\delta \left( \frac{z - y_m}{2} \right) \leq \frac{1}{2} (J(z) - J(y_m)) . \quad (4.5) \]
Corollary 2

Assume that $J$ is twice differentiable in the vicinity of $y_m$ and satisfies the finite increment relation

$$J \left( \frac{y_m + z}{2} \right) = J(y_m) + \left\langle J'(y_m), \frac{z - y_m}{2} \right\rangle +$$

$$+ \frac{1}{2} \left\langle J'' \left( y_m + \xi \frac{z + y_m}{2} \right), \frac{z - y_m}{2}, \frac{z - y_m}{2} \right\rangle,$$

where $\xi \in (0, 1)$. Since $J'(y_m) = 0_{Y^*}$, we have another estimate:

$$\gamma_\delta(z - y_m) + \frac{1}{8} \left\langle J'' \left( (1 + \frac{\xi}{2})y_m + \frac{\xi}{2}z \right), z - y_m \right\rangle \leq$$

$$\leq \frac{1}{2} (J(z) - J(y_m)). \quad (4.6)$$
Example 1

Consider a self-adjoint operator $A \in \mathcal{L}(H, H)$ defined on a Hilbert space $H$ with scalar product $(.,.)$. Assume that it satisfies the condition

$$\alpha_1 \|y\|^2 \leq G(y) := (Ay, y) \leq \alpha_2 \|y\|^2, \quad \forall y \in H.$$

For $J(y) = G(y) + (\ell, y), \quad \ell \in H$ we have

$$\frac{1}{2}G(y) + \frac{1}{2}G(z) - G\left(\frac{y+z}{2}\right) =$$

$$= \frac{1}{4}(Ay, y) + \frac{1}{4}(Az, z) - \frac{1}{8}(A(y+z), y+z) =$$

$$= \frac{1}{8}(A(z - y), (z - y)),$$

the functional $G$ is uniformly convex in any ball with

$$\Upsilon(z - y) = \frac{1}{8}(A(z - y), (z - y)).$$

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
Thus, from (4.4) we have
\[
\frac{1}{8} \langle A(z - y_m), (z - y_m) \rangle \leq \frac{1}{2} (J(z) - J(y_m)), \quad \forall z
\]

However (4.6) gives a better estimate
\[
\frac{1}{2} \langle A(z - y_m), (z - y_m) \rangle \leq J(z) - J(y_m).
\] (4.7)

Note that for quadratic type functionals this estimate holds as equality. Indeed,
\[
J(z) - J(y_m) = \langle Ay_m + \ell, z - y_m \rangle + \frac{1}{2} \langle A(z - y_m), z - y_m \rangle.
\]

and the minimizer \( y_m \) satisfies the relation
\[
\langle Ay_m + \ell, y \rangle = 0, \quad \forall y \in Y.
\]

Therefore, (4.7) holds as equality.
Theorem

Let $J_1$ and $J_2$ be uniformly convex in $B(0_Y, \delta)$ with functionals $\Upsilon_{1\delta}$ and $\Upsilon_{2\delta}$, respectively. Then the functional

$$\mu_1 J_1 + \mu_2 J_2,$$

where $\mu_1, \mu_2 \geq 0$, is uniformly convex in $B(0_Y, \delta)$ with

$$\Upsilon_{\delta} = \mu_1 \Upsilon_{1\delta} + \mu_2 \Upsilon_{2\delta}.$$ 

Proof.

The proposition follows directly from definition of uniform convexity.
Example 2

Consider the functional

\[ J(y) = \frac{1}{2}(Ay, y) + (\ell, y) + \Psi(y), \]

where \( \Psi(y) \) is a convex and l.s.c. functional. Applying the above Theorem with \( \mu_1 = \mu_2 = 1 \),

\[ J_1(y) = \frac{1}{2}(Ay, y) + (\ell, y) \quad J_2(y) = \Psi(y), \]

we see that \( J \) is uniformly convex with functional \( \Upsilon \) defined in Example 1.
Theorem

Let $J_1$ and $J_2$ be uniformly convex in $\mathcal{B}(0_Y, \delta)$ with functionals $\Upsilon_{1\delta}$ and $\Upsilon_{2\delta}$, respectively. Then the functional

$$J(y) = \max\{J_1(y), J_2(y)\}$$

is uniformly convex in $\mathcal{B}(0_Y, \delta)$ with

$$\Upsilon_\delta = \min\{\Upsilon_{1\delta}, \Upsilon_{2\delta}\}.$$
**Proof.** We have

\[
\frac{1}{2} J(y) + \frac{1}{2} J(z) - J\left(\frac{y+z}{2}\right) = \frac{1}{2} \max\{J_1(y), J_2(y)\} + \\
+ \frac{1}{2} \max\{J_1(z), J_2(z)\} - \max\left\{J_1\left(\frac{y+z}{2}\right), J_2\left(\frac{y+z}{2}\right)\right\}.
\]

Assume that

\[
\max\left\{J_1\left(\frac{y+z}{2}\right), J_2\left(\frac{y+z}{2}\right)\right\} = J_1\left(\frac{y+z}{2}\right).
\]

Then

\[
\frac{1}{2} (J(y) + J(z)) - J\left(\frac{y+z}{2}\right) \geq \\
\geq \frac{1}{2} (J_1(y) + J_1(z)) - J_1\left(\frac{y+z}{2}\right) \geq \Upsilon_1 \delta (z - y).
\]
If we have an opposite situation, i.e.,

$$\max \{ J_1 \left( \frac{y+z}{2} \right), J_2 \left( \frac{y+z}{2} \right) \} = J_2 \left( \frac{y+z}{2} \right),$$

then

$$\frac{1}{2} J(y) + \frac{1}{2} J(z) - J \left( \frac{y+z}{2} \right) \geq \gamma_{2\delta}(z - y).$$

Thus, in both cases the lower bound is given by the functional

$$\min \{ \gamma_{1\delta}(z - y), \gamma_{2\delta}(z - y) \}.$$
Example 3. Power growth functionals

Let

\[ G(y) = \frac{1}{\alpha} \int_{\Omega} |y|^\alpha \, dx \quad F(v) = \int_{\Omega} fv \, dx, \]

where \( \alpha > 1 \). Then Problem \( \mathcal{P} \) is to minimize the functional

\[ J_\alpha(v) := \int_{\Omega} \left( \frac{1}{\alpha} |\nabla v|^\alpha + fv \right) \, dx \]

over the space \( \mathbf{V} = \{ v \in H^\alpha(\Omega) \mid v = 0 \text{ on } \partial\Omega \} \).
Problem $\mathcal{P}^*$ is to maximize the functional

$$I_{\alpha^*}^*(y^*) = -\frac{1}{\alpha^*} \int_{\Omega} |y^*|^{\alpha^*} \, dx$$

over the set

$$Q_f^* = \left\{ y^* \in Y^* := L^{\alpha^*}(\Omega, \mathbb{R}^n) \bigg| \int_{\Omega} y^* \cdot \nabla w dx = \int_{\Omega} f w dx \quad \forall w \in V \right\}.$$
For $\alpha \geq 2$ uniform convexity of $\mathbf{G}(\mathbf{y})$ follows from the first Clarkson’s inequality

$$
\int_{\Omega} \left| \frac{y_1+y_2}{2} \right|^\alpha \, dx + \int_{\Omega} \left| \frac{y_1-y_2}{2} \right|^\alpha \, dx \leq \frac{1}{2} \int_{\Omega} (|y_1|^\alpha + |y_2|^\alpha) \, dx,
$$

which is valid for all $y_1, y_2 \in \mathbf{Y}$.

See S. L. Sobolev. *Some Applications of Functional Analysis in Mathematical Physics.* Hence, we observe that in this case

$$
\gamma(z) = \frac{1}{\alpha} \|z\|_{\alpha, \Omega}^\alpha.
$$

and

$$
\frac{1}{\alpha 2^\alpha} \int_{\Omega} |\nabla(v - u)|^\alpha \, dx \leq \frac{1}{2} \left( J_\alpha(v) - l_\alpha^*(q^*) \right), \forall q^* \in Q_f^*.
$$
For $1 < \alpha \leq 2$, the functional $G$ is also uniformly convex. This fact follows from the second Clarkson's inequality

$$
\left( \int_\Omega \left( \frac{y_1 + y_2}{2} \right)^\alpha \, dx \right)^{\frac{1}{\alpha-1}} + \left( \int_\Omega \left( \frac{y_1 - y_2}{2} \right)^\alpha \, dx \right)^{\frac{1}{\alpha-1}} \leq \left( \frac{1}{2} \int_\Omega (|y_1|^\alpha + |y_2|^\alpha) \, dx \right)^{\frac{1}{\alpha-1}}.
$$

However, in this case, the functional $\Upsilon_\delta$ depends on the radius $\delta$ of a ball $\mathcal{B}(0_Y, \delta)$ that contains $y_1$ and $y_2$, so that the estimate holds with

$$
\Upsilon_\delta(z) = \delta^{\frac{\alpha-2}{\alpha-1}} \kappa \|z\|^{\frac{\alpha}{\alpha-1}}_{\alpha, \Omega},
$$

where $\kappa = \frac{1}{\kappa_0 + 1}$ and $\kappa_0$ is the integer part of $\frac{1}{\alpha-1}$. 
Lecture 5.
FUNCTIONAL A POSTERIORI ESTIMATES. GENERAL APPROACH.
Main goal of the lecture

We expose the general approach to deriving two-sided functional estimates of the deviations from exact solutions of linear elliptic type problems having the operator form

\[ \Lambda^* A \Lambda u + \ell = 0 \]

where \( \Lambda \) and \( A \) are linear bounded operators and \( A \) is symmetric and positive definite.
Lecture plan

- Two–sided a posteriori estimates for linear elliptic type problems;
- Properties: computability, consistency, reliability;
- Relationships with other error estimation methods;
Problem in the abstract form

Many problems can be presented in the following form: find $u \in V_0 + u_0$ such that

$$\langle A\Lambda u, \Lambda w \rangle + \langle \ell, w \rangle = 0 \quad \forall w \in V_0. \quad (5.1)$$

Here $V_0$ is a subspace of a reflexive Banach space $V$, e.g., $V = H^1$, $V_0 = H^1_0$. $\Lambda : V \to U$ is a bounded linear operator, e.g., $\Lambda = \nabla$. $U$ is a Hilbert space with scalar product $(\cdot, \cdot)$ and norm $\| \cdot \|$, e.g., $U = L^2$. $\ell \in V_0^*$, is a linear functional in the dual space, e.g., in $H^{-1}$. In general, we may assume that

$$\langle \ell, w \rangle = (f, w) + (g, \Lambda w).$$

$A \in L(U, U)$ is a self-adjoint operator.
Assumptions

We assume that

\[ V \text{ is compactly embedded in } U \]  \hspace{1cm} (5.2)

and the operators \( \Lambda \) and \( A \) satisfy the relations

\[ c_1 \|y\|^2 \leq (Ay, y) \leq c_2 \|y\|^2, \quad \forall y \in U, \]  \hspace{1cm} (5.3)

\[ \|\Lambda w\| \geq c_3 \|w\|_V, \quad \forall w \in V_0, \]  \hspace{1cm} (5.4)
For our analysis, it is convenient to introduce two more norms:

\[ \| y \| := (A y, y)^{1/2}, \quad \| y \|_* := (A^{-1} y, y)^{1/2}, \]

where \( A^{-1} \) is the operator inverse to \( A \). The respective spaces \( Y \) and \( Y^* \) contain elements of \( U \) equipped with the norms \( \| \cdot \| \) and \( \| \cdot \|_* \), respectively.

Problem (5.1) is equivalent to following problem.

**Problem \( \mathcal{P} \).** Find \( u \in V_0 + u_0 \) such that

\[ J(u) = \inf_{v \in V_0 + u_0} J(u) := \inf \mathcal{P}, \]

where

\[ J(v) = \frac{1}{2} \| \Lambda v \|^2 + \langle \ell, v \rangle. \]
On the set \((V_0 + u_0) \times Y^*\), we define the Lagrangian

\[
L(v, y) = (y, \Lambda v) - \frac{1}{2} \| y \|^2 + \langle \ell, v \rangle
\]

and the functional

\[
I^*(y) = \inf_{v \in V_0 + u_0} L(v, y) = \begin{cases} 
(y, \Lambda u_0) - \frac{1}{2} \| y \|^2 + \langle \ell, u_0 \rangle, & y \in Q^*_\ell, \\
-\infty, & y \notin Q^*_\ell,
\end{cases}
\]

where \(Q^*_\ell := \{ y \in Y^* \mid (y, \Lambda w) + \langle \ell, w \rangle = 0, \quad \forall w \in V_0 \}\).

Note that since

\[
(y, \Lambda(u_0 + w) + \langle \ell, (u_0 + w) \rangle = (y, \Lambda u_0) + \langle \ell, u_0 \rangle
\]

we see that \(I^*\) does not depend on the form of \(u_0\) inside \(\Omega\).
The problem dual to $\mathcal{P}$ is as follows. 

**Problem $\mathcal{P}^*$**. Find $p \in Q^*_\ell$ such that

$$I^*(p) = \sup_{y \in Q^*_\ell} I^*(y) := \sup_{y \in Q^*_\ell} \mathcal{P}^* \leq \inf \mathcal{P}. $$

The minimizer $u$ satisfies and the maximizer $p$ satisfies the stationarity conditions

$$(\Lambda u, \Lambda w) + \langle \ell, w \rangle = 0 \quad \forall w \in V_0,$$

$$(\Lambda u_0 - \mathcal{A}^{-1} p, y) = 0, \quad \forall y \in Q^*_0,$$

where $Q^*_0 := \{ y \in Y^* \mid (y, \Lambda w) = 0, \quad \forall w \in V_0 \}$. 

We see that $\mathcal{A} \Lambda u \in Q^*_\ell$. 

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
Take
\[ I^*(A\Lambda u) = (A\Lambda u, \Lambda u_0) - \frac{1}{2} \| A\Lambda u \|_*^2 + \langle \ell, u_0 \rangle \]

and set \( u_0 = u \). We obtain
\[ I^*(A\Lambda u) = (A\Lambda u, \Lambda u) - \frac{1}{2} \| A\Lambda u \|_*^2 + \langle \ell, u \rangle \leq \sup P^*. \]

Since \( \| A\Lambda u \|_*^2 = (A^{-1}A\Lambda u, A\Lambda u) = \| \Lambda u \|_2^2 \), we see that
\[ I^*(A\Lambda u) = J(u) = \inf P \]

Thus
\[ \sup P^* = \inf P \]
The relation $I^*(p) = J(u)$ means that

$$(p, \Lambda u) - \frac{1}{2} \| p \|^2_* + \langle \ell, u \rangle = \frac{1}{2} \| \Lambda u \|^2 + \langle \ell, u \rangle,$$

which is equivalent to the relation

$$D(\Lambda u, p) = \frac{1}{2} \| \Lambda u \|^2 + \frac{1}{2} \| p \|^2_* - (p, \Lambda u) = 0.$$

From the above we see that $\Lambda u$ and $p$ are joined by a certain relation:

$$p = \mathcal{A}\Lambda u$$

This is the so-called duality relation for the pair $(u, p)$. 
Let \( \mathbf{v} \in \mathbf{V}_0 + \mathbf{u}_0 \) and \( \mathbf{y} \in \mathbf{Y}^* \) be some approximations of \( \mathbf{u} \) and \( \mathbf{p} \), respectively. Our goal is to obtain two-sided estimates of the quantities \( \| \Lambda (\mathbf{v} - \mathbf{u}) \| \) and \( \| \mathbf{y} - \mathbf{p} \|_* \) that are norms of deviations from the exact solutions \( \mathbf{u} \) and \( \mathbf{p} \).

First, we establish the following basic result.

**Theorem**

*For any \( \mathbf{v} \in \mathbf{V}_0 + \mathbf{u}_0 \) and \( \mathbf{q} \in \mathbf{Q}_\ell^* \),

\[
\| \Lambda (\mathbf{v} - \mathbf{u}) \|^2 + \| \mathbf{q} - \mathbf{p} \|^2_* = 2 (J(\mathbf{v}) - I^*(\mathbf{q})),
\]

(5.5)

\[
\| \Lambda (\mathbf{v} - \mathbf{u}) \|^2 + \| \mathbf{q} - \mathbf{p} \|^2_* = 2 D(\Lambda \mathbf{v}, \mathbf{q}).
\]

(5.6)
Proof

By the stationarity relations, we have

\[
\frac{1}{2} \| \Lambda(v - u) \|^2 = J(v) - J(u) + \\
(\mathcal{A} \Lambda u, \Lambda(u - v)) + \langle \ell, u - v \rangle = \\
= J(v) - J(u).
\]

Analogously

\[
\frac{1}{2} \| q - p \|^2_* = I^*(p) - I^*(q) + (\Lambda u_0 - \mathcal{A}^{-1} p, p - q) = \\
= I^*(p) - I^*(q).
\]

Since \( J(u) = I^*(p) \), we sum two relations and obtain \((5.5)\). For \( q \in Q^*_\ell \) the difference \( J(v) - I^*(q) \) is equal to \( D(\Lambda v, q) \), so that \((5.6)\) follows from \((5.5)\).
The estimates (5.5) and (5.6) are valid only for \( q \in Q_\ell^* \), which poses some difficulties. Below it is shown how we can overcome this drawback. First, we establish one subsidiary result.

**Theorem**

Let \( q \in Q_\ell^* \), \( v \in V_0 + u_0 \), \( \beta \in \mathbb{R}_+ \), and \( y \in Y^* \). Then

\[
J(v) - I^*(q) \leq (1 + \beta)D(\Lambda v, y) + \frac{1 + \beta}{2\beta} \| q - y \|_*^2. \tag{5.7}
\]

Note that

\[
D(\Lambda v, y) = \frac{1}{2} (\mathcal{A} \Lambda v, \Lambda v) + \frac{1}{2} (\mathcal{A}^{-1} p, p) - (y, \Lambda u) = \]

\[
= (\mathcal{A} \Lambda v - y, \Lambda v - \mathcal{A}^{-1} y) = \]

\[
= (\mathcal{A}(\Lambda v - \mathcal{A}^{-1} y, \Lambda v - \mathcal{A}^{-1} y) = \| \Lambda v - \mathcal{A}^{-1} y \|. \]

S. Repin  
*RICAM, Special Radon Semester, Linz, 2005.*

**LECTURES ON A POSTERIORI ERROR CONTROL**
Proof

For any $y \in Y^*$, we have

$$J(v) - I^*(q) = \frac{1}{2} \left( \| \Lambda v \|^2 + \| y \|_{\ast}^2 \right) +$$

$$+ \frac{1}{2} \left( \| q \|_{\ast}^2 - \| y \|_{\ast}^2 \right) - (\Lambda u_0, q) + \langle \ell, v - u_0 \rangle.$$

Since $\langle \ell, v - u_0 \rangle = (q, \Lambda (u_0 - v))$, we find that

$$J(v) - I^*(q) = \frac{1}{2} \left( \| \Lambda v \|^2 + \| y \|_{\ast}^2 \right) + \frac{1}{2} \left( \| q \|_{\ast}^2 - \| y \|_{\ast}^2 \right) - (q, \Lambda v) =$$

$$= D(\Lambda v, y) + \left( y - q, \Lambda v - A^{-1}y \right) + \frac{1}{2} \| q - y \|_{\ast}^2.$$

This relation yields (5.7) if we use the Young’s inequality

$$2 \left( y - q, \Lambda v - A^{-1}y \right) \leq \beta \| \Lambda v - A^{-1}y \|^2 + \beta^{-1} \| y - q \|_{\ast}^2.$$
Another form of the estimate

Introduce the quantity

\[ d_\ell^2(y) := \inf_{q \in Q_\ell^*} \| q - y \|_*^2, \]

which is the distance to \( Q_\ell^* \). Then, (5.7) imply the estimate

\[ \frac{1}{2} \| \Lambda(v - u) \|_2^2 \leq (1 + \beta) D(\Lambda v, y) + \left(1 + \frac{1}{\beta}\right) \frac{1}{2} d_\ell^2(y) \]

where \( v \in V_0 + u_0 \) and \( y \in Y^* \). We rewrite this estimate as

\[ \frac{1}{2} \| \Lambda(v - u) \|_2^2 \leq \mathcal{M}(v, \beta), \quad \forall v \in V_0 + u_0, \quad \beta \in \mathbb{R}_+, \]

where

\[ \mathcal{M}(v, \beta) := \inf_{y \in Y^*} \left\{ (1 + \beta) D(\Lambda v, y) + \left(1 + \frac{1}{\beta}\right) \frac{1}{2} d_\ell^2(y) \right\}. \]
Above estimate is sharp for any $\beta$!

Theorem

For any $\beta \in \mathbb{R}_+$,

\[
\frac{1}{2} \| \Lambda(v - u) \|^2 = \mathcal{M}(v, \beta).
\]

Proof. Set $y = \lambda p + (1 - \lambda)\mathcal{A}\Lambda v$. Then

\[
D(\Lambda v, y) = \frac{1}{2} \lambda^2 \| \Lambda(v - u) \|^2.
\]

Since

\[
d_\ell^2(y) \leq \| p - y \|_*^2 = (1 - \lambda)^2 \| p - \mathcal{A}\Lambda v \|_*^2 = (1 - \lambda)^2 \| \mathcal{A}\Lambda(u - v) \|_*^2 =
\]

\[
= (1 - \lambda)^2 \| \Lambda(u - v) \|^2,
\]

we obtain

\[
\mathcal{M}(v, \beta) \leq \frac{1}{2} \left( (1 + \beta)\lambda^2 + \left( 1 + \frac{1}{\beta} \right) (1 - \lambda)^2 \right) \| \Lambda(v - u) \|^2.
\]
The right-hand side attains its minimal value at $\lambda = 1/(1 + \beta)$, which leads to the estimate

$$\frac{1}{2} \| \Lambda(v - u) \|^2 \geq \mathcal{M}(v, \beta), \quad \forall v \in V_0 + u_0, \quad \beta \in \mathbb{R}_+.$$ 

Recalling that the inverse inequality has already been established, we arrive at the required conclusion

Now, we proceed to finding computable upper bounds for the quantity $\mathbf{d}_{\ell}$. The first step is given by

**Theorem**

$$\frac{1}{2} \mathbf{d}_{\ell}^2(y) = \sup_{w \in V_0} \left\{ -\frac{1}{2} \| \Lambda w \|^2 - \langle \ell, w \rangle - (y, \Lambda w) \right\}.$$
Proof

This assertion comes from that \( \inf \mathcal{P} = \sup \mathcal{P}^* \). Indeed,

\[
\frac{1}{2} d^2_\ell(y) = - \sup_{\eta^* \in Q^*_\ell} \left\{ - \frac{1}{2} \| y - \eta^* \|_2^2 \right\} = - \sup_{\eta^* \in Q^*_\ell - y} \left\{ - \frac{1}{2} \| \eta^* \|_2^2 \right\},
\]

where \( Q^*_\ell - y := \{ \eta^* \in Y^* | \eta^* = \varepsilon^* - y, \varepsilon^* \in Q^*_\ell \} \).

In other words, \( \eta^* \in Q^*_\ell - y \) if

\[
(\eta^*, \Lambda w) = -\langle \ell, w \rangle - (y, \Lambda w), \quad \forall w \in V_0.
\]

The right-hand side of this relation is a linear continuous functional. We denote it by \( \ell_y \) and rewrite the relation as follows:

\[
(\eta^*, \Lambda w) + \langle \ell_y, w \rangle = 0 \quad \forall w \in V_0.
\]
Then, $Q_\ell^* - y = Q_{\ell y}^*$ and

$$\frac{1}{2} d_\ell^2(y) = - \sup_{\eta^* \in Q_{\ell y}^*} \left\{ \frac{1}{2} \| \eta^* \|^2 \right\}.$$

This maximization problem is a form of Problem $P^*$ if set $u_0 = 0$ and $\ell = \ell_y$. Since $\sup P^* = \inf P$, we have

$$\frac{1}{2} d_\ell^2(y) = - \inf_{w \in V_0} \left\{ \frac{1}{2} \| \Lambda w \|^2 + \langle \ell_y, w \rangle \right\} =$$

$$= - \inf_{w \in V_0} \left\{ \frac{1}{2} \| \Lambda w \|^2 + \langle \ell, w \rangle + (y, \Lambda w) \right\} =$$

$$= \sup_{w \in V_0} \left\{ -\frac{1}{2} \| \Lambda w \|^2 - \langle \ell, w \rangle - (y, \Lambda w) \right\}.$$

□
Corollary

We arrive at the conclusion that the majorant $\mathcal{M}(v, \beta)$ has a minimax form

$$
\mathcal{M}(v, \beta) = \inf_{y \in Y^*} \sup_{w \in V_0} \left\{ (1+\beta)D(\Lambda v, y) + \frac{1+\beta}{\beta} \left( -(y, \Lambda w) - J(w) \right) \right\}. \quad (5.8)
$$

Further, we use (5.8) for deriving upper and lower error bounds.
Upper estimates of $\| v - u \|$

In the relation

$$\mathcal{M}(v, \beta) \leq (1 + \beta)D(\Lambda v, y) +$$

$$+ \left(1 + \frac{1}{\beta}\right) \sup_{w \in V_0} \left\{ -\frac{1}{2} \| \Lambda w \|^2 - \langle \ell, w \rangle - (y, \Lambda w) \right\},$$

we will estimate the value of supremum. Let $\Lambda^*$ be the operator conjugate to $\Lambda$, i.e.,

$$(y, \Lambda w) = \langle \Lambda^* y, w \rangle, \quad \forall w \in V_0.$$

Then

$$\langle \ell, w \rangle + \langle y, \Lambda w \rangle = \langle \ell + \Lambda^* y, w \rangle \leq \| \ell + \Lambda^* y \| \| \Lambda w \|.$$
Here

\[ \left\| \ell + \Lambda^* y \right\| := \sup_{w \in V_0} \frac{\langle \ell + \Lambda^* y, w \rangle}{\|\Lambda w\|} < +\infty. \]

To prove that the value of the negative norm is finite we estimate the numerator as follows:

\[
\langle \ell + \Lambda^* y, w \rangle \leq \|\ell\|_{V_0^*} \|w\|_V + \|y\| \|\Lambda w\| \leq \left( c_3^{-1} \|\ell\|_{V_0^*} + \|y\| \right) \|\Lambda w\| \leq c_1^{-1/2} \left( c_3^{-1} \|\ell\|_{V_0^*} + \|y\| \right) \|\Lambda w\|.
\]

We see that

\[
\sup_{w \in V_0} \left\{ -\frac{1}{2} \|\Lambda w\|^2 - \langle \ell, w \rangle - (y, \Lambda w) \right\} \leq \sup_{w \in V_0} \left\{ -\frac{1}{2} \|\Lambda w\|^2 + \left\| \ell + \Lambda^* y \right\| \|\Lambda w\| \right\} \leq \sup_{t > 0} \left\{ -\frac{1}{2} t^2 + \left\| \ell + \Lambda^* y \right\| t \right\} = \frac{1}{2} \left\| \ell + \Lambda^* y \right\|^2.
\]

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
Thus, we obtain
\[
\frac{1}{2} \left\| \Lambda(v - u) \right\|^2 \leq (1 + \beta)D(\Lambda v, y) + \frac{1 + \beta}{2\beta} \left\| \ell + \Lambda^* y \right\|^2. \tag{5.9}
\]

This estimate contains the norm $\| \cdot \|$ defined via a sup-relation. We replace it by the norm in a Hilbert space $\mathbf{U}$ provided that $\ell$ belongs to a narrower set. Assume that
\[
\ell \in \mathbf{U} \subset \mathbf{V}_0^*,
\]
\[
y \in \mathbf{Q}^* := \{ z^* \in \mathbf{Y}^* | \Lambda^* z^* \in \mathbf{U} \}.
\]

Note that $\mathbf{Q}^*$ can be endowed with the norm
\[
\| y \|^2_{\mathbf{Q}^*} := \| y \|^2_* + \| \Lambda^* y^* \|^2_{\mathbf{U}}.
\]

If $\ell \in \mathbf{U}$, then $\mathbf{Q}^*$ contains the exact solution $p$ of the dual problem! This fact is important for the proof of the sharpness of the Majorant.
Majorant of the deviation

Then

$$\langle \ell + \Lambda^* y, w \rangle = (\ell + \Lambda^* y, w) \quad w \in V_0.$$ 

$$\| \ell + \Lambda^* y \| = \sup_{w \in V_0} \frac{\langle \ell + \Lambda^* y, w \rangle}{\| \Lambda w \|} \leq \sup_{w \in V_0} \frac{\| \ell + \Lambda^* y \| \| w \|}{\| \Lambda w \|} \leq \| \ell + \Lambda^* y \| c_1^{-1} \sup_{w \in V_0} \frac{\| w \|}{\| \Lambda w \|} \leq c_1^{-1} c_3^{-1} \| \ell + \Lambda^* y \|.$$ 

Here $c_1$ and $c_3$ are the constants in (5.3) and (5.4). Denote $c^2 = c_1^{-2} c_3^{-2}$. Now, the Majorant is represented in the form

$$\frac{1}{2} \| \Lambda (v - u) \|^2 \leq M_\oplus (v, \beta, y) :=$$

$$= (1 + \beta) D(\Lambda v, y) + \frac{1 + \beta}{2 \beta} c^2 \| \ell + \Lambda^* y \|^2. \quad (5.10)$$
Deviation Majorant for the problem $\Lambda^* A \Lambda u + \ell = 0$

\[
(\Lambda\Lambda(v - u), \Lambda(v - u)) \leq \\
\leq (1 + \beta) \left( (\Lambda\Lambda v, \Lambda v) + (A^{-1}y, y) - 2(y, \Lambda v) \right) + \\
+ \frac{1 + \beta}{\beta} \c^2 \| \ell + \Lambda^* y \|^2.
\]

In the above, $v \in V_0 + u_0$, $\beta > 0$, $y \in U$. 

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
Theorem

For any $v \in V_0 + u_0$,

$$\frac{1}{2} \| \Lambda(u - v) \|_2^2 = \inf_{\beta > 0, y \in Q^*} M_\oplus(v, \beta, y).$$

If $\ell \in U$, then $p \in Q^*$ and, therefore,

$$\inf_{y \in Q^*} M_\oplus(v, \beta, y) \leq M_\oplus(v, \epsilon, p) = (1 + \epsilon) \frac{1}{2} \| \Lambda(u - v) \|_2^2,$$

where $\epsilon > 0$ may be taken arbitrarily small.

Hence, the majorant $M_\oplus$ is reliable and exact.
Lower estimates

Recall the minimax form of the Majorant

$$\mathcal{M}(v, \beta) = \inf_{y \in Y^*} \sup_{w \in V_0} \left\{ (1 + \beta)D(\Lambda v, y) + \frac{1 + \beta}{\beta} \left( -(y, \Lambda w) - J(w) \right) \right\}.$$ 

Since $\sup \inf \leq \inf \sup$, we have

$$\mathcal{M}(v, \beta) \geq \sup_{w \in V_0} \inf_{y \in Y^*} \left\{ (1 + \beta)D(\Lambda v, y) - \left( 1 + \frac{1}{\beta} \right) \left( \frac{1}{2} \| \Lambda w \|^2 + \langle \ell, w \rangle + (y, \Lambda w) \right) \right\}.$$
Thus, for any $w \in V_0$

$$\mathcal{M}(v, \beta) \geq \inf_{y \in Y^*} \left\{ (1 + \beta) \left( \frac{1}{2} \| y \|_*^2 - (y, \Lambda v) \right) - \left( 1 + \frac{1}{\beta} \right) (y, \Lambda w) \right\} +$$

$$+ (1 + \beta) \frac{1}{2} \| \Lambda v \|_*^2 - \left( 1 + \frac{1}{\beta} \right) \left( \frac{1}{2} \| \Lambda w \|_*^2 \right),$$

Evidently, this estimate is also valid for the function $\beta w$, which yields

$$\mathcal{M}(v, \beta) \geq (1 + \beta) \inf_{y \in Y^*} \left\{ \frac{1}{2} \| y \|_*^2 - (y, \Lambda (v + w)) \right\} +$$

$$+ (1 + \beta) \left( \frac{1}{2} \| \Lambda v \|_*^2 - \frac{\beta}{2} \| \Lambda w \|_*^2 - \langle \ell, w \rangle \right).$$
Note that

\[
\inf_{y \in Y^*} \left\{ \frac{1}{2} \| y \|^2 - (y, \Lambda(v+w)) \right\} \geq
\]

\[
\geq \inf_{y \in Y^*} \left\{ \frac{1}{2} \| y \|^2 - \| y \| \| \Lambda(v+w) \| \right\} = -\frac{1}{2} \| \Lambda(v+w) \|^2.
\]

Thus, we obtain

\[
\mathcal{M}(v, \beta) \geq (1 + \beta) \left\{ -\frac{1}{2} \| \Lambda(v+w) \|^2 +
\right.
\]

\[
+ \frac{1}{2} \| \Lambda v \|^2 - \frac{\beta}{2} \| \Lambda w \|^2 - \langle \ell, w \rangle \right\} =
\]

\[
= (1 + \beta) \left\{ -\langle \mathcal{A} \Lambda v, \Lambda w \rangle - \frac{1 + \beta}{2} \| \Lambda w \|^2 - \langle \ell, w \rangle \right\}.
\]
In

$$(1 + \beta)\left\{ -(A\Lambda v, \Lambda w) - \frac{1 + \beta}{2} \| \Lambda w \|^2 - \langle \ell, w \rangle \right\}.$$}

$w$ is an arbitrary function in $V_0$. We may replace

$$w \quad \text{by} \quad \frac{w}{1 + \beta}.$$}

Such a replacement leads to the Minorant $M_{\ominus}(v, w)$ that gives a lower bound of the deviation from exact solution:

For any $w \in V_0$,

$$\frac{1}{2} \| \Lambda(v - u) \|^2 \geq -\frac{1}{2} \| \Lambda w \|^2 -(A\Lambda v, \Lambda w) - \langle \ell, w \rangle \quad (5.11)$$
Minorant is sharp

It is easy to see that

$$\sup_{w \in V_0} M_{\Theta}(v, w) = \frac{1}{2} \parallel \Lambda(v - u) \parallel^2.$$ 

Indeed, take $w = u - v$.

$$M_{\Theta}(v, u - v) = -\frac{1}{2} \parallel \Lambda(u - v) \parallel^2 - (\mathcal{A} \Lambda v, \Lambda u - v) - \langle \ell, u - v \rangle.$$

Represent the last two terms as follows:

$$-(\mathcal{A} \Lambda v, \Lambda(u - v)) - \langle \ell, u - v \rangle =$$

$$= -(\mathcal{A} \Lambda v, \Lambda(u - v)) + (\mathcal{A} \Lambda u, \Lambda(u - v)) =$$

$$= (\mathcal{A} \Lambda(u - v), \Lambda(u - v)) = \parallel \Lambda(u - v) \parallel^2$$

so that this choice of $w$ gives the true error.
Remark.

We outline that for the exact solution $M_\ominus = M_\oplus = 0$! Indeed, assume that $v$ coincides with $u$. In this case,

$$M_\ominus(u, w) = -\frac{1}{2} \| \Lambda w \|^2 - (A\Lambda u, \Lambda w) - \langle \ell, w \rangle = -\frac{1}{2} \| \Lambda w \|^2$$

and, therefore,

$$\sup_{w \in V_0} M_\ominus(u, w) = 0.$$  

The same is true for the majorant. Indeed, set $\hat{y} = A\Lambda u$. Then,

$$M_\oplus(u, \beta, \hat{y}) = (1 + \beta)D(\Lambda u, \hat{y}) + \frac{1 + \beta}{2\beta} c^2 \| \ell + \Lambda^* A\Lambda u \|^2 = 0.$$  

Thus,

$$\inf_{y \in Y^*} M_\oplus(u, \beta, y) = 0.$$
In many cases, error estimates in terms of the dual variable (that may represent "flux" or "stress") is as important as the error control of the primal variable.

Error estimates for the dual variable in the dual energy norm \( \| \cdot \|_* \) can be obtained by the arguments similar to those used above. Let \( y \in Y^* \) be an approximation of \( p \). For any \( q \in Q_\ell^* \), we obtain (from the triangle inequality and Young inequalities with \( \gamma > 0 \))

\[
\| y - p \|_*^2 \leq (1 + \gamma) \| y - q \|_*^2 + \left( 1 + \frac{1}{\gamma} \right) \| q - p \|_*^2.
\]

Recall that (see (5.5)) \( \| q - p \|_*^2 \leq 2 (J(v) - I^*(q)) \).
Therefore,

\[ \| y - p \|_2^2 \leq (1 + \gamma) \| y - q \|_2^2 + 2 \left(1 + \frac{1}{\gamma}\right) (J(u) - I^*(q)) \leq \]

\[ \leq (1 + \gamma) \| y - q \|_2^2 + 2 \left(1 + \frac{1}{\gamma}\right) (J(v) - I^*(q)) = \]

Recall that

\[ J(v) - I^*(q) \leq (1 + \beta)D(\Lambda v, y) + \left(1 + \frac{1}{\beta}\right) \frac{1}{2} \ell^2(y) \]

so that the right-hand side is estimated by

\[ (1 + \gamma) \left(1 + \frac{1}{\gamma} + \frac{1}{\beta \gamma}\right) \ell^2 + 2(1 + \beta) \left(1 + \frac{1}{\gamma}\right) D(\Lambda v, y). \]
Therefore,

\[
\frac{1}{2} \| y - p \|_2^2 \leq (1 + \gamma) \left( 1 + \frac{1}{\gamma} + \frac{1}{\beta \gamma} \right) I \ell + \Lambda^* y I^2 + (1 + \beta) \left( 1 + \frac{1}{\gamma} \right) D(\Lambda v, y). \tag{5.12}
\]

Rewrite this estimate as follows:

\[
\frac{1}{2} \| y - p \|_2^2 \leq M^*_\oplus (y, v, \beta, \gamma),
\]

where \( M^*_\oplus \) denotes the right-hand side of (5.12). This estimate holds for any \( y \in Y^* \), positive parameters \( \beta, \gamma \), and any \( v \in V_0 + u_0 \). Here \( v \) is a ”free” function in \( V_0 + u_0 \). This ”freedom” can be used to make the estimate sharper.
Computability of two–sided estimates

By **computability** we mean that upper and lower estimates can be computed with any a priori given accuracy by solving finite-dimensional problems. In the case considered, they are certain problems for quadratic type integral functionals whose minimization (maximization) is performed by well-known methods.
Let \( \{Y_i^*\}_{i=1}^{\infty} \) and \( \{V_{0i}\}_{i=1}^{\infty} \) be two sequences of finite-dimensional subspaces that are dense in \( Q^* \) and \( V_0 \), respectively, i.e., for any given \( \varepsilon > 0 \) and arbitrary elements \( y \in Y^* \) and \( w \in V_0 \), one can find a natural number \( k_\varepsilon \) such that

\[
\inf_{\tilde{w} \in V_{0i}} \| \tilde{w} - w \|_V \leq \varepsilon, \\
\inf_{\tilde{y} \in Y_i^*} \| \tilde{y} - y \|_{Q^*} \leq \varepsilon, \\
\forall i \geq k_\varepsilon.
\]

Let us show that sequences of two-sided bounds converging to the actual error can be evaluated by minimizing the Majorant on \( \{Y_i^*\} \) and maximizing the Minorant on \( \{V_{0i}\} \).
Take a small $\varepsilon > 0$. Then there exists a number $k$ and elements $w_k \in V_{0k}$ and $p_k \in Y_{0k}^*$ satisfying the conditions

$$\|w_k - (u - v)\|_V \leq \varepsilon, \quad \|p_k - p\|_{Q^*} \leq \varepsilon.$$ 

Define two quantities defined by solving finite–dimensional problems, namely

$$M_k^\oplus = \inf_{y_k \in Y_{0k}^*} M_\oplus(v, \beta, y_k), \quad M_k^\odot = \sup_{w_k \in V_{0k}} M_\odot(v, w_k).$$

By the definition

$$M_\odot(v, w_k) \leq M_k^\odot \leq \frac{1}{2} \|u - v\|^2 \leq M_k^\oplus \leq M_\oplus(v, \beta, p_k).$$
The quantities $M^k_\oplus$ and $M^k_\ominus$ are computable (they require solving finite dimensional problems for quadratic type functionals). We will that

$$M^k_\oplus \rightarrow \frac{1}{2} \| \Lambda (v - u) \|^2,$$

$$M^k_\ominus \rightarrow \frac{1}{2} \| \Lambda (v - u) \|^2$$

as the dimensionality $k$ tends to $+\infty$. 
Consider the upper estimates.

\[ M_\oplus(v, \beta, p_k) = (1 + \beta)D(\Lambda v, p_k) + \frac{1 + \beta}{2\beta}c^2\| \ell + \Lambda^* p_k \|^2 = \]

\[ = (1 + \beta)D(\Lambda v, p_k) + \frac{1 + \beta}{2\beta}c^2\| \Lambda^* (p_k - p) \|^2. \]

Here

\[ D(\Lambda v, p_k) = \frac{1}{2}(\Lambda v - A^{-1}p_k, A\Lambda v - p_k) = \]

\[ = \frac{1}{2} \left( \Lambda(v - u) - A^{-1}(p_k - p), A\Lambda(v - u) - (p_k - p) \right) = \]

\[ = \frac{1}{2} \| \Lambda(v - u) \|^2 + \| p_k - p \|^2_\star - (\Lambda(v - u), p_k - p). \]

From the latter estimate we see that

\[ D(\Lambda v, p_k) \leq \frac{1}{2} \| \Lambda(v - u) \|^2 + \varepsilon \| \Lambda(v - u) \| + \frac{1}{2}\varepsilon^2. \]
Since
\[ \| \Lambda^*(p_k - p) \|_{Q^*} \leq \varepsilon, \]
we find that
\[
M^k_\oplus \leq M_\oplus(v, \varepsilon, p_k) = \\
= (1 + \varepsilon) \left( \frac{1}{2} \| \Lambda(v - u) \|^2 + \varepsilon \| \Lambda(v - u) \| + \frac{1}{2} \varepsilon^2 \right) + \frac{1 + \varepsilon}{2\varepsilon} c^2 \varepsilon^2 = \\
= \frac{1}{2} \| \Lambda(v - u) \|^2 + c_4 \varepsilon + o(\varepsilon^2).
\]
where \( c_4 = \frac{1}{2} \left( c + 2 \| \Lambda(v - u) \| + \| \Lambda(v - u) \|^2 \right) \). Thus, we conclude that
\[
M^k_\oplus \rightarrow \frac{1}{2} \| \Lambda(v - u) \|^2 \quad \text{as} \quad k \rightarrow \infty.
\]
It is worth noting that the constant $c_4$ in the convergence term with $\varepsilon$ depends on the norm of $(v-u)$, so that we can await that for a good approximation convergence of the upper bounds to the exact value of the error is faster than in the case where $\|v-u\|$ is considerable. This phenomenon was observed in many numerical experiments. In general, finding an upper bound for a precise approximation takes less CPU time than for a coarse one.
Consider the lower estimates.

\[ M_{\ominus}(v, w_k) = -\frac{1}{2} \| \Lambda w_k \|^2 - (\mathcal{A} \Lambda v, \Lambda w_k) - \langle \ell, w_k \rangle = \]
\[ = -\frac{1}{2} \| \Lambda w_k \|^2 + (\mathcal{A} \Lambda (u - v), \Lambda w_k) = \]
\[ = \frac{1}{2} \| \Lambda (u - v) \|^2 - \frac{1}{2} \| \Lambda (w_k - (u - v)) \|^2 \geq \]
\[ \geq \frac{1}{2} \| \Lambda (u - v) \|^2 - \frac{1}{2} c_2 \| \Lambda (w_k - (u - v)) \|^2. \]

This implies the estimate

\[ \frac{1}{2} \| \Lambda (u - v) \|^2 \geq M_k^\ominus \geq \frac{1}{2} \| \Lambda (u - v) \|^2 - c_5 \varepsilon^2, \]

where \( c_5 > 0 \) depends on the norm of \( \Lambda \). Thus,

\[ M_k^\ominus \rightarrow \frac{1}{2} \| \Lambda (u - v) \|^2 \quad \text{as} \ k \rightarrow \infty. \]
Having $M^k_\oplus$ and $M^k_\ominus$, one can define the number

$$\eta_k := \frac{M^k_\oplus}{M^k_\ominus} \geq 1,$$  \hspace{1cm} (5.14)

which gives an idea of the quality of the error estimation. From the above it follows that

$$\eta_k \rightarrow 1, \quad \text{as } k \rightarrow +\infty.$$
Relationships with other methods

$\mathbf{M}_\oplus(v, \beta, y)$ involves an arbitrary function $y$. We are aimed to show that some special choices of it lead to known error estimates. We assume that $\langle l, w \rangle = (g, w)$, where $g \in U$, so that $p \in Q^* \subset Q_i^*$ and

$$Q_\ell^* := \{y \in Q^* \mid (\Lambda^* y + g, w) = 0, \ \forall w \in V_0\}.$$  

First, we select $y$ as follows

$$y_1^* = A\Lambda v. \quad (5.15)$$

Other variants arise if we set

$$y = \Pi y_1^*, \quad (5.16)$$

where $\Pi$ is a certain continuous mapping.
Residual based estimate

If $\Pi$ is the identity mapping of $Y^*$, i.e., $y = y_0^*$, then

$$D(\Lambda v, y_0^*) = 0.$$ 

Use the majorant in the form (5.9):

$$\frac{1}{2} \| \Lambda (v - u) \|^2 \leq (1 + \beta) D(\Lambda v, y) + \frac{1 + \beta}{2\beta} \| \ell + \Lambda^* y \|^2.$$ 

Now, it contains only the second term, which after the minimization with respect to $\beta$ gives

$$\| \Lambda (v - u) \|^2 \leq \| \ell + \Lambda^* A \Lambda v \| =$$

$$\sup_{w \in V_0} \frac{(g, w) + (A \Lambda v, \Lambda w)}{\| \Lambda w \|}. \quad (5.17)$$

If $v$ is obtained by FEM and $v = u_h \in V_h := V_{0h} + u_0$, (5.17) is estimated by using Galerkin orthogonality.
If in the functional a posteriori error estimate is applied to a FEM solution $u_h$ then we may select the variable $y$ in the simplest way as $y = \Lambda u_h$. Then, if $u_h$ is a Galerkin approximation, we can use this fact and obtain at an upper bound given by the residual type a posteriori error estimate that involve integral terms associated with finite elements and interelement jumps.
Estimates using post–processing of the dual variable

In $M_\oplus(v, \beta, y)$ the best choice is $y = p \in Q^*$. Therefore, if $y_0^* \notin Q^*$ then its mapping $Q^*$ could be a better approximation of $p$. Let us denote such a mapping by $\Pi_1$. We obtain

$$y_1^* = \Pi_1 y_0^* \in Q^*$$

and the quantity $M_\oplus(v, \beta, y_1^*)$, which leads to the error majorant

$$M^{(1)}_\oplus(v) = \inf_{\beta \in \mathbb{R}^+} \left\{ (1 + \beta)D(\Lambda v, \Pi_1(\mathcal{A}\Lambda v)) + \frac{1 + \beta}{2\beta} c^2 \| \ell + \Lambda^* \Pi_1(\mathcal{A}\Lambda v) \|^2 \right\}.$$
Particular case

In the simplest case associated with the problem

$$\Delta u + f = 0, \quad u = u_0 \quad \text{on } \partial \Omega$$

we have

$$M_{\ominus}^{(1)}(u_h) =$$

$$\inf_{\beta \in \mathbb{R}_+} \left\{ (1 + \beta) \| \nabla u_h - \Pi_1(\nabla u_h) \|^2 + \frac{(1 + \beta) C_{\Omega}^2}{2\beta} \| f + \text{div} \Pi_1(\nabla u_h) \|^2 \right\}.$$

If \( \Pi_1 \) is a gradient averaging operator, then the first term in the right-hand side is the difference between the original and averaged gradient, i.e. it coincides with a gradient averaging indicator. However, as we have seen in previous lectures, such an indicator cannot provide a reliable upper bound of the error. The second term in the right-hand side shows what is necessary to add in order to provide the reliability.
Diagram that shows connections with other methods

Problem $P$ (primal)

SPACE $V$

Problem $P^*$ (dual)

SPACE $V^*$
Estimates based on the "equilibration" of the dual variable

Let $\Pi_2$ maps $Y^*$ to the set $Q^*_\ell$. Define

$$y_2^* = \Pi_2 y_0^* \in Q^*_\ell.$$ (5.20)

Then,

$$\Lambda^* y_2^* + \ell = 0,$$

so that the Majorant has only the first term:

$$M_\oplus^{(2)}(v) = D(\Lambda v, y_2^*).$$

$\Pi_2$ is natural to call an equilibration operator. In general, it is rather difficult to construct an "exact mapping" $\Pi_2$ to $Q^*_\ell$. One may use an operator $\tilde{\Pi}_2$, which provides an approximate "equilibration". In this case, the second term of the Majorant does not vanish and should be taken into account.
A priori projection type error estimates

As an exercise, we now will derive classical a priori projection type error estimates from a functional a posteriori estimate. Let \( u_h \in V_h \) be a Galerkin approximation of \( u \). We have

\[
\| \Lambda(u - u_h) \| \leq 2(1 + \beta)D(\Lambda u_h, y) + \left( 1 + \frac{1}{\beta} \right) \| \Lambda^* y + \ell \|^2
\]

Set here \( y = \mathcal{A} \Lambda v_h \), where \( v_h \) is an arbitrary element of \( V_h \). Then,

\[
\| \Lambda^* y + \ell \| = \sup_{w \in V_0} \frac{(y - p, \Lambda w)}{\| \Lambda w \|} = \sup_{w \in V_0} \frac{(\mathcal{A} \Lambda (v_h - u), \Lambda w)}{\| \Lambda w \|} \leq \| \Lambda (v_h - u) \|.
\]
It is easy to see that
\[
D(\Lambda u_h, A\Lambda v_h) = J(v_h) - J(u_h).
\]
Indeed,
\[
D(\Lambda u_h, A\Lambda v_h) = \frac{1}{2} (A\Lambda v_h, \Lambda v_h) + \langle \ell, v_h \rangle - \\
- \frac{1}{2} (A\Lambda u_h, \Lambda u_h) - \langle \ell, u_h \rangle + \\
+ (A\Lambda u_h, \Lambda(u_h - v_h)) + \langle \ell, u_h - v_h \rangle.
\]
Since \( u_h \in V_h \) is a Galerkin approximation, the last two terms vanish and we obtain the relation.
We know that
\[
\| \Lambda(u_h - u) \|^2 = 2(J(u_h) - J(u)), \\
\| \Lambda(v_h - u) \|^2 = 2(J(v_h) - J(u)).
\]
Therefore,

\[ 2D(\Lambda u_h, \mathcal{A}\Lambda v_h) = 2(J(v_h) - J(u)) - 2(J(u_h) - J(u)) = \]
\[ = \| \Lambda(v_h - u) \|^2 - \| \Lambda(u_h - u) \|^2. \]

Now, the error estimate comes in the form

\[ \| \Lambda(u - u_h) \| \leq (1 + \beta)(\| \Lambda(v_h - u) \|^2 - \| \Lambda(u_h - u) \|^2) + \]
\[ + \left(1 + \frac{1}{\beta}\right) \| \Lambda(v_h - u) \|^2. \]

Thus, we obtain

\[ (2 + \beta) \| \Lambda(u - u_h) \|^2 \leq \]
\[ \leq (1 + \beta) \| \Lambda(v_h - u) \|^2 + \left(1 + \frac{1}{\beta}\right) \| \Lambda(v_h - u) \|^2, \]
We see that

\[ \| \Lambda(u - u_h) \|^2 \leq \left( 1 + \frac{1}{\beta(2 + \beta)} \right) \| \Lambda(u - v_h) \|^2. \]

Since \( \beta \) is an arbitrary positive number, we arrive at the projection type error estimate

\[ \| \Lambda(u - u_h) \| \leq \inf_{v_h \in V_h} \| \Lambda(u - v_h) \|. \]
Finally, we note that functional a posteriori estimates also imply a projection type error estimate of a different type. Let us set $v = u_h$, $y = y_h := A\nabla u_h$. Since

$$D(\Lambda u_h, y_h) = 0,$$

we have

$$\| \Lambda (u_h - u) \|^2 \leq \| y_h - q \|^2 \quad \forall q \in Q^*_\ell.$$
From here, it follows the estimate

\[ \| \Lambda(u - u_h) \| \leq \inf_{q \in Q^*_\ell} \| y_h - q \|^*, \]

which is in a sense \textbf{dual} to the first one. It shows that an upper bound of the error is also given by the distance in the space \( Y^* \) between the "Galerkin flux" \( \Lambda \nabla u_h \) and the set \( Q^*_\ell \) that contains the solution of the dual problem.
Lecture 6.
FUNCTIONAL A POSTERIORI ESTIMATES. LINEAR ELLIPTIC PROBLEMS.
Main goal of the lecture

In the previous lecture we have analyzed the abstract linear problem of the form

\[ \Lambda^* \mathcal{A} \Lambda u + \ell = 0 \]

and obtained an estimate

\[ \frac{1}{2} \| \Lambda(v - u) \|^2 \leq (1 + \beta)D(\Lambda v, y) + \frac{1 + \beta}{2\beta} \| \ell + \Lambda^* y \|^2. \]

In the present lecture, we discuss particular forms of this general estimate for some elliptic type boundary–value problems.
Lecture plan

- Diffusion equation with Dirichlet boundary conditions;
- Diffusion equation with Neumann boundary conditions;
- Diffusion equation with mixed boundary conditions;
- Linear elasticity with mixed boundary conditions;
Let $\mathcal{A}$ is produced by a matrix $A = \{a_{ij}\} = \{a_{ji}\}$, $V = H^1(\Omega)$, where $\Omega$ is a Lipschitz domain, $U = L^2(\Omega, \mathbb{R}^n)$, and $\Lambda w = \nabla w$. Let the entries of $A$ be bounded at almost all points of $\Omega$ and such that

$$c_1|\xi|^2 \leq a_{ij}\xi_i\xi_j \leq c_2|\xi|^2, \quad \forall \xi \in \mathbb{R}^n. \tag{6.1}$$

Then, the spaces $Y$ and $Y^*$ have the norms

$$\|y\|^2 = \int_{\Omega} A y \cdot y \, dx, \quad \|y\|^2_* = \int_{\Omega} A^{-1} y \cdot y \, dx.$$
Dirichlet boundary conditions

We begin with the problem

\[ \text{div} A \nabla u = f \quad \text{in} \quad \Omega, \quad (6.2) \]
\[ u = u_0 \quad \text{on} \quad \partial \Omega. \quad (6.3) \]

In this case, \( V_0 = \overset{\circ}{H}^1(\Omega) \) and \( u \) meets the integral identity

\[ \int_{\Omega} A \nabla u \cdot \nabla w \, dx + \langle f, w \rangle = 0, \quad \forall w \in V_0. \quad (6.4) \]

The relation \((y, \Lambda w) = \langle \Lambda^* y, w \rangle\) has the form

\[ \int_{\Omega} y \cdot \nabla w \, dx = \langle -\text{div} y, w \rangle, \]

where \( \Lambda^* = -\text{div} \) and \( \text{div} y \) is in \( H^{-1}(\Omega) \).
The operator $\Lambda$ satisfies the required inequality

$$c_{\Omega} \| \nabla w \| \geq \| w \|, \quad \forall w \in \overset{\circ}{H}^1(\Omega).$$

Upper estimates of $\| v - u \|$ for an approximation $v \in V_0 + u_0$ follow from the general estimate presented in Lecture 5. We have

$$\frac{1}{2} \int_{\Omega} A \nabla (v - u) \cdot \nabla (v - u) \, dx \leq M_{\oplus}(v, \beta, y),$$

where

$$M_{\oplus}(v, \beta, y) =$$

$$1 + \frac{\beta}{2} \int_{\Omega} \left( \nabla v - A^{-1} y \right) \cdot \left( A \nabla v - y \right) \, dx + \frac{1 + \beta}{2\beta} \frac{c_1^2}{c_2} \| \text{div} y - f \|^2$$

(6.5)
Certainly, the above estimate is applicable for the case $f \in L^2(\Omega)$ so that

$$\langle f, w \rangle = \int_{\Omega} fw \, dx,$$

and for $y \in H(\Omega, \text{div})$.

Let $\{Y_k^*\}$ be finite-dimensional subspaces of $Y^*$ such that

$$Y_k^* \in H(\Omega, \text{div}) \quad \text{for all } k = 1, 2, \ldots;$$

$$\dim Y_k^* \to +\infty \quad \text{as } k \to \infty.$$

We obtain computable upper bounds

$$M_k^\oplus = \inf_{\substack{y \in Y_k^* \\ \beta \in \mathbb{R}^+}} \left\{ \frac{1+\beta}{2} \int_{\Omega} (\nabla v - A^{-1}y) \cdot (A \nabla v - y) \, dx + \frac{1+\beta}{2\beta} \frac{c_1^2}{c_2} \| \text{div } y - f \|_{\Omega}^2 \right\}.$$ (6.6)
Lower estimates follow

We have

\[ \frac{1}{2} \int_{\Omega} A \nabla (v - u) \cdot \nabla (v - u) \, dx \geq M_{\Theta}(v, w), \quad \forall w \in V_0, \]

where

\[ M_{\Theta}(v, w) = -\frac{1}{2} \int_{\Omega} A \nabla w \cdot \nabla w \, dx - \int_{\Omega} A \nabla v \cdot \nabla w \, dx - \langle f, w \rangle. \]

Let \( \{ V_{0k} \} \) be finite-dimensional subspaces such that

\[ V_{0k}^* \in V_0 \quad \text{for all } k = 1, 2, \ldots; \]

\[ \dim V_{0k} \to +\infty \quad \text{as } k \to \infty. \]

Find the numbers

\[ M_k = \sup_{w_k \in V_{0k}} M_{\Theta}(v, w_k). \quad (6.7) \]
Both sequences $M^k_\ominus$ and $M^k_\oplus$ tend to $\frac{1}{2} \| v - u \|^2$ as $k \to \infty$, provided that $\{Y^*_k\}$ and $\{V_{0k}\}$ possess necessary approximation properties (limit density).

Note that if $v$ is a Galerkin approximation computed on $V_{0k}$, then $M_\ominus(v, w_k) = 0$. This means that to obtain a sensible lower estimate in this case, one must always use a finite-dimensional subspace that is larger than $V_{0k}$.
Consider the Neumann boundary condition
\[ \nu \cdot A \nabla u + F = 0 \quad \text{on} \quad \partial \Omega, \quad (6.8) \]
where \( \nu \) is the vector of unit outward normal to \( \partial \Omega \). To apply the general scheme we set
\[ V_0 := \left\{ v \in H^1(\Omega) \mid \int_\Omega v \, dx = 0 \right\} \]
and define \( \Lambda^* y \in V_0^* \) by the relation
\[ \langle \Lambda^* y, w \rangle = \int_\Omega y \cdot \nabla w \, dx, \quad \forall w \in V_0. \]
If \( y \) is sufficiently regular then
\[ \langle \Lambda^* y, w \rangle = \int_\Omega (-\text{div} y) w \, dx + \int_{\partial \Omega} (y \cdot \nu) w \, dx. \]
Therefore, in such a case

$$\Lambda^* y = [-\text{div} y \mid \Omega; \quad y \cdot \nu \mid \partial \Omega]$$

Also, we assume that $F$ and $f$ satisfy the equilibrium condition

$$\int_{\Omega} f \, dx + \int_{\partial \Omega} F \, dx = 0.$$ 

Assume that $f \in L^2(\Omega)$ and $F \in L^2(\partial \Omega)$. Then the Neumann problem has a solution defined by the integral identity

$$\int_{\Omega} A \nabla u \cdot \nabla w \, dx + \langle \ell, w \rangle = 0, \quad \forall w \in V_0,$$

where

$$\langle \ell, w \rangle = \int_{\Omega} fw \, dx + \int_{\partial \Omega} Fw \, ds.$$
In general, \( \ell + \Lambda^* y \) is estimated in terms of the norms
\[
\| \text{div } y - f \|_{H^{-1}} \quad \text{and} \quad \| y \cdot \nu + F \|_{H^{-1/2}}.
\]
However, if we assume that \( y \) possesses a certain regularity, so that
\[
y \in Q^*(\Omega) := \{ y \in Y^* | \text{div } y \in L^2(\Omega), y \cdot \nu \in L^2(\partial\Omega) \},
\]
then
\[
\langle \ell + \Lambda^* y, w \rangle = \int_{\Omega} (f - \text{div } y)w \, dx + \int_{\partial\Omega} (F + y \cdot \nu)w \, ds
\]
and, therefore,
\[
|\langle \ell + \Lambda^* y, w \rangle| \leq \| \text{div } y - f \|_{2,\Omega} \| w \|_{2,\Omega} + \| y \cdot \nu + F \|_{2,\partial\Omega} \| w \|_{2,\partial\Omega}. \quad (6.9)
\]
Let the constant $c_\Omega$ be defined as
\[
\frac{1}{c^2_{(\Omega,\partial\Omega)}} = \inf_{w \in V_0} \frac{\int_{\Omega} A \nabla w \cdot \nabla w \, dx}{\|w\|^2_{2,\Omega} + \|w\|^2_{2,\partial\Omega}}.
\]
Since the trace operator is bounded, this constant is finite. Therefore, (6.9) implies the estimate
\[
|\langle \ell + \Lambda^*y, w \rangle| \leq c_{(\Omega,\partial\Omega)} \left( \| \text{div } y - f \|^2_{2,\Omega} + \|y \cdot \nu + F\|^2_{2,\partial\Omega} \right)^{1/2} \| \Lambda w \|^2
\]
and the second term of the majorant is calculated as follows:
\[
\| \ell + \Lambda^*y \| = \sup_{w \in V_0} \frac{\langle \ell + \Lambda^*y, w \rangle}{\| \Lambda w \|} \leq c_{(\Omega,\partial\Omega)} \left( \| \text{div } y - f \|^2_{2,\Omega} + \|y \cdot \nu + F\|^2_{2,\partial\Omega} \right)^{1/2}.
\]
The term $D(\Lambda v, y)$ is defined as in the Dirichlet problem.
We see that the Majorants $M_\oplus$ for the two main boundary-value problems have different values of $c_\Omega$. In addition, the Neumann problem majorant contains an extra term

$$\|y \cdot \nu + F\|_{2, \partial \Omega}$$

that penalizes violations of the Neumann boundary condition. It is worth noting that if the given $F$ can be exactly reproduced by $y \cdot \nu$ for $y$ in a certain finite dimensional subspace $Y^*_k$, then one can compute $M_k^\oplus$ as

$$M_k^\oplus = \inf_{y \in Y^*_k, \ y \cdot \nu = F \text{ on } \partial \Omega} \left\{ \frac{1 + \beta}{2} \int_\Omega (\nabla v - A^{-1}y) \cdot (A \nabla v - y) \, dx + \frac{1 + \beta}{2\beta} c_{(\Omega, \partial \Omega)}^2 \| \text{div} \, y - f \|_{\Omega}^2 \right\}. \quad (6.10)$$
Mixed boundary conditions

Let \( \partial \Omega \) consist of two measurable nonintersecting parts \( \partial_1 \Omega \) and \( \partial_2 \Omega \), on which different boundary conditions are given:

\[
\begin{align*}
\mathbf{u} &= \mathbf{u}_0 \quad \text{on} \quad \partial_1 \Omega, \\
\nu \cdot A \nabla \mathbf{u} + \mathbf{F} &= 0 \quad \text{on} \quad \partial_2 \Omega.
\end{align*}
\]

Set

\[
V_0 := \left\{ \mathbf{v} \in H^1(\Omega) \mid \mathbf{v} = 0 \quad \text{on} \quad \partial_1 \Omega \right\}
\]

and

\[
\langle \Lambda^* \mathbf{y}, \mathbf{w} \rangle = \int_{\Omega} \mathbf{y} \cdot \nabla \mathbf{w} \, dx, \quad \forall \mathbf{w} \in V_0.
\]
Assume that
\[ f \in L^2(\Omega), \quad F \in L^2(\partial_2 \Omega). \]
and \( y \) possesses an extra regularity, namely,
\[ y \in Q^*(\Omega) := \left\{ y \in Y^* \mid \text{div } y \in L^2(\Omega), \ y \cdot \nu \in L^2(\partial_2 \Omega) \right\}. \]
Then, for any \( w \subset V_0 \), we have
\[
\langle \ell + \Lambda^* y, w \rangle = \int_{\Omega} (\text{div } y - f)w \, dx + \int_{\partial_2 \Omega} (y \cdot \nu + F)w \, ds,
\]
Note that \( p \in Q^*(\Omega) \)!
Now, we obtain

\[ |\langle \ell + \Lambda^* y, w \rangle| \leq \| \text{div} y - f \|_{2,\Omega} \| w \|_{2,\Omega} + \| y \cdot \nu + F \|_{2,\partial_2 \Omega} \| w \|_{2,\partial_2 \Omega}. \]

Let \( \gamma \) and \( \gamma^* \) be two numbers such that \( \gamma > 1, \quad \gamma^* > 1, \)
\( \frac{1}{\gamma} + \frac{1}{\gamma^*} = 1. \) Use the algebraic inequality

\[
ab + cd \leq \sqrt{\gamma a^2 + \gamma^* c^2} \sqrt{\frac{1}{\gamma} b^2 + \frac{1}{\gamma^*} d^2}.
\]

Then

\[ |\langle \ell + \Lambda^* y, w \rangle| \leq \left( \gamma \| \text{div} y - f \|_{2,\Omega}^2 + \gamma^* \| y \cdot \nu + F \|_{2,\partial_2 \Omega}^2 \right)^{1/2} \times \]
\[ \times \left( \frac{1}{\gamma} \| w \|_{2,\Omega}^2 + \frac{1}{\gamma^*} \| w \|_{2,\partial_2 \Omega}^2 \right)^{1/2}. \]
Since (Friederichs type inequality)
\[ \| w \|_{2,\Omega}^2 \leq C_F^2(\Omega) \| \nabla w \|_{2,\Omega}^2, \quad \forall w \in V_0, \]
and (trace inequality)
\[ \| w \|_{2,\partial_2\Omega}^2 \leq C_{tr}^2(\Omega, \partial_2\Omega) \| w \|_{1,2,\Omega}^2, \quad \forall w \in V_0, \]
we find that
\[ \frac{1}{\gamma} \| w \|_{2,\Omega}^2 + \frac{1}{\gamma^*_*} \| w \|_{2,\partial_2\Omega}^2 \leq \]
\[ \leq C_F^2 \frac{1}{\gamma} \| \nabla w \|^2 + C_{tr}^2 \frac{1}{\gamma^*_*} \left( \| w \|_{2,\Omega}^2 + \| \nabla w \|_{2,\Omega}^2 \right) \leq \]
\[ \leq \left( C_F^2 \frac{1}{\gamma} + C_{tr}^2 \frac{1}{\gamma^*_*} \left( 1 + C_F^2 \right) \right) \| \nabla w \|_{2,\Omega}^2. \]
Therefore, there exist a positive constant $C_\gamma$ such that

$$\frac{1}{C_\gamma^2} = \inf_{w \in V_0} \frac{\int_{\Omega} A \nabla w \cdot \nabla w \, dx}{\frac{1}{\gamma} \| w \|_{2,\Omega}^2 + \frac{1}{\gamma^*} \| w \|_{2,\partial_2 \Omega}^2}.$$  

The value of this constant can be estimated numerically by minimizing the above quotient on a sufficiently representative finite dimensional subspace. Besides, if $C_F$ and $C_{tr}$ are estimated, then

$$C_\gamma^2 \leq \hat{C}_\gamma^2 := \left( C_F^2 \frac{1}{\gamma} + C_{tr}^2 (1 + C_F^2) \frac{1}{\gamma^*} \right) c_{1}^{-1},$$

so that an upper bound of $C_\gamma$ is directly computed. Now,

$$|\langle \ell + \Lambda^* y, w \rangle| \leq$$

$$\leq \hat{C}_\gamma \left( \gamma \| \text{div} y - f \|_{2,\Omega}^2 + \gamma^* \| y \cdot \nu + F \|_{2,\partial_2 \Omega}^2 \right)^{1/2} \| \nabla w \|.$$
From this estimate, we obtain

\[ \|\ell + \Lambda^* y\|^2 \leq \hat{C}_2^2 \left( \gamma \|\text{div} \ y - f\|^2_{2,\Omega} + \frac{\gamma}{\gamma - 1} \|y \cdot \nu + F\|^2_{2,\partial_2 \Omega} \right). \]

Consider first the case, in which we simply set \( \gamma = \gamma^* = 2 \). Then

\[ \hat{C}_2^2 \left( \gamma = 2 \right) := \hat{C}_2^2 = \frac{1}{2} \left( C_F^2 + C_{tr}^2 (1 + C_F^2) \right) c_1^{-1}, \]

\[ \|\ell + \Lambda^* y\|^2 \leq 2 \hat{C}_2^2 \left( \|\text{div} \ y - f\|^2_{2,\Omega} + \|y \cdot \nu + F\|^2_{2,\partial_2 \Omega} \right). \]

and we find that

\[ M_\oplus(v, \beta, y) = \frac{1 + \beta}{2} \int_\Omega (\nabla v - A^{-1} y) \cdot (A \nabla v - y) \, dx + \]

\[ + \frac{1 + \beta}{2\beta} \hat{C}_2^2 \left( \|\text{div} \ y - f\|^2_{2,\Omega} + \|y \cdot \nu + F\|^2_{2,\partial_2 \Omega} \right). \quad (6.11) \]

This Majorant gives an upper bound of the deviation for any \( v \in V_0 + u_0, \ y \in Q^*, \) and \( \beta > 0. \)
A more exact estimate is obtained if we define $\gamma$ by minimizing of the quantity

$$
\left( C_F^2 \frac{1}{\gamma} + C_{tr}^2 (1 + C_F^2) \frac{1}{\gamma^*} \right) \left( \gamma \| \text{div} \, y - f \|_{2,\Omega}^2 + \gamma^* \| \nu y + F \|_{2,\partial_2 \Omega}^2 \right) =
C_F^2 \frac{\gamma^*}{\gamma} \| \nu \cdot y + F \|_{2,\partial_2 \Omega}^2 + \frac{\gamma^*}{\gamma} C_{tr}^2 (1 + C_F^2) \| \text{div} \, y - f \|_{2,\Omega}^2 + \text{const}(\gamma).
$$

Denote

$$
\rho_1 = \| \text{div} \, y - f \|_{2,\Omega}^2, \quad \rho_2 = \| \nu \cdot y + F \|_{2,\partial_2 \Omega}^2,
\kappa_1 = C_F^2, \quad \kappa_2 = C_{tr}^2 (1 + C_F^2).
$$

Then the problem is

$$
\min_{\gamma} \left( \kappa_1^2 \rho_2^2 \frac{1}{\gamma} + (\gamma - 1) \kappa_2^2 \rho_1^2 \right).
$$

Its minimum is attained at $\hat{\gamma} = 1 + \kappa_1 \rho_2 / \kappa_2 \rho_1$. 
In other words, we observe that the lowest estimate of the term $\| \ell + \Lambda^* y \|$ is attained if

$$\gamma = \hat{\gamma} := 1 + \frac{\| y \cdot \nu + F \|_{2, \partial_2 \Omega} C_F}{\| \text{div} \ y - f \|_{2, \Omega} C_{\text{tr}} (1 + C_F^2)^{1/2}}.$$ 

Let us find the respective upper bound. We need to calculate

$$\left( \frac{\kappa_1^2}{\hat{\gamma}} + \frac{\kappa_2^2}{\hat{\gamma}^*} \right) (\hat{\gamma} \rho_1^2 + \hat{\gamma}^* \rho_2^2) =$$

$$= \frac{1}{\hat{\gamma}} (\kappa_1^2 + (\hat{\gamma} - 1) \kappa_2^2) \frac{\hat{\gamma}}{\hat{\gamma} - 1} ((\hat{\gamma} - 1) \rho_1^2 + \rho_2^2) =$$

$$= \frac{\kappa_2 \rho_1}{\kappa_1 \rho_2} (\kappa_1^2 + \frac{\kappa_1 \rho_2}{\kappa_2 \rho_1} \kappa_2^2) (\rho_1^2 \frac{\kappa_1 \rho_2}{\kappa_2 \rho_1} + \rho_2^2) = \kappa_2 \rho_1 (\kappa_1 + \frac{\rho_2}{\rho_1} \kappa_2) (\rho_1 \frac{\kappa_1}{\kappa_2} + \rho_2) =$$

$$= (\kappa_1 \rho_1 + \rho_2 \kappa_2) (\rho_1 \kappa_1 + \rho_2 \kappa_2) = (\kappa_1 \rho_1 + \rho_2 \kappa_2)^2.$$
By recalling the definitions of $\kappa_1$, $\kappa_2$, $\rho_1$, and $\rho_2$ we obtain

$$]\ell + \Lambda^* y [^2 \leq \left( C_F \| \text{div} \, y - f \|_{2,\Omega} + 
+ C_{tr}(1 + C_F^2)^{1/2} \| y \cdot \nu + F \|_{2,\partial_2 \Omega} \right)^2 c_1^{-2}$$

and we have

$$M_\oplus(v, \beta, y) = \frac{1 + \beta}{2} \int_{\Omega} (\nabla v - A^{-1} y) \cdot (A \nabla v - y) \, dx + 
+ \frac{1 + \beta}{2\beta} \left( C_F \| \text{div} \, y - f \|_{2,\Omega} + 
+ C_{tr}(1 + C_F^2)^{1/2} \| y \cdot \nu + F \|_{2,\partial_2 \Omega} \right)^2 c_1^{-2}. \quad (6.12)$$

Majorant vanishes if and only if $v = u$ and $y = A \nabla u$, it is continuous with respect to the convergence of $v$ in $V$ and $y$ in $Q$. 

S. Repin

LECTURES ON A POSTERIORI ERROR CONTROL

RICAM, Special Radon Semester, Linz, 2005.
Lower estimates

Lower estimates for the problems considered follow from the general ones obtained in the previous lecture. They have the form

$$\frac{1}{2} \int_{\Omega} A \nabla (v - u) \cdot \nabla (v - u) \, dx \geq M_\Theta(v, w), \quad \forall w \in V_0,$$

where

$$M_\Theta(v, w) = -\frac{1}{2} \int_{\Omega} A \nabla (w - v) \cdot \nabla w \, dx - \int_{\Omega} fw \, dx - \int_{\partial_2 \Omega} Fw \, ds.$$

Here $V_0$ depends on the type of boundary conditions, and the integral over $\partial_2 \Omega$ must be eliminated in the case of Dirichlet problem.
Linear elasticity

**Classical statement.** The classical formulation is as follows:

Find a tensor-valued function $\sigma^*$ (stress) and a vector-valued function $u$ (displacement) that satisfy the system of equations

\[
\sigma^* = L\varepsilon(u) \quad \text{in} \quad \Omega, \quad \text{(Hooke’s law)}
\]
\[
div\sigma^* = f \quad \text{in} \quad \Omega, \quad \text{(Equilibrium equation)}
\]
\[
u = u_0 \quad \text{on} \quad \partial_1\Omega,
\]
\[
\sigma^*\nu + F = 0 \quad \text{on} \quad \partial_2\Omega.
\]

where $\varepsilon(u)$ is a symmetric part of the tensor $\nabla u$. 

S. Repin

LEC TURES ON A POSTERIORI ERROR CONTROL

RICAM, Special Radon Semester, Linz, 2005.
Here $\Omega$ is a bounded domain with Lipschitz boundary $\partial \Omega$ that consists of two disjoint parts $\partial_1 \Omega$ and $\partial_2 \Omega$, $|\partial_1 \Omega| > 0$, $\mathbf{f}$ and $\mathbf{F}$ are given forces and $\mathbb{L} = \{L_{ijkm}\}$ is the tensor of elasticity constants, which is subject to the conditions

$$C_1 |\varepsilon|^2 \leq \mathbb{L} \varepsilon : \varepsilon \leq C_2 |\varepsilon|^2, \quad \forall \varepsilon \in \mathbb{M}_{n \times n}^n,$$

and

$$L_{ijkm} = L_{jikm} = L_{kmij}, \quad L_{ijkm} \in L^\infty(\Omega).$$
Generalized solution

Let

\[ f \in L^2(\Omega, \mathbb{R}^n), \quad F \in L^2(\partial_2 \Omega, \mathbb{R}^n). \]

Then, a generalized solution \( u \in V_0 + u_0 \) is defined by the identity

\[ \int_\Omega \nabla (u) : \nabla (w) \, dx + \langle \ell, w \rangle = 0, \quad \forall w \in V_0, \quad (6.13) \]

where

\[ \langle \ell, w \rangle = \int_\Omega f \cdot w \, dx + \int_{\partial_2 \Omega} F \cdot w \, ds. \]
Assume that $u$ is a smooth function and it satisfies the identity

$$\int_{\Omega} \mathbb{L} \varepsilon(u) : \varepsilon(w) \, dx + \langle \ell, w \rangle = 0, \quad \forall w \in V_0,$$

Then,

$$\int_{\Omega} (f - \text{div} \mathbb{L} \varepsilon(u)) \cdot w \, dx + \int_{\partial_2 \Omega} \left( (\mathbb{L} \varepsilon(u))\nu + F \right) \cdot w \, ds = 0,$$

$$\forall w \in V_0,$$

and we observe that in such a case the equilibrium equation and the Neumann boundary condition are satisfied in the classical sense.
Variational formulation

Note that the relation (6.13) is the Euler’s equation for the functional

$$J(v) = \frac{1}{2} \int_{\Omega} \mathbb{L}\varepsilon(v) : \varepsilon(v) \, dx + \langle \ell, v \rangle.$$ 

Therefore, the respective boundary–value problem may be considered as a minimization problem for $J(v)$ on the set

$$V_0 := \{ v \in H^1(\Omega, \mathbb{R}^n) \mid v = u_0 \text{ on } \partial_1 \Omega \}.$$

To prove existence of a minimizer we must show the coercivity of $J(v)$ on $V_0$. The key role in this belongs to the so–called Korn’s inequality.
In the Dirichlet problem

\[ J(v) = \frac{1}{2} \int_{\Omega} \mathbb{L}\varepsilon(v) : \varepsilon(v) \, dx + < \ell, v > \geq \]

\[ \geq \frac{C_1}{2} \|\varepsilon(v)\|^2 - \|f\| \|v\| = \]

\[ = \frac{C_1}{2} \|\varepsilon(u_0 + w)\|^2 - \|f\| \|u_0 + w\| \geq \]

\[ \geq \frac{C_1}{2} (\|\varepsilon(u_0)\| - \|\varepsilon(w)\|)^2 - \|f\| \|u_0\| - \|f\| \|w\|. \]

Thus, if we can prove that

\[ \|\varepsilon(w)\| \geq c \|\nabla w\| \quad \forall w \in H^1(\Omega), \]

then we would establish the coercivity of \( J \).
Korn's inequality

This inequality is required in various aspects of the mathematical analysis of elasticity problems. In the general form it states the equivalence of two norms:

\[ \|w\|_{1,2,\Omega} := \left( \int_\Omega \left( |\nabla w|^2 + |w|^2 \right) \, dx \right)^{1/2}, \]

and

\[ \|w\|_{1,2,\Omega} := \left( \int_\Omega \left( |\varepsilon(w)|^2 + |w|^2 \right) \, dx \right)^{1/2}. \]
Korns’s inequality in $\overset{\circ}{H}^1$

For the functions in $\overset{\circ}{H}^1(\Omega)$ this fact is not difficult to prove. Indeed,

$$
\int_{\Omega} |\varepsilon(w)|^2 \, dx = \frac{1}{2} \| \nabla w \|^2 + \frac{1}{2} \int_{\Omega} \sum_{ij} w_{i,j}w_{j,i} \, dx = \frac{1}{2} \| \nabla w \|^2 - \frac{1}{2} \int_{\Omega} \sum_{ij} w_i w_{j,i} \, dx = \frac{1}{2} \| \nabla w \|^2 + \frac{1}{2} \int_{\Omega} \sum_{ij} w_{i,i}w_{j,j} \, dx = \frac{1}{2} \| \nabla w \|^2 + \frac{1}{2} \int_{\Omega} \sum_i |w_{i,i}|^2 \, dx.
$$

Thus,

$$
\| \nabla w \| \leq \sqrt{2} \| \varepsilon(w) \| \quad \forall w \in \overset{\circ}{H}^1(\Omega).
$$

(6.14)
By (6.14) we prove that the energy functional of the elasticity problem for the case of Dirichlet boundary conditions is coercive, i.e.,

\[ J(v_k) \rightarrow +\infty, \quad \text{as} \quad \|\nabla v_k\| \rightarrow +\infty. \]
Rigid deflections

In the analysis of elasticity problems one more notion is often required. It is the so-called *Space of Rigid Deflections* that we denote $\text{RD}(\Omega)$. This space is the kernel of the operator $\varepsilon(w)$, i.e. it contains vector-valued functions $w$ such that

$$\varepsilon(w) = 0.$$ 

It can be defined as follows:

$$\text{RD}(\Omega) := \{ w = w_0 + \omega_0 x \mid w_0 \in \mathbb{R}^n, \omega_0 \in \mathbb{M}^{n \times n} \},$$

where $\omega_0(w) = \frac{1}{2}(\nabla w - (\nabla w)^T)$ is a skew-symmetric tensor associated with "rigid rotations". 

S. Repin

*RICAM, Special Radon Semester, Linz, 2005.*

*LECTURES ON A POSTERIORI ERROR CONTROL*
Implications of the Korn’s inequality

**Theorem**

Let $\Omega$ be a Lipschitz domain and $\partial_1 \Omega$ is a nonempty connected part of the boundary. Then,

$$\|u\|_{1,p,\Omega} \leq C \left( \int_{\Omega} |\varepsilon(u)|^p \, dx \right)^{\frac{1}{p}} \quad \forall u \in V, \quad p \in (1, 2] \quad (6.15)$$

**Proof.** Assume the opposite. Then, for any $m \in \mathbb{N}$ we can find $v^{(m)}$ such that $v^{(m)} \in V$ and

$$\|v^{(m)}\|_{1,p,\Omega} > m \left( \int_{\Omega} |\varepsilon(v^{(m)})|^p \, dx \right)^{\frac{1}{p}}.$$
Set \( w^{(m)} = \frac{v^{(m)}}{\|v^{(m)}\|_{1,p,\Omega}} \), then

\[
\|w^{(m)}\|_{1,p,\Omega} = 1 \quad \text{and} \quad \frac{1}{m} \geq \left( \int_{\Omega} |\varepsilon(w^{(m)})|^p \, dx \right)^{\frac{1}{p}}.
\]

Therefore,

\[
\begin{align*}
& w^{(m)} \rightharpoonup w \quad \text{in} \quad W^1_p(\Omega, \mathbb{R}^n), \\
& w^{(m)} \rightarrow w \quad \in \quad L^p(\Omega, \mathbb{R}^n), \\
& \|\varepsilon(w^{(m)})\|_{p,\Omega} \rightarrow 0 \quad \text{in} \quad L^p(\Omega, \mathbb{R}^n).
\end{align*}
\]

From here we conclude that \( \varepsilon(w) = 0 \).
Indeed, by the fact that a norm is weakly lower semicontinuous, we have

$$0 = \liminf_m \| \varepsilon(w^{(m)}) \|_{p, \Omega} \geq \| \varepsilon(w) \|_{p, \Omega}.$$ 

Thus, \( w \in RD(\Omega) \cap V \). There is only one such a function: \( w = 0 \). It means that \( w^{(m)} \to 0 \) in \( L^p \). Now, we apply Korn’s inequality

$$\| w^{(m)} \|_{1, p, \Omega} \leq C \left( \int_{\Omega} \left( |\varepsilon(w^{(m)})|^p + |w^{(m)}|^p \right) dx \right)^{1/p} \to 0 \quad m \to \infty,$$

which shows that \( \| w^{(m)} \|_{1, p, \Omega} \) tends to zero. But for any \( m \)

\( \| w^{(m)} \|_{p, 1, \Omega} = 1 \), so that such a behavior is impossible. We have arrived at a contradiction that proves the Theorem.
Another similar result is required for the Neumann problem. Define the set

\[ V = \left\{ v \in W^1_p(\Omega) \mid \int_{\Omega} v \cdot w \, dx = 0 \quad \forall w \in RD(\Omega) \right\} . \]

**Theorem**

*Let \( \Omega \) be a bounded domain with Lipschitz boundary \( \partial\Omega \). Then*

\[ \|u\|_{1,p,\Omega} \leq C \left( \int_{\Omega} |\varepsilon(u)|^p \, dx \right)^{\frac{1}{p}} \quad \forall u \in V. \quad (6.16) \]
Proof. By the same arguments as before, we obtain a sequence \( w^{(m)} \in V \) such that

\[
\begin{align*}
  w^{(m)} &\rightharpoonup w \quad \text{in} \quad W^{1}_p(\Omega, R^n), \\
  w^{(m)} \rightarrow w \quad \text{in} \quad L^p(\Omega, R^n), \\
  \|\varepsilon(w^{(m)})\|_{p,\Omega} &\rightarrow 0 \quad \text{in} \quad L^p(\Omega, R^n).
\end{align*}
\]

By the arguments similar to those in the previous Theorem, we find that \( \varepsilon(w) = 0 \) and, thus, \( w \in RD(\Omega) \). In addition, for any \( \tilde{w} \in RD \), we have

\[
0 = \int_{\Omega} w^{(m)} \cdot \tilde{w} \, dx = \int_{\Omega} w \cdot \tilde{w} \, dx.
\]

But \( w \in RD \), so that \( \|w\| = 0 \), and by applying Korn’s inequality we prove that \( \|w^{(m)}\|_{1,p,\Omega} \) tends to zero, what leads to a contradiction.

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
Estimates of deviations

Let $\mathbf{v}$ and $\mathbf{y}$ be some approximations of $\mathbf{u}$ and $\mathbf{\sigma}^*$. Estimates of $\mathbf{v} - \mathbf{u}$ and $\mathbf{y} - \mathbf{\sigma}^*$ follow from the general scheme if we set

$$U = L^2(\Omega, M_{s \times n}^n), \quad V = H^1(\Omega, \mathbb{R}^n),$$

$$V_0 = \{ \mathbf{w} \in V \mid \mathbf{w} = 0 \text{ on } \partial_1 \Omega \},$$

$$\| \mathbf{y} \|_2^2 = \int_{\Omega} \mathbf{L} \mathbf{y} : \mathbf{y} \, dx, \quad \| \mathbf{y} \|_{2}\star^2 = \int_{\Omega} \mathbf{L}^{-1} \mathbf{y} : \mathbf{y} \, dx,$$

and $\Lambda \mathbf{v} = \varepsilon(\mathbf{v}) := \frac{1}{2} \left( \nabla \mathbf{v} + (\nabla \mathbf{v})^T \right)$. In this case,

$$\langle \Lambda^* \mathbf{y}, \mathbf{w} \rangle = \int_{\Omega} \mathbf{y} : \varepsilon(\mathbf{w}) \, dx, \quad \forall \mathbf{w} \in V_0,$$
Now $y$ is a tensor-valued function and $y_\nu = y_{ij}\nu_j$ is a vector-function defined on $\partial\Omega$. If

$$y \in Q^* := \{ y \in Y^* \mid \text{div} y \in L^2(\Omega, \mathbb{M}^{n\times n}), \; y_\nu \in L^2(\partial_2\Omega, \mathbb{R}^n) \}.$$ 

then

$$\langle \Lambda^* y, w \rangle = -\int_\Omega \text{div} y \cdot w \, dx + \int_{\partial_2\Omega} (y_\nu) \cdot w \, d\Gamma$$

so that

$$\Lambda^* y = \{-\text{div} y \mid \Omega, \; (y_\nu) \mid \partial_2\Omega\}.$$
Upper estimates

By applying the general estimate, we obtain the following upper estimate:

\[
\frac{1}{2} \int_{\Omega} \mathbb{L} \varepsilon(v - u) : \varepsilon(v - u) \, dx \leq M_{\oplus}(v, \beta, y),
\]

where

\[
M_{\oplus}(v, \beta, y) = \frac{1 + \beta}{2} D(\varepsilon v, y) + \frac{1 + \beta}{2\beta} \mathbf{I} \Lambda^* y + \ell \mathbf{I}^2
\]

and

\[
D(\varepsilon(v), y) = \frac{1}{2} \int_{\Omega} \left( \mathbb{L} \varepsilon(v) : \varepsilon(v) \mathbb{L}^{-1} y : y - 2 \varepsilon(v) : y \right) \, dx = \int_{\Omega} (\varepsilon(u) - \mathbb{L}^{-1} y) : (\mathbb{L} \varepsilon(u) - y) \, dx.
\]
If $y \in Q^*$, then

$$
\| \Lambda^* y + \ell \| = \sup_{w \in V_0} \frac{\langle \Lambda^* y + \ell, w \rangle}{\| \Lambda w \|} = \sup_{w \in V_0} \frac{\langle \Lambda^* y + \ell, w \rangle}{\| \Lambda w \|} = \frac{\int_{\Omega} (y : \varepsilon(w) + f \cdot w) \, dx + \int_{\partial_2 \Omega} F \cdot w \, ds}{\| \varepsilon(w) \|} = \sup_{w \in V_0} \frac{\int_{\Omega} (f - \text{div} \, y) \cdot w \, dx + \int_{\partial_2 \Omega} (F + y \nu) \cdot w \, ds}{\| \varepsilon(w) \|} \leq \sup_{w \in V_0} \frac{\| f - \text{div} \, y \|_{2, \Omega} \| w \|_{2, \Omega} + \| F + y \nu \|_{\partial_2 \Omega} \| w \|_{\partial_2 \Omega}}{\| \varepsilon(w) \|}.
$$
Let $C_\Omega$ be a constant in the inequality
\[
\int_\Omega |w|^2 \, dx + \int_{\partial_2 \Omega} |w|^2 \, ds \leq C_\Omega^2 \|\varepsilon(w)\|^2_\Omega, \quad \forall w \in V_0.
\]

Note that the existence of such a constant follows from the Korn’s inequality. Indeed, the inequality
\[
\int_\Omega |w|^2 \, dx + \int_{\partial_2 \Omega} |w|^2 \, ds \leq \hat{C}_\Omega^2 \|\nabla(w)\|^2_\Omega, \quad \forall w \in V_0.
\]

for the tensor–gradient $\nabla(w)$ follows from the Friederichs type inequality for the vector–valued functions and the respective trace theorems. By (6.15) we recall that for the functions in $V_0$
\[
\|\nabla(w)\|_\Omega \leq C \|\varepsilon(w)\|_\Omega
\]

with a certain constant $C$ and the estimate follows.
In practice, values of $C_{\Omega}$ can be estimated by minimizing the quotient

$$
\| \varepsilon(w) \|_{\Omega}^2 \over \int_{\Omega} |w|^2 \, dx + \int_{\partial_2 \Omega} |w|^2 \, ds
$$

over sufficiently representative finite dimensional space $V_{0h} \subset V_0$. Let us now return to finding an upper bound of the quantity $\| \Lambda^* y + \ell \|$. By the inequality $ab + cd \leq \sqrt{a^2 + c^2} \sqrt{b^2 + d^2}$, we obtain

$$
\| \Lambda^* y + \ell \| \leq \left( \| \text{div } y - f \|_{\Omega}^2 + \| F + y\nu \|_{\partial_2 \Omega}^2 \right)^{1/2} \sup_{w \in V_0} \| \varepsilon(w) \| \leq C_{\Omega} c_1^{-1/2} \left( \| \text{div } y - f \|_{\Omega}^2 + \| F + y\nu \|_{\partial_2 \Omega}^2 \right)^{1/2}
$$
Error Majorant for mixed boundary conditions

Hence, we arrive at the Majorant $M_\oplus$:

$$
M_\oplus(\varepsilon(v), y) = \frac{1+\beta}{2} \int_\Omega (\varepsilon(u) - \mathbb{I}^{-1}y) : (\mathbb{I} \varepsilon(u) - y) \, dx +
$$

$$
+ \frac{1+\beta}{2\beta c_1} C_\Omega^2 \left( \| \text{div} \ y - f \|_{\Omega}^2 + \| F + y\nu \|_{\partial^2 \Omega}^2 \right). \quad (6.17)
$$

It has a clear physical meaning. The first term of $M_\oplus$ is nonnegative and vanishes if and only if

$$
y = \mathbb{I} \varepsilon(v).
$$

It penalizes violations of the Hooke's law. The meaning of the second term is obvious: it contains $L^2$-norms of other two relations, which gives errors in the equilibrium equation and boundary condition for the stress tensor.
Thus, the majorant not only gives an idea of the overall value of the error, but also shows its physically sensible parts.

Let \( \{ Y^*_k \} \subset H^1(\Omega, M^{n \times n} \) be a collection of finite-dimensional subspaces that satisfy the limit density condition. Then, (6.17) generates a sequence of computable upper bounds

\[
M^k_\oplus = \inf_{y \in Y^*_k, \beta \in \mathbb{R}_+} \left\{ \frac{1+\beta}{2} \int_{\Omega} \left( \mathbb{L} \epsilon(v) : \epsilon(v) + \mathbb{L}^{-1} y : y - 2\epsilon(v) : y \right) dx \right. \\
+ \frac{1+\beta}{2\beta c_1} C^2_\Omega \left( \| \text{div} \ y - f \|_{\Omega}^2 + \| F + y \nu \|_{\partial_2 \Omega}^2 \right) \right\},
\]

which tends to the exact value of the error.
Lower estimates also follow from the general theory. We have
\[ \frac{1}{2} \int_{\Omega} \mathbb{L} \varepsilon(v - u) : \varepsilon(v - u) \, dx \geq M_{\ominus}(v, w), \quad \forall w \in V_0, \]
where
\[ M_{\ominus}(v, w) = -\frac{1}{2} \int_{\Omega} \mathbb{L} \varepsilon(w) : \varepsilon(w) \, dx - \int_{\Omega} \mathbb{L} \varepsilon(v) : \varepsilon(w) \, dx - \int_{\Omega} f \cdot w \, dx - \int_{\partial_2 \Omega} F \cdot w \, ds. \]

By the same arguments as for the diffusion equation one can prove that
\[ \frac{1}{2} \int_{\Omega} \mathbb{L} \varepsilon(v - u) : \varepsilon(v - u) \, dx = \sup_{w \in V_0} M_{\ominus}(v, w_k). \]
By the maximization the functional $M_\Theta$ on a sequence of finite-dimensional spaces $V_{0k} \subset V_0$, we obtain a sequence of computable lower bounds

$$M^k_\Theta = \sup_{w \in V_{0k}} M_\Theta(v, w).$$

If the spaces $V_{0k}$ satisfy the limit density condition stated, then the sequence of numbers $\{M_\Theta\}$ tends to $\frac{1}{2}\|\varepsilon(v-u)\|^2$. 
FUNCTIONAL A POSTERIORI ESTIMATES. FOURTH ORDER EQUATIONS.
Linear elliptic equations of the fourth order

Now, we consider the problem

$$\nabla \cdot \nabla \cdot (B \nabla \nabla u) = f \quad \text{in} \quad \Omega,$$

$$u = \frac{\partial u}{\partial \nu} = 0 \quad \text{on} \quad \partial \Omega. \quad (6.18)$$

Here $\Omega \subset \mathbb{R}^2$, $\nu$ denotes the outward unit normal to the boundary, and $B = \{b_{ijkl}\} \in \mathcal{L}(M_s^{2 \times 2}, M_s^{2 \times 2})$. We assume that $b_{ijkl} = b_{jikl} = b_{klij}$,

$$\alpha_1|\eta|^2 \leq B\eta : \eta \leq \alpha_2|\eta|^2, \quad \forall \eta \in M_s^{2 \times 2},$$

and

$$f \in L^2(\Omega), \quad b_{ijkl} \in L^\infty(\Omega).$$
To apply the general scheme, we set

\[ U = L^2(\Omega, \mathbb{M}_s^{2\times2}), \quad V = H^2(\Omega), \]

\[ V_0 = \{ w \in V \mid w = \frac{\partial w}{\partial \nu} = 0 \text{ on } \partial \Omega \}, \]

and define \( \Lambda \) as the Hessian operator. Now, the basic integral identity has the form

\[
\int_{\Omega} B \nabla \nabla u : \nabla \nabla w \, dx = \int_{\Omega} fw \, dx \quad \forall w \in V_0. \quad (6.20)
\]

By \( B^{-1} \) we denote the inverse tensor, which satisfies the double inequality

\[
\alpha_2^{-1} |\eta|^2 \leq B^{-1} \eta : \eta \leq \alpha_1^{-1} |\eta|^2, \quad \forall \eta \in \mathbb{M}_s^{2\times2},
\]
The spaces $Y$ and $Y^*$ are equipped with norms

$$
\| y \|^2 = \int_{\Omega} B y : y \, dx; \quad \| y \|_{*}^2 = \int_{\Omega} B^{-1} y : y \, dx,
$$

$$
\langle \ell, w \rangle = -\int_{\Omega} f w \, dx,
$$

and

$$
Q^*_\ell = \{ y \in Y^* \mid \int_{\Omega} y : \nabla \nabla w \, dx = \int_{\Omega} f w \, dx, \quad \forall w \in V_0 \}. 
$$

Since

$$
\| \nabla \nabla w \| \geq \alpha_3 \| w \|_{2,2,\Omega} \quad \forall w \in V_0,
$$

we have the required version of the coercivity condition

$$
\| \Lambda w \| \geq c_3 \| w \|_V.
$$

Problem (6.18) and (6.19) is associated with two variational problems.
Problem $\mathcal{P}$. Find $u \in V_0$ such that

$$J(u) = \inf_{v \in V_0} J(v),$$

where

$$J(v) = \frac{1}{2} \int_{\Omega} B \nabla \nabla v : \nabla \nabla v \, dx - \int_{\Omega} fw \, dx.$$

Problem $\mathcal{P}^*$. Find $p \in Q^*_{\ell}$ such that

$$I^*(p) = \sup_{\forall q \in Q^*_{\ell}} I^*(q),$$

where

$$I^*(q) = -\frac{1}{2} \int_{\Omega} B^{-1}q : q \, dx.$$
By In this case, the two basic relations for deviations derived in Lecture 5 come in the form:

\[ \| \nabla \nabla (v - u) \|^2 + \| q - p \|^2_* = 2(J(v) - I^*(q)), \]  
(6.21)

and

\[ \| \nabla \nabla (v - u) \|^2 + \| q - p \|^2_* = 2D(\nabla \nabla v, q) = \]
\[ = \int_\Omega \left( B \nabla \nabla v : \nabla \nabla v + B^{-1}q : q - 2\nabla \nabla v : q \right) \, dx, \]  
(6.22)

which hold for any \( v \in V_0 \) and \( q \in Q^*_\ell \).

Also, from the general theory it readily follows the first a posteriori estimate:

\[ \frac{1}{2} \| \nabla \nabla (v - u) \|^2 \leq \left( 1 + \beta \right) D(\nabla \nabla v, y) + \left( 1 + \frac{1}{\beta} \right) \frac{d^2_\ell(y)}{2}, \]  
(6.23)

where \( d^2_\ell(y) = \inf_{q \in Q^*_\ell} \| q - y \|^2_* \).
Note that

\[ \int_{\Omega} y : \nabla \nabla w \, dx = \int_{\Omega} (\text{divdiv } y) w \, dx, \quad \forall w \in V_0, \]

so that \( \Lambda^* : Y^* \to H^{-2}(\Omega) \) is the operator \text{divdiv}. Next,

\[ \langle \ell + \Lambda^* y, w \rangle = \int_{\Omega} (y : \nabla \nabla w - fw) \, dx \]

and, therefore,

\[ d_\ell^2(y) = I \ell + \Lambda^* y I = \sup_{w \in V_0} \frac{\int_{\Omega} (y : \nabla \nabla w - fw) \, dx}{\| \nabla \nabla w \|}. \]
If

\[ y \in H(\text{divdiv}, \Omega) := \left\{ y \in L^2(\Omega, M_s^{n \times n}) \mid \text{divdiv } y \in L^2(\Omega) \right\}, \]

then this quantity is estimated by the relation

\[
\ell + \Lambda^* y \leq \sup_{w \in V_0} \frac{\| \text{divdiv } y - f \|_\Omega \| w \|_\Omega}{\| \nabla \nabla w \|} \leq \frac{C_{1\Omega}}{\alpha_1} \| \text{divdiv } y - f \|_\Omega,
\]

in which \( C_{1\Omega} \) is a constant in the inequality

\[
\| w \|_\Omega \leq C_{1\Omega} \| \nabla \nabla w \|_\Omega \quad \forall w \in V_0.
\]

Now, we obtain the first variant of a posteriori estimate for the biharmonic type problem.
First a posteriori estimate

\[
\frac{1}{2} \| \nabla \nabla (v - u) \|^2 \leq (1 + \beta) D(\nabla \nabla v, y) + \left(1 + \frac{1}{\beta}\right) \frac{C^2_{1,\Omega}}{2\alpha^2_1} \| \text{div} \text{div} y - f \|^2_{\Omega}, \quad (6.24)
\]

Here, \( y \) is an arbitrary tensor-valued function from \( H(\text{div} \text{div}, \Omega) \) and \( \beta \) is a positive real number. However, this is rather demanding in relation to the dual variable \( y \) (which must have square summable \( \text{divdiv} \)). To avoid technical difficulties that rises from this condition, we estimate the negative norm in a different way.
\[ \| \ell + \Lambda^* y \| = \sup_{w \in V_0} \frac{\int_{\Omega}(y : \nabla \nabla w - fw) \, dx}{\| \nabla \nabla w \|} = \sup_{w \in V_0} \frac{\int_{\Omega}(y : \nabla \nabla w + \eta \cdot \nabla w + \text{div} \eta w - fw) \, dx}{\| \nabla \nabla w \|} \leq \frac{C_{2\Omega}}{\alpha_1} \| \text{div} y - \eta \|_{\Omega} + \frac{C_{1\Omega}}{\alpha_1} \| \text{div} \eta - f \|_{\Omega}. \]

Here, \( \eta \) is an arbitrary vector-valued function from \( H(\text{div}, \Omega) \) and \( C_{2\Omega} \) is a constant in the inequality

\[ \| \nabla w \|_{\Omega} \leq C_{2\Omega} \| \nabla \nabla w \|_{\Omega} \quad \forall w \in V_0. \]
Second a posteriori estimate

Then, we arrive at the estimate

\[
\frac{1}{2} \left\| \nabla \nabla (v - u) \right\|^2 \leq (1 + \beta) D(\nabla \nabla v, y) + \\
+ \left( 1 + \frac{1}{\beta} \right) \frac{1}{2\alpha^2_1} \left( C_{2\Omega} \left\| \text{div} y - \eta \right\|_{\Omega} + C_{1\Omega} \left\| \text{div} \eta - f \right\|_{\Omega} \right)^2 ,
\]

in which \( y \in \Sigma_{\text{div}}(\Omega) \) and \( \eta \in H(\text{div}, \Omega) \).

This estimate was obtained in

Note that

$$\|w\| \leq C_F \|\nabla w\|_\Omega \leq C_F C_{2\Omega} \|\nabla \nabla w\|_\Omega \quad \forall w \in V_0.$$  

where $C_F$ is a constant in the Friederichs inequality. Therefore, $C_{1\Omega} \leq C_F C_{2\Omega}$. In view of this, we obtain a slightly different form of the deviation estimate:

$$\frac{1}{2} \|\nabla \nabla (v - u)\|^2 \leq (1 + \beta)D(\nabla \nabla v, y) + \left(1 + \frac{1}{\beta}\right) \frac{C_{2\Omega}^2}{2\alpha_1^2} \left(\|\text{div} y - \eta\|_\Omega + C_F \|\text{div} \eta - f\|_\Omega\right)^2,$$  

(6.26)

For boundary conditions of other types, the deviation majorants can be derived by arguments similar to those used in Lecture 6.
Lower estimates of the deviation from $u$

Lower estimates follow from the general estimate discussed in Lecture 5. We have

$$\frac{1}{2} \| \nabla \nabla (v - w) \|^2 \geq M_\Theta(v, w) \quad w \in V_0, \quad (6.27)$$

where

$$M_\Theta(v, w) := -\frac{1}{2} \| \nabla w \|^2 - \int_\Omega (B \nabla \nabla v : \nabla \nabla w - fw) dx.$$
Lecture 7.
FUNCTIONAL A POSTERIORI ESTIMATES. STOKES PROBLEM.
Lecture plan

- Stokes problem;
- Inf-sup condition;
- A posteriori estimates for solenoidal approximations;
- A posteriori estimates for non-solenoidal approximations;
- A posteriori estimates for problems with condition \( \text{div} v = \phi \);
- A posteriori estimates for problems on a subspace.
Stokes problem

George Stokes

**Classical formulation** of the Stokes problem: find a vector–valued function $u$ (velocity) and a scalar–valued function $p$ (pressure) that satisfy the relations

$$-\nu\Delta u = f - \nabla p \quad \text{in } \Omega, \tag{7.1}$$
$$\text{div} u = 0 \quad \text{in } \Omega, \tag{7.2}$$
$$u = u_0 \quad \text{on } \partial\Omega, \tag{7.3}$$

where $u_0$ is a given function such that $\text{div} u_0 = 0$. 

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
Nomenclature

Let smooth solenoidal functions with compact supports in $\Omega$ form the set be denoted by $\dot{J}^\infty(\Omega)$. The closure of $\dot{J}^\infty(\Omega)$ with respect to the norm $\|\nabla v\|$ is the space $\dot{J}^{\frac{1}{2}}(\Omega)$.

Next, $W := W^1_2(\Omega, \mathbb{R}^d)$ and $\Sigma := L_2(\Omega, \mathbb{M}^{d \times d})$, where $\mathbb{M}^{d \times d}$ is the space of symmetric $d \times d$ matrixes (tensors), whose scalar product is denoted by two dots. $W_0$ is a subspace of $W$ that contains functions with zero traces on $\partial \Omega$.

$W_0 + u_0$ contains functions of the form $w + u_0$, where $w \in V_0$. Analogously, $\dot{J}^{\frac{1}{2}}(\Omega) + u_0$ contains functions of the form $w + u_0$, $w \in \dot{J}^{\frac{1}{2}}(\Omega)$.

The operator $\varepsilon(v) := \frac{1}{2} (\nabla v + (\nabla v)^T)$ acts from $W$ to $\Sigma$. 

S. Repin

LECTURES ON A POSTERIORI ERROR CONTROL
We will also use the Hilbert space $\Sigma_{\text{div}}(\Omega)$, which is a subspace of $\Sigma$ that contains tensor–valued functions $\tau$, such that $\text{div}\tau \in L^2$. The scalar product in this space is defined by the relation

$$(\tau, \eta) := \int_{\Omega} (\tau : \eta + \text{div}\tau \cdot \text{div}\eta) \, dx.$$ 

By $L^2(\Omega)$ we denote the space of square summable functions with zero mean. Henceforth, we assume that

$$f \in L^2(\Omega, \mathbb{R}^d), \quad u_0 \in W^1_2(\Omega, \mathbb{R}^d),$$
Generalized solution can be defined by the integral identity. It is a function $u \in \mathring{\mathbf{H}}^1_2(\Omega) + u_0$ that meets the relation

$$\int_{\Omega} \nu \nabla(u) : \nabla(v) \, dx = \int_{\Omega} f \cdot v \, dx \quad \forall v \in \mathring{\mathbf{H}}^1_2(\Omega).$$  \hspace{1cm} (7.4)

It is well known that $u$ exists and unique and can be viewed as the minimizer of the functional

$$I(v) = \int_{\Omega} \left( \frac{\nu}{2} |\nabla(v)|^2 - f \cdot v \right) \, dx$$

on the set $\mathring{\mathbf{H}}^1_2(\Omega) + u_0$. Thus, the problem

$$\inf_{v \in \mathring{\mathbf{H}}^1_2(\Omega) + u_0} I(v)$$

presents a variational formulation of the Stokes problem.
Existence of a minimizer follows from known properties of convex lower semicontinuous functionals. In addition, the Stokes problem can be presented in a minimax form.

Let \( L : (W_0 + u_0) \times \hat{L}_2(\Omega) \to \mathbb{R} \) be defined as follows:

\[
L(v, q) = \int_{\Omega} \left( \frac{\nu}{2} |\nabla v|^2 - f \cdot v - q \text{div} v \right) \, dx.
\]

Now, \( u \) and \( p \) are defined as a saddle-point that satisfies the relations

\[
L(u, q) \leq L(u, p) \leq L(v, p) \quad \forall v \in W_0 + u_0, \ p \in \hat{L}_2(\Omega).
\]
First, we recall some basic results that has been established when the solvability of the Stokes problem was investigated. Works of O.A. Ladyzhenskaya made a grate contribution to the mathematical theory of viscous incompressible fluids.

**Lemma 1.**

For any vector–valued function $a \in W^{1/2}_{2}(\partial \Omega)$ satisfying the condition $\int_{\partial \Omega} a \cdot \nu \, dx = 0$ there exists a function $\bar{u} \in W_{0}$ such that $\text{div}\bar{u} = 0$ and

$$\|\nabla \bar{u}\| \leq \kappa_{1}(\Omega)\|a\|_{1/2,\partial \Omega},$$

(7.5)

where $\kappa_{1}(\Omega)$ is a positive constant that depends on $\Omega$. 

---

S. Repin


**LECTURES ON A POSTERIORI ERROR CONTROL**
This lemma implies another proposition, which is of great importance for the analysis of problems defined on solenoidal fields.

**Lemma 2**

For any $f \in L^2(\Omega)$ there exists a function $\bar{u} \in W_0$ satisfying the relation $\text{div} \bar{u} = f$ and the condition

$$\|\nabla \bar{u}\| \leq \kappa_2(\Omega) \|f\|,$$

(7.6)

where $\kappa_2(\Omega)$ is a positive constant that depends on $\Omega$.

Lemma 2 implies several important corollaries that we discuss below.
Inf-Sup condition

Lemma 2 is related to the inequality known in the literature as the Inf-Sup or LBB (Ladyzhenskaya–Babuška–Brezzi)–condition that reads: there exists a positive constant $C_\Omega$ such that

$$\inf_{\phi \in L_2(\Omega), \phi \neq 0} \sup_{w \in W_0, w \neq 0} \frac{\int_{\Omega} \phi \text{div} w \, dx}{\|\phi\| \|\nabla w\|} \geq C_\Omega. \quad (7.7)$$

S. Repin

LECTURES ON A POSTERIORI ERROR CONTROL

RICAM, Special Radon Semester, Linz, 2005.
Lemma 2 implies LBB condition

By Lemma 2, any $\phi \in L_2(\Omega)$ has a counterpart function $v_\phi \in W_0$ that meets the conditions

$$\text{div} v_\phi = \phi, \quad \|\nabla v_\phi\| \leq \kappa_2(\Omega) \|\phi\|.$$

In this case,

$$\sup_{v \in W_0, \|v\| \neq 0} \frac{\int_\Omega \phi \text{div} v \, dx}{\|\nabla v\| \|\phi\|} \geq \frac{\int_\Omega \phi \text{div} v_\phi \, dx}{\|\nabla v_\phi\| \|\phi\|} = \frac{\|\phi\|}{\|\nabla v_\phi\|} \geq \frac{1}{\kappa_2(\Omega)}$$

and, consequently, Inf-Sup condition holds with

$$C_\Omega = \frac{1}{\kappa_2(\Omega)}.$$
It is easy to observe that the Inf-Sup condition can be presented in the form

\[ \sup_{w \in W_0 \setminus \{0\}} \frac{\int_{\Omega} p \, \text{div} w \, dx}{\| \nabla w \|} \geq C_\Omega \| p \| \quad \text{for all } p \in L^2(\Omega). \]

We may consider the expression in the left–hand side of the above inequality as the norm of \( \nabla p \) in the space topologically dual to \( W_0 \), namely

\[ \left[ \nabla p \right] := \sup_{w \in W_0} \frac{\langle \nabla p, w \rangle}{\| \nabla w \|}. \]

Then, we arrive to the Nečas inequality.
Nečas inequality

\[ \left\| p \right\| \leq \kappa_2 \left\| \nabla p \right\| \quad \forall \ p \in H^1_0(\Omega), \]  

(7.8)

In the later paper, it is also shown that the well–known Korn’s inequality follows from Inf-Sup condition. Constants $C_\Omega$ and $\kappa_2$ play an important role in the numerical analysis of the Stokes problem as well as in the theoretical one.
Existence of a saddle point

Existence of a saddle point of $L(v, q)$ follows from Lemma 2 and known results of the minimax theory. In a simplified version these results reads:

Lagrangian $L(v, q)$ possess a saddle point provided that
(a) it is convex and continuous with respect to the first variable and concave and continuous with respect to the second one;
(b) for a certain $\bar{q}$ the functional $v \mapsto -L(v, \bar{q})$ is coercive (or the set of admissible $v$ is compact);
(c) or a certain $\bar{v}$ the functional $q \mapsto -L(\bar{v}, q)$ is coercive (or the set of admissible $q$ is compact.)
Since

\[ J(v) = \sup_{q \in \Sigma} L(v, q) \geq L(\bar{q}, v), \]

we observe that (b) means that \( J(v) \) is coercive. Analogously, (c) means that the functional \( -I(q) \), where

\[ I(q) = \inf_{q \in V_0 + u_0} L(v, q) \leq L(q, \bar{v}), \]

is coercive.

In other words, for a continuous convex-concave Lagrangian existence of a saddle point mainly depends on the coercivity properties of the two dual functionals generated by it.
Let us apply these results to the Stokes problem. It is easy to see that for any \( q \in L_2(\Omega) \) the mapping

\[
L(v, q) = \int_{\Omega} \left( \frac{\nu}{2} |\nabla v|^2 - f \cdot v - q \text{div} v \right) \, dx.
\]

is convex and continuous (in \( W \)) and there exists an element \( \bar{q} \in L_2(\Omega) \) (e.g., \( \bar{q} = 0 \)) such that \( L(v, \bar{q}) \to +\infty \) if \( \|v\|_V \to +\infty \).

The mapping \( q \mapsto L(v, q) \) is affine and continuous (in \( L_2(\Omega) \)) for any \( v \in V \). Therefore, existence of a saddle point is guaranteed provided that the coercivity condition

\[
\lim_{\|q\| \to +\infty} \inf_{v \in W_0 + u_0} L(v, q) = -\infty \tag{7.9}
\]

is established. By Lemma 2 we can prove this fact.
Consider the functional

\[ I(q) := \inf_{v \in W_0 + u_0} L(v, q) \]

and the variational problem

\[ I(p) = \sup_{q \in L_2(\Omega)} I(q) \quad (7.10) \]

for the pressure function. Note that the functional \( I \) has no explicit integral-type form and is defined as a supremum–functional. The solvability of this problem follows from the coercivity condition (7.9). To prove (7.9) we apply Lemma 2.
Coercivity of the variational problem for the pressure function

Indeed, by Lemma 2 for any \( q \in L_2(\Omega) \) we find \( v_q \in W_0 \) such that

\[
\text{div} v_q = q \quad \text{and} \quad \| \nabla v_q \| \leq \kappa_2 \| q \|.
\]

Take \( v = \mu v_q + u_0 \) and recall that \( \text{div} u_0 = 0 \). Then,

\[
\inf_{v \in W_0 + u_0} L(v, q) \leq \int_{\Omega} \left( \frac{\nu}{2} |\nabla (\mu v_q + u_0)|^2 - f \cdot (\mu v_q + u_0) - q \text{div} (\mu v_q + u_0) \right) dx \leq \\
\leq \int_{\Omega} \left( \frac{\nu}{2} |\nabla u_0|^2 - f \cdot u_0 \right) dx + \mu (\nu \| \nabla u_0 \| + C_\Omega \| f \|) \| \nabla v_q \| + \\
+ \frac{\nu \mu^2}{2} \| \nabla v_q \|^2 - \mu \| q \|^2 \leq \int_{\Omega} \left( \frac{\nu}{2} |\nabla u_0|^2 - f \cdot u_0 \right) dx + \\
+ \mu (\nu \| \nabla u_0 \| + C_\Omega \| f \|) \kappa_2 \| q \| + \mu \left( \frac{\nu \mu \kappa_2^2}{2} - 1 \right) \| q \|^2,
\]

where \( C_\Omega \) is a constant in the Friederichs inequality.

S. Repin

LECTURES ON A POSTERIORI ERROR CONTROL
We see that

\[
L(q) \leq c_1(u_0, f, \nu) + \mu(\nu \|\nabla u_0\| + C_\Omega \|f\|) \kappa_2 \|q\| + \\
+ \mu \left( \frac{\nu \mu \kappa_2^2}{2} - 1 \right) \|q\|^2.
\]

Set here \( \mu = \frac{1}{\nu \kappa_2^2} \). Then

\[
\inf_{v \in W_0 + u_0} L(v, q) \leq c_1 + c_2 \|q\| - \frac{1}{2 \nu \kappa_2^2} \|q\|^2 \to -\infty \quad \text{as} \quad \|q\| \to +\infty.
\]

Thus, we observe that the constant \( \kappa_2 \) arises in the quadratic term that provides the required coercivity property of the pressure functional.
Estimates of the distance to the set of solenoidal fields

Now we are concerned with the estimates of the distance between a function \( \hat{v} \in H^1 \) and the space of solenoidal functions.

**Estimates in L\(_2\)-norm.** An estimate of the distance between \( \hat{v} \) and the space

\[
J_2^1(\Omega) := \left\{ v \in W_2^1(\Omega) \mid \text{div} v = 0 \right\}
\]

in L\(_2\)-norm follow from the solvability of the Dirichlet problem for the Laplace operator. It is as follows:

\[
\inf_{v_0 \in J_2^1} \| \hat{v} - v_0 \| \leq C_F \| \text{div} \hat{v} \|,
\]

where \( C_F \) is the constant in the Friederichs inequality.
Proof. Indeed, since the problem

$$\Delta \phi = f,$$

has a solution $\phi \in \overset{\circ}{W}^1_2(\Omega)$ for any $f \in L^2(\Omega)$, we conclude that for any $f$ there exists $v_f = \nabla \phi$ such that

$$\text{div} v_f = f \quad \text{and} \quad \|v_f\| \leq C_F \|f\|.$$

Set $f = \text{div} \hat{v}$. Then,

$$\text{div}(v_f - \hat{v}) = 0,$$

so that $v_0 = v_f - \hat{v}$ belongs to $J^1_2$ and we observe that

$$\|\hat{v} - v_0\| \leq C_F \|\text{div} \hat{v}\|.$$
Estimates in $H^1$-norm. Let now $\hat{v} \in W_0$. Set $f = \text{div}\hat{v}$. Since 

$$\int_{\Omega} \text{div}\hat{v} \, dx = \int_{\partial\Omega} v \cdot \nu \, ds = 0,$$

we see that $f \in \mathcal{L}_2(\Omega)$. Then, by Lemma 2, one can find $u_f \in W_0$ such that 

$$\text{div}u_f = \text{div}\hat{v}, \quad \text{and} \quad \|\nabla u_f\| \leq \kappa_2(\Omega)\|\text{div}\hat{v}\|.$$

In other words, there exists a solenoidal field $w_0 = (\hat{v} - u_f) \in W_0$ such that 

$$\|\nabla (\hat{v} - w_0)\| = \|\nabla u_f\| \leq \kappa_2(\Omega)\|\text{div}\hat{v}\|. $$

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
This fact can be presented in another form

\[ \inf_{\mathbf{v} \in \mathring{J}_2(\Omega)} \| \nabla (\mathbf{\hat{v}} - \mathbf{v}) \| \leq \kappa_2(\Omega) \| \text{div} \mathbf{\hat{v}} \|. \quad (7.11) \]

Thus, for the functions with zero traces the distance to \( \mathring{J}_2(\Omega) \) in a strong norm is also measured via \( \| \text{div} \mathbf{\hat{v}} \| \), but with a different factor: \( \kappa_2(\Omega) \).
Comments on the value of $C_\Omega$

Note that $C_\Omega$ can be estimated throughout the constant $C_F$ and the constant $C_P$ in the Poincare inequality. Indeed,

$$C_\Omega = \inf_{q \in L^2, \ q \neq 0} \mathcal{E}(q),$$

$$\mathcal{E}(q) = \sup_{w \in W_0, \ w \neq 0} \frac{\int_{\Omega} q \cdot \nabla w \, dx}{\|q\| \|\nabla w\|}.$$ 

For $q \in \widetilde{W}(\Omega) := L^2 \cap W^{1,2}_0(\Omega)$ we have

$$\mathcal{E}(q) = \sup_{w \in W_0, \ w \neq 0} \frac{\int_{\Omega} \nabla q \cdot w \, dx}{\|q\| \|\nabla w\|} \leq \frac{\|\nabla q\|}{\|q\|} \sup_{w \in W_0, \ w \neq 0} \frac{\|w\|}{\|\nabla w\|} \leq C_F \frac{\|\nabla q\|}{\|q\|}.$$
Let $C_P$ be the smallest constant in the inequality

$$
\|q\| \leq C_P \|\nabla q\|, \quad q \in \tilde{W}(\Omega),
$$

i.e.,

$$
\inf_{q \in \tilde{W}(\Omega), q \neq 0} \frac{\|\nabla q\|}{\|q\|} = \frac{1}{C_P}.
$$

Then

$$
C_\Omega = \inf_{q \in L_2, q \neq 0} \mathcal{E}(q) \leq \inf_{q \in \tilde{W}(\Omega), q \neq 0} \mathcal{E}(q) \leq \frac{C_F}{C_P}.
$$
LBB-condition can be written in the form

$$\|p\| \leq C_\Omega^{-1} \int \nabla p \, \text{d}x \quad \forall \, p \in L_2,$$

what amounts

$$C_\Omega \leq \frac{\int \nabla p \, \text{d}x}{\|p\|}$$

we see the meaning of this constant: $C_\Omega$ is the infimum of $H^{-1}$ norms of functions such that $\|p\| = 1$ and $\int_\Omega p \, \text{d}x = 0$.

**Proposition 1**

If $\Omega \in \mathbb{R}^n$ then

$$\frac{\|\nabla p\|_{(-1)}}{\|p\|} \leq n \quad \forall p \in L_2(\Omega).$$
Proof.

\[
\sup_{w \in W_0, \ w \neq 0} \int_\Omega \frac{p \text{ div} w \ dx}{\| \nabla w \|} = \sum_{t=1}^n \int_\Omega \frac{p \ w_{t,t} \ dx}{\| \nabla w \|} \leq \sum_{t=1}^n \sup_{w_t \in W_0, \ w_t \neq 0} \int_\Omega \frac{p \ w_{t,t} \ dx}{\| \nabla w \|}.
\]
Since

\[ \| \nabla w \|^2 = \int_{\Omega} \left( \sum_{t,s=1}^{n} w_{t,s}^2 \right) dx \geq \int_{\Omega} w_{t,t}^2 dx \quad \forall t = 1, 2, \ldots n \]

we have

\[ \sup_{w \in W_0, w \neq 0} \int_{\Omega} p \, \text{div} w \, dx \leq \sum_{t=1}^{n} \sup_{w_t \in W_0, w_t \neq 0} \int_{\Omega} p w_{t,t} \, dx \leq \sum_{t=1}^{n} \sup_{\eta \in L_2, \eta \neq 0} \int_{\Omega} p \eta \, dx = \sum_{t=1}^{n} \| p \| = n \| p \|. \]
Proposition 2

If \( n = 1 \) then \( C_\Omega = 1 \).

Let \( \Omega = (a, b) \). Due to Proposition 1 we see that \( C_\Omega \leq 1 \). Let \( p \) be an arbitrary function from the set \( L_2 \). Then, the function

\[
\mathbf{w}^{(p)} = \int_a^x p \, dx \in W_0.
\]

Really, \( \mathbf{w}^{(p)}(a) = 0 \), \( \mathbf{w}^{(p)}(b) = \int_a^b p \, dx = 0 \) and \( \mathbf{w}^{(p)'} = p \in L_2(a, b) \). Thus,

\[
\sup_{\mathbf{w} \in W_0, \mathbf{w} \neq 0} \frac{\int_\Omega p \, \mathbf{w}' \, dx}{\| \mathbf{w}' \|} \geq \frac{\int_\Omega p \, \mathbf{w}^{(p)'} \, dx}{\| \mathbf{w}^{(p)'} \|} = \frac{\int_\Omega p^2 \, dx}{\| p \|} = \| p \|.
\]

Thus, \( C_\Omega \geq 1 \) and we arrive at the required result.
These estimates give a certain presentation on the value of $C_{\Omega}$. However, we are mainly interested in the estimate from below, what imposes a task more complicated than the finding the constant in the Friederichs inequality.

In principle, one could determine $C_{\Omega}$ by the following arguments.

Let $w_p \in W_0$ be a function such that

$$\Delta w_p = \nabla p, \quad w_p = 0 \text{ on } \partial \Omega.$$ 

Then,

$$-\int_{\Omega} \nabla w_p : \nabla v \, dx = \int_{\Omega} \nabla p \cdot v \, dx \quad \forall v \in W_0$$

and, thus, we have

$$\int_{\Omega} |\nabla w_p|^2 \, dx = \int_{\Omega} p \, \text{div} w_p \, dx \quad \forall v \in W_0.$$
Therefore,

\[
C_\Omega := \inf_{p \in L_2 \setminus \{0\}} \sup_{w \in W_0 \setminus \{0\}} \frac{\int_{\Omega} p \text{div} w \, dx}{\|p\| \|\nabla w\|} \geq \inf_{p \in L_2 \setminus \{0\}} \frac{\int_{\Omega} p \text{div} w_p \, dx}{\|p\| \|\nabla w_p\|} = \inf_{p \in L_2 \setminus \{0\}} \frac{\|\nabla w_p\|}{\|p\|}.
\]

Thus, finding \( C_\Omega \) requires the minimization of this quotient with respect to all \( p \in L_2 \setminus \{0\} \), where \( w_p \) is taken as the solution of the above defined linear problem. Certainly, such a task (for some \( \Omega \)) might be solved by only analytical methods. However, the minimization on a subspace of \( L_2 \setminus \{0\} \) may give a presentation on the value of \( C_\Omega \).
The value of $C_{LBB}$ is known for several model domains:

- Rectangular domain $(0, 1) \times (0, L)$, \quad $L \geq 1$
  
  see G. Stoyan, M. Olshanskij, E. Chizhonkov

$$\frac{\sin \frac{\pi}{8}}{L} \leq C_{LBB} \leq \frac{\pi}{2\sqrt{3}L}$$

- unitary disc with radius 1
  
  see L. Halpern

$$C_{LBB} = \frac{1}{\sqrt{2}}$$

Concerning numerical computation of $C_{LBB}$ see the works of G. Stoyan, M. Olshanskij, E. Chizhonkov
On $\mathbb{C}_\Omega$ for the square domain

Let

$$\Omega = \mathbb{Q} := \{ x \in \mathbb{R}^n \mid x_i \in (-\pi, \pi), \ i = 1, 2, \ldots, n \}.$$  

We are interested in the value of the quotient

$$\inf_{p \in \mathbb{L}_2} \frac{\| \nabla p \|}{\| p \|_Q}.$$ 

Represent $p$ as a series with respect to the trial functions

$$p_{ij}^{(1)} = \sin ix \sin jy, \quad p_{ij}^{(2)} = \sin ix \cos jy,$$

$$p_{ij}^{(3)} = \cos ix \sin jy, \quad p_{ij}^{(4)} = \cos ix \cos jy,$$

where $i, j = 0, 1, 2, \ldots$. Then

$$p(x, y) = \sum_{i, j = 0}^{\infty} \sum_{s = 1}^{4} a_{ij}^{(s)} p_{ij}^{(s)}.$$  

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
Here, the first nonzero coefficients are

\[ a_{00}^{(4)} = \frac{1}{4\pi^2} \int_{\Omega} p \, dx\,dy , \]
\[ a_{i0}^{(2)} = \frac{1}{2\pi^2} \int_{\Omega} p \sin ix \, dx\,dy , \]
\[ a_{0j}^{(3)} = \frac{1}{2\pi^2} \int_{\Omega} p \sin jy \, dx\,dy , \]
\[ a_{i0}^{(4)} = \frac{1}{2\pi^2} \int_{\Omega} p \cos ix \, dx\,dy , \]
\[ a_{0j}^{(2)} = \frac{1}{2\pi^2} \int_{\Omega} p \cos jy \, dx\,dy , \]
Other coefficients are as follows:

\[
a_{ij}^{(1)} = \frac{1}{\pi^2} \int_{\Omega} p \sin ix \sin jy \, dxdy ,
\]

\[
a_{ij}^{(2)} = \frac{1}{\pi^2} \int_{\Omega} p \sin ix \cos jy \, dxdy ,
\]

\[
a_{ij}^{(3)} = \frac{1}{\pi^2} \int_{\Omega} p \cos ix \sin jy \, dxdy ,
\]

\[
a_{ij}^{(4)} = \frac{1}{\pi^2} \int_{\Omega} p \cos ix \cos jy \, dxdy .
\]

We have

\[
\|p\|_Q^2 = \pi^2 \sum_{i,j=0}^{\infty} \lambda_{ij} \left[ (a_{ij}^{(1)})^2 + (a_{ij}^{(2)})^2 + (a_{ij}^{(3)})^2 + (a_{ij}^{(4)})^2 \right],
\]

where \(\lambda_{00} = 0, \lambda_{01} = 2, \lambda_{10} = 2\) and \(\lambda_{ij} = 1\) for all \(i,j \geq 1\).
Let us take a finite number of elements in the Fourier series for $p$:

$$
p = \sum_{i,j=0}^{N} \sum_{s=1}^{4} a^{(s)}_{ij} p^{(s)}_{ij},
$$

where $a^{(s)}_{ij}$ are the above defined coefficients. Since

$$\|\nabla p\| = \sup_{v \in W_0} \int_{\Omega} p \text{div} v \, dx,$$

we need to introduce the system of trial functions in $W_0(Q)$. It is given by the system of eigenfunctions for the problem

$$\Delta w = \mu w \quad w|_{\partial Q} = 0.$$
This system is

$$\phi_{\alpha\beta} = \sin \frac{\alpha}{2}(x + \pi) \sin \frac{\beta}{2}(y + \pi).$$

In this case,

$$\phi_{\alpha\beta,1} = \frac{\alpha}{2} \cos \frac{\alpha}{2}(x + \pi) \sin \frac{\beta}{2}(y + \pi),$$

$$\phi_{\alpha\beta,2} = \frac{\beta}{2} \sin \frac{\alpha}{2}(x + \pi) \cos \frac{\beta}{2}(y + \pi).$$

Take a finite number $M$ of basic functions in the representation of $v$, namely we set

$$v = v^M = (v_1^M, v_2^M), \quad v_1^M = \sum_{\alpha,\beta=1}^{M} b_{\alpha\beta} \phi_{\alpha\beta}, \quad v_2^M = \sum_{\alpha,\beta=1}^{M} c_{\alpha\beta} \phi_{\alpha\beta}.$$

The set of all such functions we denote $W_0^M$. In this case, we can obtain a lower bound for the required norm. Really, we have
\[
\| \nabla p \|^{(M)} := \sup_{v^M \in W_0^M} \int_{\Omega} p \text{div} v^M \, dx \leq \| \nabla p \| = \sup_{v \in W_0} \int_{\Omega} p \text{div} v \, dx.
\]

Thus, we may hope to estimate the value of the quotient

\[
\inf_{p \in L_2} \frac{\| \nabla p \|}{\| p \|_Q}.
\]

by taking \(N, M \to +\infty, M = \kappa N\) \(\kappa\) is essentially larger than 1 (typically 8-20). Numerical results for different \(N\) are exposed below.
Minimizer $p_n$ for $n = 8, 36$ and $120$. 
In order to clarify the main ideas of our approach we rewrite the classical Stokes system in a somewhat different form:

\[
\begin{align*}
\text{div}\sigma &= \nabla p - f \quad \text{in } \Omega, \quad (7.12) \\
\text{div}u &= 0 \quad \text{in } \Omega, \quad (7.13) \\
\sigma &= \nu \nabla u \quad \text{in } \Omega, \quad (7.14) \\
u &= 0 \quad \text{on } \partial\Omega. \quad (7.15)
\end{align*}
\]

This system involves one additional variable \( \sigma \) that corresponds to the field of stresses. Now we may regard the Stokes problem as the problem of finding a triplet of functions \((u, \sigma, p)\).
Primal and Dual Problems

Functional formulations of the above problem are given in natural "energy" set for this velocity–stress-pressure setting, which is

\[ \mathcal{E} := \hat{H}^{1/2}(\Omega) \times \Sigma \times \hat{L}_2. \]

**Problem** \( \mathcal{P} \). Find \( u \in \hat{H}^{1/2}(\Omega) \) such that

\[ J(u) \leq J(v) \quad \text{for all } v \in \hat{H}^{1/2}(\Omega), \]

where

\[ J(v) = \int_{\Omega} \left( \frac{\nu}{2} |\nabla v|^2 - f \cdot v \right) \, dx. \]

We denote the exact lower bound of this problem by \( \inf \mathcal{P} \).
Let $\Sigma = L^2(\Omega, M^{n \times n})$ and $L : J^1_2(\Omega) \times \Sigma(\Omega) \to \mathbb{R}$ be the Lagrangian

$$L(v, \tau) = \int_\Omega \left( \tau : \nabla v - \frac{1}{2\nu} |\tau|^2 \right) \, dx - \int_\Omega fv \, dx$$

that together with Problem $\mathcal{P}$ generates the dual problem

$$\sup_{\tau \in \Sigma} \inf_{v \in J^1_2(\Omega)} L(v, \tau)$$

which is Problem $\mathcal{P}^*$: find $\sigma \in \Sigma_f(\Omega)$ such that

$$I^*(\sigma) = \sup_{\tau \in \Sigma_f(\Omega)} I^*(\tau), \quad I^*(\tau) = -\frac{1}{2\nu} \int_\Omega |\tau|^2 \, dx$$

where

$$\Sigma_f(\Omega) := \left\{ \tau \in \Sigma(\Omega) \mid \int_\Omega \tau : \nabla w \, dx = \int_\Omega fw \, dx \text{ for all } w \in J^1_2(\Omega) \right\}.$$
From the general theorems of convex analysis it follows

**Theorem (1)**

*There exists a unique minimizer $u$ of problem $\mathcal{P}$ and unique maximizer $\sigma$ of problem $\mathcal{P}^*$. These two functions meet the equalities*

\[
I^*(\sigma) = \sup \mathcal{P}^* = \inf \mathcal{P} = I(u),
\]

(7.16)

\[
\sigma = \nu \nabla u.
\]

(7.17)
Basic error estimate

The basic error relation for the Stokes problem is given by the following theorem (S. Repin, 2002).

Theorem (2)

For any \( v \in \mathcal{J}_2^1(\Omega) \) and any \( \tau_f \in \Sigma_f \), we have

\[
\nu \| \nabla (v - u) \|^2 + \frac{1}{\nu} \| \tau_f - \sigma \|^2 = 2 \left( J(v) - I^*(\tau_f) \right). \tag{7.18}
\]
Proof of Theorem 2

The minimizer $u$ of problem $\mathcal{P}$ satisfies the relation (??). Therefore, we obtain

$$J(v) - J(u) = \int_{\Omega} \left( \frac{\nu}{2} |\nabla v|^2 - \frac{\nu}{2} |\nabla u|^2 - f \cdot (v - u) \right) \, dx =$$

$$= \int_{\Omega} \left( \frac{\nu}{2} |\nabla (v - u)|^2 + \nu \nabla u : \nabla (v - u) - f \cdot (v - u) \right)$$

$$= \frac{\nu}{2} \int_{\Omega} |\nabla (v - u)|^2 \, dx \quad \text{for all } v \in J_1^2(\Omega).$$

Since $J(u) = \inf \mathcal{P}$, we conclude that

$$\frac{\nu}{2} \|\nabla (v - u)\|^2 = J(v) - \inf \mathcal{P} \quad \text{for all } v \in J_1^2(\Omega).$$
The next step is to derive a similar relation for the dual problem. For this purpose, we note that the maximizer $\sigma$ of problem $\mathcal{P}^*$ satisfies the relation

$$\int_{\Omega} \sigma : (\tau_f - \sigma) \, dx = 0 \quad \text{for all } \tau_f \in \Sigma_f(\Omega).$$

By virtue of this relation, we find that

$$\sup \mathcal{P}^* - l^*(\tau_f) = l^*(\sigma) - l^*(\tau_f) = \frac{1}{2\nu} \|\tau_f - \sigma_f\|^2 \quad \tau_f \in \Sigma_f(\Omega).$$

Since $\inf \mathcal{P} = \sup \mathcal{P}^*$ we sum the two equalities and obtain

$$\nu \|\nabla(v - u)\|^2 + \frac{1}{\nu} \|\tau_f - \sigma\|^2 = 2 \left( J(v) - l^*(\tau_f) \right).$$
Stokes problem is a particular case of the abstract problem we investigated in Lecture 5:

Find $u \in V_0 + u_0$ such that

$$(\mathcal{A}\Lambda u, \Lambda w) + \langle \ell, w \rangle = 0 \quad \forall w \in V_0.$$ 

In this case $V_0 = \overset{0}{\mathcal{J}}^1_2(\Omega)$, $V$ is a subspace of $H^1$ containing solenoidal fields, $\Lambda = \nabla$ (tensor-gradient), $U = \Sigma$, $\mathcal{A}y = \nu y$, and

$$\langle \ell, w \rangle = - \int_{\Omega} fw \, dx$$
Thus, we can apply the estimate

$$\frac{1}{2} \| \Lambda (v-u) \|^2 \leq (1 + \beta) D(\Lambda v, y) + \frac{1 + \beta}{2\beta} \| \ell + \Lambda^* y \|^2,$$

(7.19)

where $\| y \|^2 = \int_{\Omega} \nu |y|^2 dx$ and

$$\| \ell + \Lambda^* y \| = \sup_{w \in V_0} \frac{\langle \ell + \Lambda^* y, w \rangle}{\| \Lambda w \|} = \sup_{w \in J^1_2(\Omega)} \frac{\int_{\Omega} (\nabla w : y - fw) dx}{\| \nabla w \|} = \sup_{w \in J^1_2(\Omega)} \frac{\int_{\Omega} (\nabla w : y - fw - q \text{div} w) dx}{\| \nabla w \|} \leq \sup_{w \in H^1(\Omega)} \frac{\int_{\Omega} (\nabla w : y - fw - q \text{div} w) dx}{\| \nabla w \|} \forall q \in L^2(\Omega).$$
If
\[ y \in \Sigma_{\text{div}}(\Omega) := \{ y \in \Sigma \mid \text{div} y \in L^2(\Omega, \mathbb{R}^n) \} \]
and \( q \in H^1 \), we have
\[
\sup_{w \in H^1(\Omega)} \frac{\int_{\Omega} (\nabla w : y - fw - q\text{div}w)\,dx}{\| \nabla w \|} = \sup_{w \in H^1(\Omega)} \frac{\int_{\Omega} (f - \nabla q + \text{div} y) \cdot w\,dx}{\| \nabla w \|}
\]
Since
\[ \| w \| \leq C_\Omega \| \nabla w \| = C_\Omega \nu^{-1/2} \| \nabla w \| , \]
we obtain
\[ \ell + \Lambda^* y \leq C_\Omega \nu^{-1/2} \| f - \nabla q + \text{div} y \| \]
Further,

\[ D(\nabla v, y) = \int_{\Omega} \left( \frac{1}{2} \nu \nabla v : \nabla v + \frac{1}{2} \nu^{-1} y : y - \nabla v : y \right) \, dx = \]

\[ = \frac{1}{2\nu} \| y - \nu \nabla v \|^2. \]

Now, from (7.19) we obtain

\[ \frac{\nu}{2} \| \nabla (u - v) \|^2 \leq (1 + \beta) \frac{1}{2\nu} \| y - \nu \nabla v \|^2 + \frac{1 + \beta}{2\beta\nu} C_{\Omega}^2 \| f - \nabla q + \text{div} y \|^2, \]

or

\[ \nu^2 \| \nabla (u - v) \|^2 \leq (1 + \beta) \| y - \nu \nabla v \|^2 + \frac{1 + \beta}{\beta} C_{\Omega}^2 \| f - \nabla q + \text{div} y \|^2. \]
By the minimization with respect to $\beta$ we derive the first basic estimate for the Stokes problem:

$$\nu \| \nabla (u - v) \| \leq \| y - \nu \nabla v \| + C_\Omega \| f - \nabla q + \text{div} y \|. \quad (7.20)$$

Here $v$ is any conforming approximation of $u$ and $y$ is any tensor–function in $\Sigma_{\text{div}}(\Omega)$ and $q \in H^1$ is an ”image” of the pressure function.

Non-solenoidal approximations

If the function $\hat{v} \in V_0 + u_0$ does not satisfy the incompressibility condition, then the estimate of its deviation from $u$ can be obtained as follows.

By Lemma 2 for the function $\hat{v}_0 := \hat{v} - u_0$ one can find a function $w_0 \in J^{1/2}(\Omega)$ such that

$$\| \nabla (\hat{v}_0 - \hat{w}_0) \| \leq k(\Omega) \| \text{div} \hat{v}_0 \| .$$

Then,

$$\nu \| \nabla (u - \hat{v}) \| = \nu \| \nabla (u - \hat{v}_0 - u_0) \| \leq$$

$$\leq \nu \| \nabla (u - (\hat{w}_0 + u_0)) \| + \nu \| \nabla (\hat{v}_0 - \hat{w}_0) \| .$$

Use (7.20) to estimate the first norm in the right–hand side of this inequality.
We obtain

\[ \nu \| \nabla (u - \hat{v}) \| \leq \| \nu \nabla (\hat{w}_0 + u_0) - y \| + C_\Omega \| \text{div} y + f - \nabla q \| + \]
\[ + \nu \| \nabla (\hat{v}_0 - \hat{w}_0) \| \leq \| \nu \nabla \hat{v} - y \| + \]
\[ + C_\Omega \| \text{div} y + f - \nabla q \| + 2\nu \| \nabla (\hat{v}_0 - \hat{w}_0) \|. \]

Hence, we arrive at the estimate

\[ \nu \| \nabla (u - \hat{v}) \| \leq \| \nu \nabla (\hat{v}) - y \| + C_\Omega \| \text{div} y + f - \nabla q \| + \frac{2\nu}{C_\Omega} \| \text{div} \hat{v} \|. \tag{7.21} \]

Three terms in the right–hand side of the estimate present three natural parts of the error, namely errors in the constitutive law, differential equation and incompressibility condition.
Another form of the Majorant

Set \( y = \eta + qI \), where \( I \) is the unit tensor and \( \eta \in \Sigma_{\text{div}}(\Omega)(\Omega) \). Then the Majorant comes in the form

\[
\nu \| \nabla(u - \hat{v}) \| \leq \| \nu \nabla(\hat{v}) - \eta - qI \| + C_\Omega \| \text{div} \eta + f \| + \frac{2\nu}{C_\Omega} \| \text{div} \hat{v} \| \quad (7.22)
\]

Thus, if the constants \( c_\Omega \) and \( C_\Omega \) are known (or we know suitable upper bounds for them), then \((7.21)\) and \((7.22)\) provides a way of practical estimation the deviation of \( \hat{v} \) from \( u \).
Practical implementation

To use the above estimates in practice we should select certain finite dimensional subspaces

$$\Sigma_k \quad \text{and} \quad Q_k$$

for the functions $y$ (or $\eta$) and $q$, respectively. Minimization of the right-hand side of the estimates with respect to $y$ and $q$ gives an estimate of the deviation, which will be the sharper the greater is the dimensionality of the subspaces used. Numerical testing of the estimates has been performed in E. Gorshkova and S. Repin. Error control of the approximate solution to the Stokes equation using a posteriori error estimates of functional type. In *European Congress on Computational Methods in Applied Sciences and Engineering, ECCOMAS 2004*, Jyväskylä, 24-28 July, 2004 (electronic).
Estimates for the pressure field

Let \( q \in L_2 \) be an approximation of the pressure field \( p \). Then \( (p - q) \in L_2 \) and the Inf-Sup condition implies the relation

\[
\sup_{w \in V_0, \ w \neq 0} \frac{\int_{\Omega} (p - q) \text{div} w \, dx}{\|p - q\| \|\nabla w\|} \geq C_\Omega.
\]

Thus, for any small positive \( \epsilon \) there exists a nonzero function \( w_{pq}^\epsilon \in V_0 \) such that

\[
\int_{\Omega} (p - q) \text{div} w_{pq}^\epsilon \, dx \geq (C_\Omega - \epsilon)\|p - q\| \|\nabla w_{pq}^\epsilon\|.
\]
Since
\[ \int_{\Omega} \nu \nabla u : \nabla w_{pq}^\varepsilon \, dx = \int_{\Omega} \left( f \cdot w_{pq}^\varepsilon + p \, \text{div} w_{pq}^\varepsilon \right) \, dx, \]
we have
\[ \int_{\Omega} (p - q) \, \text{div} w_{pq}^\varepsilon \, dx = \]
\[ = \int_{\Omega} \left\{ \nu \nabla (u - \hat{v}) : \nabla w_{pq}^\varepsilon + \nu \nabla \hat{v} : \nabla w_{pq}^\varepsilon + \nabla q \cdot w_{pq}^\varepsilon - f \cdot w_{pq}^\varepsilon \right\} \, dx \]
\[ = \int_{\Omega} \nu \nabla (u - \hat{v}) : \nabla w_{pq}^\varepsilon \, dx + \int_{\Omega} (\nu \nabla \hat{v} - y : \nabla w_{pq}^\varepsilon) \, dx \]
\[ + \int_{\Omega} (y : \nabla w_{pq}^\varepsilon + \nabla q \cdot w_{pq}^\varepsilon - f \cdot w_{pq}^\varepsilon) \, dx, \]
where \( \hat{v} \) is an arbitrary function in \( W_0 + u_0 \) and \( y \) as an arbitrary tensor–valued function in \( \Sigma \).
Above relations lead to the estimates

\[ \|p - q\| \leq \frac{1}{(C_{\Omega} - \epsilon)\|\nabla w_{pq}\|} \times \left[ \int_{\Omega} (\nu \nabla(u - \hat{v}) : \nabla(w_{pq}) + (\nu \nabla(\hat{v}) - y) : \nabla(w_{pq})) \, dx \right. \]

\[ + \int_{\Omega} (-w_{pq} \cdot \text{div} y + \nabla q \cdot w_{pq} - f \cdot w_{pq}) \, dx \left. \right] \]

\[ \leq \frac{1}{(C_{\Omega} - \epsilon)} \left[ \nu \|\nabla(u - \hat{v})\| + \|\nu \nabla(\hat{v}) - y\| + C_{\Omega} \|\text{div} y + f - \nabla q\| \right]. \]

The first term in the right–hand side of this inequality is estimated by (7.21).
Deviation estimate for the pressure function

Since $\epsilon$ may be taken arbitrarily small, we obtain the following estimate for the deviation from the exact pressure field:

$$\frac{1}{2} \|p - q\| \leq \frac{\nu}{C_\Omega^2} \| \text{div}\hat{v} \| + \frac{1}{C_\Omega} \| \nu \nabla(\hat{v}) - y \| + \frac{C_\Omega}{C_\Omega} \| \text{div}y + f - \nabla q \|. \quad (7.23)$$

It is easy to see that the right-hand side of (7.23) consists of the same terms as the right-hand side of (7.21) and vanishes if and only if, $\hat{v} = u$, $y = \sigma$ and $p = q$. However, in this case, the dependence of the penalty multipliers from the constant $C_\Omega$ is stronger.
Problems with condition \( \text{div} u = \phi \).

In many cases, divergence–free condition is replaced by

\[
\text{div} u = \phi \quad \text{in } \Omega,
\]

where \( \phi \) is a given function in \( L_2 \). For such functions, we have the problem: find \( u \) that is equal to \( u_0 \) on \( \partial \Omega \) and

\[
-\text{div} \sigma + \nabla p = f \quad \text{in } \Omega,
\]

\[
\sigma = \nu \nabla u \quad \text{in } \Omega,
\]

Let \( u_\phi \in W_0, \text{div} u_\phi = \phi \). By setting \( u = \bar{u} + u_\phi \) and \( \bar{u}_0 = u_0 - u_\phi \), we present the boundary–value problem as follows: find \( \bar{u} \in J_2(\Omega) + \bar{u}_0 \) such that

\[
-\text{div} \tilde{\sigma} + \nabla \tilde{p} = \tilde{f} \quad \text{in } \Omega, \quad \tilde{f} = f + \nu \text{div} \nabla (u_\phi) \in H^{-1},
\]

\[
\tilde{\sigma} = \nu \nabla \bar{u} \quad \text{in } \Omega.
\]
Assume that $u$ is approximated by a certain $v \in V_0 + u_0$. Let $v$ be presented in the form $v = \bar{v} + u_\phi$. Now, we apply (7.21) to a "shifted" system and obtain

$$\|\nabla (u - v)\| = \|\nabla (\bar{u} - \bar{v})\| \leq$$

$$\leq \|\nu \nabla \bar{v} - y\| + \left[ \text{div} y + \bar{f} - \nabla q \right] + \frac{2\nu}{C_{\text{LBB}}} \|\text{div} \bar{v}\|.$$

Set here $y = -\nu \nabla u_\phi + \eta$, where $\eta$ is a function in $\Sigma$. Then

$$\text{div} y + \bar{f} = -\nu \text{div} \nabla u_\phi + \text{div} \eta + \bar{f} = \text{div} \eta + f$$

and $\nu \nabla \bar{v} - y = \nu \nabla (v - u_\phi) - y = \nu \nabla v - \eta$. Therefore,

$$\|\nabla (u - v)\| \leq$$

$$\leq \|\nu \nabla v - \eta\| + \left[ \text{div} \eta + f - \nabla q \right] + \frac{2\nu}{C_{\text{LBB}}} \|\text{div} v - \phi\|.$$
Problems for almost incompressible fluids

Models of almost incompressible fluids are often used for constructing sequences of functions converging to a solution of the Stokes problem. In this case, the incompressibility condition is replaced by the term that contains the divergence with a large multiplier. Let us consider a model of such a type.

We find $u_\delta \in V$ satisfying the integral identity

$$
\int_{\Omega} \left( \nu \nabla u_\delta : \nabla w + \frac{1}{\delta} \text{div} u_\delta \text{div} w \right) \, dx = \int_{\Omega} f \cdot w \, dx, \quad w \in W_0,
$$

and the boundary condition $u_\delta = u_0 \quad \partial \Omega$. It is not difficult to show (see, e.g., R. Temam [?]), that $u_\delta$ tends to $u$ (solution of the Stokes problem) in $H^1$ norm and $p_\delta = -\frac{1}{\delta} \text{div} u_\delta \in L^2$ converges to the respective pressure function $p$ in $L^2$ as $\delta \to 0$.
By (7.21) we can easily obtain an estimate of the difference between \( u \) and \( u_\delta \). Let us set in (7.21) \( y = \tau_\delta := \nu \nabla u_\delta \) and \( q = p_\delta = -\frac{1}{\delta} \text{div} u_\delta \). In this case, \( \| \nu \nabla u_\delta - \tau_\delta \| = 0 \) and

\[
\| \text{div} \tau_\delta + f - \nabla p_\delta \| = \\
= \sup_{w \in V_0} \int_{\Omega} \left( -\nu \nabla u_\delta : \nabla w + f \cdot w + p_\delta \text{div} w \right) dx \| \nabla w \| = 0.
\]

Thus, we conclude that

\[
\frac{1}{2} \| \nabla (u - u_\delta) \| \leq \frac{1}{C_{\text{LBB}}} \| \text{div} u_\delta \|,
\]

We observe that the deviation from the exact solution of the Stokes problem is controlled by the norm of the divergence of the regularized problem. Similar estimate can be obtained for the approximations constructed by means of the Uzawa algorithm.

*Algebra and Analiz (St.-Petersburg Math. J)*, 16(2004), 5

functional a posteriori estimates for the Stokes and some other problems were derived by *nonvariational* techniques. In particular, in this paper readers can find such estimates for

**Convection–diffusion equation**

\[-\text{div}A\nabla u + a \cdot \nabla u = f\]

and **Oseen problem**

\[-\nu \Delta u + \text{div}(a \otimes u) = f - \nabla p \quad \text{in } \Omega,\]

\[
\text{div}u = 0 \quad \text{in } \Omega, \\
u = 0 \quad \text{on } \partial \Omega.
\]
Generalizations

A posteriori estimates of the above discussed type can be derived in the abstract form for the whole class of problems where a solution is seeking in a subspace. Typically, we have the following diagram:

\[
\begin{array}{cccccc}
H & \leftrightarrow & W_0 & \xrightarrow{\Lambda} & U & (Y, Y^*) \\
& \downarrow & & & & \uparrow \\
H & \xrightarrow{B^*} & W^*_0 & \xleftarrow{\Lambda^*} & U \\
\end{array}
\]

**Basic problem.** Find \( p \in H \) and \( u \in V_0 \) that satisfy the relation

\[
(\mathcal{A}\Lambda u, \Lambda w) + \langle f - B^*p, w \rangle = 0 \quad \forall w \in W_0,
\]

where

\[
V_0 = \text{Ker}B := \{ v \in W_0 \mid Bv = 0 \}.
\]
Assume that

\[ \nu_1 \| y \|^2 \leq (A y, y) \leq \nu_2 \| y \|^2, \quad y \in U, \]

Let the operator \( B \) possesses the following property: there exists a constant \( \alpha \) such that for any

\[ g \in \text{Im } B := \{ z \in H \mid \exists v \in W_0 : Bv = z \} \]

one can find \( u_g \in W_0 \) such that

\[ Bu_g = g \quad \text{and} \quad \| u_g \|_W \leq \alpha \| g \|. \]

Note that such a condition is a generalization of Lemma 2.

Under the above assumption we obtain an estimate of the deviation from \( u \).
Estimate of the deviation from $u$

\[ \| \Lambda(u - \hat{v}) \| \leq \]
\[ \leq 2\sqrt{\nu_2}\alpha B\hat{v} + \| \Lambda\Lambda\hat{v} - y \|_* + \frac{1}{\sqrt{\nu_1}} \left[ f + \Lambda^*y - B^*q \right]. \]

where $\| y \| := (Ay, y)^{1/2}$, $\| y \|_* := (A^{-1}y, y)^{1/2}$ We see that the terms of the estimate present errors in the basic relations

\[
\left\{ \begin{array}{l}
\langle \Lambda^*\sigma + f - B^*p, w \rangle = 0 \quad \forall w \in V_0, \\
\sigma = A\Lambda u, \\
Bv = 0.
\end{array} \right.
\]
For the Stokes problem $\Lambda v = \nabla v$, $A = \nu I$, where $I$ denotes the identity operator and $Bv = -\operatorname{div}v$. It is easy to see that in this case $\nu_1 = \nu_2 = \nu$,

$$\| A\hat{v} - y \|_* = \frac{1}{\sqrt{\nu}} \| \nu \nabla v - y \|.$$

Since $\| \Lambda(u - \hat{v}) \| = \sqrt{\nu} \| \Lambda(u - \hat{v}) \|$, we find that the general estimate coincides with (7.21).
Literature comments.

A significant part of the difficulties arising in the process of solving such problems is related to the *incompressibility condition*. Typically, this condition is taken into account by projecting of a discrete solution to the set of solenoidal fields or by introducing appropriate penalty terms (see, e.g., A. Chorin [12], E. W. and J. G. Liu [14], V. Girault, P. A. Raviart [16], G. Heywood and R. Rannacher [17], R. Rannacher [25,26], J. Shen [32], R. Temam [33]). Stationary problems are often solved by passing to a minimax formulation and using the so-called mixed approximations for the velocity and pressure fields (see, e.g., F. Brezzi and J. Duglas [9], F. Brezzi and M. Fortin [10]).

A posteriori error estimates for approximations of the Stokes problem constructed by various types of finite element methods were obtained in numerous papers (mainly in the framework of certain modifications of the residual method (see, e.g., M. Aintworth and T. Oden [1], R. Bank and B. Welfert [5], E. Dari, R. Duran and C. Padra [13], C. Carstensen and S. Funken [11], G. Heywood and R. Rannacher [18], C. Johnson and R. Rannacher [19], C. Johnson, R. Rannacher and M. Boman [20], RICAM, Special Radon Semester, Linz, 2005. S. Repin, LECTURES ON A POSTERIORI ERROR CONTROL
J. Oden, W. Wu and M. Aintworth [24], R. Verfürth [35,36]). In these estimates, the right-hand side is given by the sum of local quantities $\eta_k$ that include additional terms that take into account violations of the incompressibility condition.

Functional type a posteriori error estimates for the Stokes problem were firstly derived by the variational techniques in [27]. Later, this method was applied to some other viscous flow problems with non-quadratic dissipative potentials [28]. In [30], another nonvariational techniques was used. It was shown that for the Stokes problem it leads to the same estimates. In [31] a posteriori estimates in local norms were derived.


Lecture 8.

ESTIMATION OF INDETERMINACY ERRORs.
Errors arising due to data indeterminacy

IN REAL LIFE PROBLEMS ALL THE DATA ARE INDETERMINATE!!!

Example 1.
Diffusion problem: find (temperature) $T$ such that

$$\text{div}k(x) \nabla T(x) + f = 0 \quad \text{in } \Omega$$

$$T(x) = T_0 \quad \text{on } \partial \Omega$$

In reality, the diffusion coefficient, temperature sources, and even the domain are not exactly known.
Example 2.
Stokes problem: find solenoidal $u$ such that

\[- \text{div} \sigma = f - \nabla p \quad \text{in } \Omega,\]
\[\nu \nabla u = \sigma \quad \text{in } \Omega,\]
\[u = u_0 \quad \text{on } \partial \Omega,\]

In Stokes, Oseen and Navier–Stokes equations the viscosity coefficient is never known exactly. Moreover, it is typically an unknown function (depending both on the spatial and time coordinates) that may depend on the temperature, contamination and other factors.
Uncertain data lead to a quite different analysis

Data uncertainty drastically changes some basic relations in the numerical analysis.
For example, let \( v \in C^2[a, b] \) and \( x, x + h \in [a, b] \). Since

\[
v(x + h) = v(x) + v'(x)h + v''(x + \theta h)\frac{h^2}{2}, \quad \theta \in (0, 1)
\]

we have the standard finite difference quotient approximation of the derivative

\[
v'(x) \approx \frac{v(x + h) - v(x)}{h}, \quad |e| \leq \frac{\mu h}{2}, \quad \mu = \max_{x \in [a, b]} |v''(x)|
\]
Assume now that $v(x)$ is defined with a certain indeterminacy, so that its real value is unknown and instead we have a function $\tilde{v}(x)$ whose values lie in the interval $[v(x) - \varepsilon, v(x) + \varepsilon]$. If we use these data to approximate the derivative, then we arrive at the following result:

$$
\left| v'(x) - \frac{\tilde{v}(x + h) - \tilde{v}(x)}{h} \right| \leq \left| v'(x) - \frac{v(x + h) - v(x)}{h} \right| + \frac{2\varepsilon}{h} \leq \mu \frac{h}{2} + \frac{2\varepsilon}{h}.
$$

We see that the error does not tend to zero as $h \to 0$. Moreover, 

$$
\min_{h} \left( \frac{\mu h}{2} + \frac{2\varepsilon}{h} \right)
$$

is attained at

$$
\tilde{h} = 2\sqrt{\frac{\varepsilon}{\mu}} \quad \text{(the best accuracy)}.
$$
Therefore, in the case of not fully determinate data the highest accuracy of the numerical differentiation is

$$|e_{\text{min}}| = 2\sqrt{\varepsilon \mu}$$

For example, if $\varepsilon = 10^{-4}$ and $\mu = 9$, then the highest accuracy is

$$|e_{\text{min}}| = 6 \times 10^{-2} \sim 10^{-1} !!!$$

and it is attained for $h \approx 0.01$.  

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
Errors in coupled problems

Effects close to those arising as a result of data indeterminacy often appear in the process of numerical simulation of **coupled systems** where certain quantities in a differential problem are defined throughout solutions of some other problems. In such systems a phenomenon of "error multiplication" may lead to a dramatic loss of the accuracy. An example below demonstrates such type effects.
"Baby" coupled problem. Find \( z(8) \), where \( z \) is the solution of the problem

\[
z'' - 9z' - 10z = 0, \quad z = z(x), \quad x \in [0, 8],
\]

\[
z(0) = 1, \quad z'(0) = a_{N-1} - a_N,
\]

where \( a \) is a solution of the system of the dimensionality \( N \)

\[
Ba = f, \quad b_{ij} = \frac{2S_i^2S_j^2}{\pi} \int_0^\pi \left( \sin(i\xi) \sin(j\xi) + \sin(i+j^2)\xi \right) d\xi,
\]

\[
i, j = 1, 2, \ldots N, \quad f_i = (i + 1)^4i, \quad S_i = \sum_{k=0}^{+\infty} \left( \frac{i}{i+1} \right)^k.
\]

For \( N = 10, 50, 100, 200 \) find \( z(8) \) analytically and compare with the result obtained by "purely numerical" approach in which sums, integrals and ODE are treated numerically.
Solution.

\[ S_i = \sum_{k=0}^{+\infty} \left( \frac{i}{i + 1} \right)^k = \sum_{k=0}^{+\infty} q^k = \frac{1}{1 - q} = i + 1 \]

\[ \sin(i \xi) \sin(j \xi) = \frac{1}{2} (\cos(i - j) \xi - \cos(i + j) \xi), \]

\[ \int_{0}^{\pi} \cos((i + j) \xi) d\xi = \frac{1}{i + j} \sin(i + j) \bigg|_{0}^{\pi} = 0, \]

\[ \int_{0}^{\pi} (\cos(i - j) \xi) d\xi = \pi \text{ if } i = j \text{ and } = 0 \text{ otherwise.} \]

Therefore,

\[ \int_{0}^{\pi} \sin(i \xi) \sin(j \xi) d\xi = \frac{\pi}{2} \text{ if } i = j \text{ and } = 0 \text{ otherwise.} \]
Thus, $B$ is a diagonal matrix with

$$b_{ii} = S_i^4 = (i + 1)^4$$

and

$$Ba = f \quad \text{is} \quad (i + 1)^4 a_i = (i + 1)^4 i \quad \Rightarrow \quad a_i = i.$$  

Solution of the equation we find in the form $z = e^{\lambda x}$, where $\lambda$ is a root of

$$\lambda^2 - 9\lambda - 10 = 0, \quad \lambda_1 = -1, \quad \lambda_2 = 10.$$ 

We have

$$z = C_1 e^{-x} + C_2 e^{10x}, \quad z(0) = C_1 + C_2 = 1$$

$$z' = -C_1 e^{-x} + 10C_2 e^{10x}, \quad z'(0) = -C_1 + 10C_2 = a_{N-1} - a_N = -1$$

From here, $C_1 = 1$ and $C_2 = 0$, so that

$$z = e^{-x}, \quad z(8) \approx 3.3546262 \times 10^{-4}$$
Principal question

Assume we have an approximation $u_h$ computed on a mesh $\mathcal{T}_h$. The question to be answered is as follows:

WHICH ERROR: APPROXIMATION or INDETERMINACY IS BIGGER?

If

Indeterminacy error $> \approx$ Approximation error

then all further computations and mesh adaptations are senseless!

We need a practical way to explicitly evaluate errors caused by indeterminacy in the problem data.
General framework

Consider the problem

$$\Lambda^* A \Lambda u = \ell \text{ in } \Omega, \quad u = u_0 \text{ on } \partial \Omega.$$  \hspace{1cm} (8.1)

where the operator $A$ and the functional $\ell$ are defined with some indeterminacy. It means that

$$A \in \mathcal{U}_A \subset \mathcal{L}(U, U),$$

$$\ell \in \mathcal{U}_\ell \subset V_0^*,$$

where $\mathcal{U}_A$ and $\mathcal{U}_\ell$ are certain bounded sets.
All possible solutions of the problem with such a data form the set

\[ \Upsilon(U_A, U_\ell) := \left\{ \tilde{u} \in V_0 + u_0 \mid \tilde{u} \text{ satisfies (8.1) for some } A \in U_A \ell \in U_\ell \right\}. \]

Let \( v \in V_0 + u_0 \) be an approximation of an unknown exact solution. Since the data are indeterminate, the error estimation problem comes in two different forms. The first problem is to find the quantity

\[ e_{\text{min}}^2(v, \Upsilon) = \frac{1}{2} \inf_{\tilde{u} \in \Upsilon} \| \Lambda(v - \tilde{u}) \|^2, \quad (8.2) \]

The quantity \( e_{\text{min}} \) measures the distance between \( v \) and the set \( \Upsilon \). It is equal to zero if \( v \) satisfies (8.1) for some pair \((A, \ell) \in U_A \times U_\ell\). This quantity provides the lowest possible bound of the true error or the error in the best-case situation.
Another task is to find the quantity

\[ e_{\text{max}}^2(v, \gamma) = \frac{1}{2} \sup_{\tilde{u} \in \gamma} \| \Lambda (v - \tilde{u}) \|^2, \]  

which shows the highest possible error. It takes into account computational errors and errors caused by indeterminacy and shows the error in the worst-case situation when the exact solution is an element of \( \gamma \) that is most distant of \( v \).

This quantity is always positive and its value gives an idea of the accuracy limit dictated by the effect of indeterminacy in the data.
Thus,

\[ e_{\text{min}}(v, \mathcal{Y}) \leq e(v) \leq e_{\text{max}}(v, \mathcal{Y}), \quad (8.4) \]

where the actual error \( e(v) \) is \textit{principally unknown} and we may only hope to find its bounds. In general, the exact values of \( e_{\text{min}} \) and \( e_{\text{max}} \) could hardly be found. However, using functional type \textit{a posteriori} estimates, one can find their computable bounds. Indeed, the majorant \( M_\oplus \) and the minorant \( M_\ominus \) explicitly depend on \( \mathcal{A} \) and \( \ell \), which opens a way for computing errors caused by the indeterminacy in values of the problem data. Below we show how such an account can be performed.

Assume that the set \( \mathcal{Y} \) is known. Our aim is to find \textit{practically computable} numbers \( e_\ominus(v, \mathcal{Y}) \) and \( e_\oplus(v, \mathcal{Y}) \) such that for any \( v \in V_0 + u_0 \) the following relations hold:

\[ e_\ominus(v, \mathcal{Y}) \leq e_{\text{min}}(v, \mathcal{Y}) \leq e_{\text{max}}(v, \mathcal{Y}) \leq e_\oplus(v, \mathcal{Y}). \quad (8.5) \]
Consider the generalized diffusion problem

\[ \text{div} A \nabla u + f = 0, \]

In this case,

\[ V = H^1(\Omega), \quad Y = L^2(\Omega, \mathbb{R}^n), \quad V_0 := H^1(\Omega), \]
\[ V_0^* = H^{-1}(\Omega), \quad \Lambda v = \nabla v, \quad \Lambda^* y^* = -\text{div} y^*, \]

and \( \mathcal{A} \) is a mapping given by the relation \( y^*(x) \rightarrow A(x)y^*(x) \), where \( A(x) \) is a symmetric positive definite matrix.
Assume the coefficients of the differential equation are defined by some "mean" elements

\[ A_0 \in L^\infty(\Omega; \mathbb{M}^{n \times n}) \quad \text{and} \quad f_0 \in L^2(\Omega) \]

and certain (bounded) variations around these values.

\[ \mathcal{U}_A := \{ A \in L^\infty(\Omega; \mathbb{M}^{n \times n}) \mid A = A_0 + \varepsilon E, \ E \in \mathcal{E} \} , \]

\[ \mathcal{U}_f := \{ f \in L^2(\Omega) \mid f = f_0 + \delta \varphi, \ \varphi \in \mathcal{F} \} , \]

where

\[ \mathcal{E} := \{ E \in L^\infty(\Omega; \mathbb{M}^{n \times n}) \mid \| E \|_{\infty,\Omega} \leq 1 \} , \]

\[ \mathcal{F} := \{ \varphi \in L^2(\Omega) \mid \| \varphi \|_{2,\Omega} \leq 1 \} . \]

We will define the influence of the above indeterminacy errors. Our analysis follows the lines of S. Repin. A posteriori error estimates taking into account indeterminacy of the problem data. *Russian J. Numer. Anal. Anal.*
Let, $\varepsilon$ and $\delta$ be small parameters characterizing the range of indeterminacy and

$$\mathcal{U}_A := \left\{ A \in L^\infty(\Omega; \mathbb{M}^{n \times n}) \mid A = A_0 + \varepsilon E, \ E \in \mathcal{E} \right\},$$

$$\mathcal{U}_f := \left\{ f \in L^2(\Omega) \mid f = f_0 + \delta \phi, \ \phi \in \mathcal{F} \right\},$$

where

$$\mathcal{E} := \left\{ E \in L^\infty(\Omega; \mathbb{M}^{n \times n}) \mid \| E \|_{\infty, \Omega} \leq 1 \right\},$$

$$\mathcal{F} := \left\{ \phi \in L^2(\Omega) \mid \| \phi \|_{2, \Omega} \leq 1 \right\}.$$

We assume that the parameter $\varepsilon$ is small enough, so that the problems remain uniformly elliptic for all possible data, so that the relation

$$c_1 |\xi|^2 \leq A_0 \xi \cdot \xi \leq c_2 |\xi|^2, \quad \forall \xi \in \mathbb{R}^n,$$

implies a similar double inequality for all $A \in \mathcal{U}_A$. 

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
Since

\[ |E\xi \cdot \xi| = |E : (\xi \otimes \xi)| \leq |E| |\xi|^2, \]

we find that

\begin{align*}
A\xi \cdot \xi & \geq A_0\xi \cdot \xi - \varepsilon |E| |\xi|^2 \geq (c_1 - \varepsilon)|\xi|^2, & \text{(8.6)} \\
A\xi \cdot \xi & \leq A_0\xi \cdot \xi + \varepsilon |E| |\xi|^2 \leq (c_2 + \varepsilon)|\xi|^2. & \text{(8.7)}
\end{align*}

Therefore, we must assume that possible "disturbances" are sufficiently small, namely

\[ \varepsilon < c_1. \]

For the inverse matrix, we have

\begin{align*}
c_2^{-1}|\xi|^2 & \leq A_0^{-1}\xi \cdot \xi \leq c_1^{-1}|\xi|^2, & \text{(8.8)} \\
(c_2 + \varepsilon)^{-1}|\xi|^2 & \leq A^{-1}\xi \cdot \xi \leq (c_1 - \varepsilon)^{-1}|\xi|^2, & \text{(8.9)}
\end{align*}

where \( A \in U_A \).
Indeterminacy is explicitly accounted by the Majorant

A principle possibility to involve indeterminacy data into the consideration is based on that **external data are explicitly presented in the Majorant.**

Indeed, we have

\[
\int_{\Omega} A \nabla (\tilde{u} - v) \cdot \nabla (\tilde{u} - v)^2 \leq \\
(1 + \beta) \int_{\Omega} \left( A \nabla v \cdot \nabla v + A^{-1} y \cdot y - 2 \nabla v \cdot y \right) \, dx + \\
\frac{(1 + \beta) C^2_{\Omega}}{\beta} \| \text{div} y + f \|^2.
\]

We do not know \( A, f \) (and also \( u \)) exactly. But we can try to express all terms in this estimate via \( \varepsilon, \delta, A_0, \) and \( f_0 \).
The left–hand side of the estimate is easy to estimate from below. Indeed,

\[ \| \nabla (v - \tilde{u}) \|^2 = \int_\Omega \left( A_0 + \varepsilon E \right) \nabla (v - u) \cdot \nabla (v - u) \, dx \geq \]

\[ \geq (c_1 - \varepsilon) \| \nabla (v - u) \|^2. \]

Then, we find that

\[ (c_1 - \varepsilon) \| \nabla (\tilde{u} - v) \|^2 \leq \]

\[ \leq \sup_{A \in \mathcal{A}, f \in F} \inf_{y, \beta} \left\{ (1 + \beta) \int_\Omega \left( A \nabla v \cdot \nabla v + A^{-1} y \cdot y - 2 \nabla v \cdot y \right) \, dx + \right. \]

\[ \left. + (1 + \beta) C^2_\Omega \| \text{div} y + f \|^2 \right\}. \]

In this estimate an approximate solution \( v \) contains both APPROXIMATION and INDETERMINACY Errors!
Basic idea

Since \( \sup \inf \leq \inf \sup \), we can change the order and obtain

\[
(c_1 - \varepsilon) \| \nabla (u - v) \|^2 \leq \\
\leq \inf_{y, \beta} \sup_{A \in \mathcal{A}, f \in \mathcal{F}} \left\{ (1 + \beta) \int_{\Omega} \left( A \nabla v \cdot \nabla v + A^{-1} y \cdot y - 2 \nabla v \cdot y \right) dx + \\
+ \frac{(1 + \beta) C^2(\Omega, A)}{\beta} \| \text{div} y + f \|^2 \right\}.
\]

Now, our aim is to find an analytical estimate for the supremum that explicitly involves indeterminacy parameters.
Now, the upper bound of the error of an approximate solution $v$ with respect to the "worst case situation" comes in the form

$$e_{\text{max}}^2(v, \gamma) \leq \frac{1}{c_1} \left( 1 + \frac{\varepsilon}{c_1 - \varepsilon} \right) \left\{ (1 + \beta) \sup_{A \in U_A} D(\nabla v, y) + \frac{(1 + \beta)}{2\beta} \sup_{A \in U_A} C^2(\Omega, A) \sup_{f \in U_f} \| \text{div} \ y - f \|^2 \right\}, \quad (8.10)$$

which is valid for any $y \in Q^*$ and $\beta > 0$. Let us consider its terms.
To obtain a transparent estimate we need to find upper bounds for the quantities

\[
\sup_{A \in \mathcal{U}_A} D(\nabla \mathbf{v}, \mathbf{y}),
\]

\[
C^2(\Omega, A),
\]

\[
\sup_{f \in \mathcal{U}_f} \| \text{div} \mathbf{y} - f \|^2.
\]
First, we analyze the functional
\[ D(\Lambda v, y) := \frac{1}{2}(A\Lambda v, \Lambda v) + \frac{1}{2}(A^{-1}y, y) - (\Lambda v, y) \]
for any \( A \in \mathcal{U}_A \). First, we rewrite the first term
\[ \int_\Omega A \nabla v \cdot \nabla v \, dx = \int_\Omega (A_0 \nabla v \cdot \nabla v + \varepsilon E \nabla v \cdot \nabla v) \, dx. \]
Now, our aim is to estimate the most complicated second term. Present the inverse matrix as follows
\[ A^{-1} = (A_0 + \varepsilon E)^{-1} = \left( A_0 (I + \varepsilon A_0^{-1}E) \right)^{-1} = (I + \varepsilon B)^{-1} A_0^{-1}, \]
where \( B = A_0^{-1}E \). Note that
\[ \varepsilon |B| = \varepsilon |A_0^{-1}E| \leq \varepsilon |A_0^{-1}| |E| \leq \varepsilon c_1^{-1} < 1, \]
and, therefore, \((I + \varepsilon B)^{-1}\) can be presented as a convergent matrix series, namely
\[(\mathbb{I} + \varepsilon B)^{-1} = \mathbb{I} + \sum_{j=1}^{\infty} (-1)^j \varepsilon^j B^j,\]

Hence, we can present the second term as a combination of known matrixes \(A_0\) and powers of \(\varepsilon\).

\[
A^{-1}y \cdot y = (\mathbb{I} + \varepsilon B)^{-1}A_0^{-1}y \cdot y = \left( \mathbb{I} + \sum_{j=1}^{\infty} (-1)^j \varepsilon^j B^j \right) A_0^{-1}y \cdot y = \\
= A_0^{-1}y \cdot y - \varepsilon BA_0^{-1}y \cdot y + \sum_{j=2}^{\infty} (-1)^j \varepsilon^j B^j A_0^{-1}y \cdot y.
\]
Since $E \in \mathcal{E}$, we have

$$
\int_{\Omega} B^j A_0^{-1} y \cdot y \, dx \leq \int_{\Omega} |A_0^{-1}|^{j+1} |E|^j |y|^2 \, dx \leq c_1^{-(j+1)} \|y\|^2
$$

$$
\int_{\Omega} A^{-1} y \cdot y \, dx \leq \int_{\Omega} (A_0^{-1} y \cdot y - \varepsilon B A_0^{-1} y \cdot y) \, dx + \left( \sum_{j=2}^{\infty} (-1)^j \varepsilon^j c_1^{-(j+1)} \right) \|y\|^2.
$$
We find that the first term $D$ explicitly depends on $\varepsilon$:

$$D(\nabla v, y) \leq D_0(\nabla v, y) + \frac{\varepsilon}{2} \int_{\Omega} (E \nabla v \cdot \nabla v - BA_0^{-1} y \cdot y) \, dx +$$

$$+ \frac{1}{2c_1} \|y\|^2 \sum_{j=2}^{\infty} \left(-\frac{\varepsilon}{c_1}\right)^j,$$

where

$$D_0(\nabla v, y) = \int_{\Omega} \left(\frac{1}{2} A_0 \nabla v \cdot \nabla v + \frac{1}{2} A_0^{-1} y \cdot y - \nabla v \cdot y\right) \, dx.$$
Since all the matrices are symmetric, we have

\[ E \nabla v \cdot \nabla v - A_0^{-1} E A_0^{-1} y \cdot y = E (\nabla v - A_0^{-1} y) \cdot (\nabla v + A_0^{-1} y). \]

Now, we obtain

\[ D(\nabla v, y) \leq D_0(\nabla v, y) + \varepsilon \int_{\Omega} E (\nabla v - A_0^{-1} y) \cdot (\nabla v + A_0^{-1} y) \, dx + \left( \frac{\varepsilon}{c_1} \right)^2 \frac{1}{2(\varepsilon + c_1)} \| y \|^2. \]

Note that the last two terms presents a positive penalty arose due to indeterminacy. All the terms in the right–hand side are directly computable!
Next,

\[ \frac{1}{C^2(\Omega, A)} = \inf_{w \in V_0} \frac{\int_{\Omega} A \nabla w \cdot \nabla w \, dx}{\|w\|^2}, \]

where

\[ \int_{\Omega} A \nabla w \cdot \nabla w \, dx \geq \int_{\Omega} A_0 \nabla w \cdot \nabla w \, dx - \varepsilon \|\nabla w\|^2. \]

Hence,

\[ \frac{1}{C^2(\Omega, A)} \geq (1 - \varepsilon c_1^{-1}) \inf_{w \in V_0} \frac{\int_{\Omega} A_0 \nabla w \cdot \nabla w \, dx}{\|w\|^2} = \frac{c_1 - \varepsilon}{c_1} \frac{1}{C^2(\Omega, A_0)} \]

and

\[ C^2(\Omega, A) \leq \left(1 + \frac{\varepsilon}{c_1 - \varepsilon}\right) C^2(\Omega, A_0). \]
For any \( g \in L^2(\Omega) \)

\[
\sup_{\varphi \in F} \int_{\Omega} (g - \varphi)^2 \, dx = \|g\|^2 + 2\|g\| + 1.
\]

By this relation, we find the value of the term

\[
\sup_{\varphi \in F} \int_{\Omega} |\text{div} \, y - f_0 - \delta \varphi|^2 \, dx,
\]

which is

\[
\|\text{div} \, y - f_0\|^2 - 2\delta \|\text{div} \, y - f_0\| + \delta^2.
\]

Then, we arrive at the final estimate. To represent its right-hand side of this estimate in a more transparent form, we introduce a number of quantities.
\[ M_{00}(v, \beta, y) := (1 + \beta)D_0(\nabla v, y) + \left(1 + \frac{1}{\beta}\right) \frac{C^2(\Omega, A_0)}{2} \|\text{div } y - f_0\|^2, \]
\[ M_{10}(v, \beta, y) := \frac{\varepsilon}{2} \left( \int_{\Omega} |(\nabla v - A_0^{-1}y) \cdot (\nabla v + A_0^{-1}y)| \, dx + \right) \]
\[ + \left(1 + \frac{1}{\beta}\right) \frac{C^2(\Omega, A_0)}{c_1 - \varepsilon} \int_{\Omega} |\text{div } y - f_0|^2 \, dx, \]
\[ M_{01}(v, \beta, y) := \delta \left(1 + \frac{1}{\beta}\right) C^2(\Omega, A_0) \|\text{div } y - f_0\|, \]
\[ M_{11}(v, \beta, y) := \varepsilon \delta \left(1 + \frac{1}{\beta}\right) \frac{C^2(\Omega, A_0)}{c_1 - \varepsilon} \|\text{div } y - f_0\|, \]
\[ M_{22}(v, \beta, y) := (1 + \beta) \left(\frac{\varepsilon}{c_1}\right)^2 \frac{\|y\|^2}{2(\varepsilon + c_1)} + \left(1 + \frac{1}{\beta}\right) \frac{c_1 C^2(\Omega, A_0)}{2(c_1 - \varepsilon)} \delta^2. \]
We obtain an upper bound of $e_\oplus(v, \Upsilon)$ in the form

$$e_\oplus^2(v, \Upsilon) = \frac{1}{c_1 - \epsilon} \inf_{y \in Q^*, \beta > 0} \left( M_{00}(v, \beta, y) + 
\right.$$

$$+ M_{01}(v, \beta, y) + M_{10}(v, \beta, y) + M_{11}(v, \beta, y) + M_{22}(v, \beta, y) \right).$$

(8.11)

The term $M_{00}(v, \beta, y)$ coincides with the majorant constructed for the ”mean” problem (with $A_0$ and $f_0$). It represents the approximation error. The terms $M_{10}$, $M_{01}$, and $M_{11}$ are given by some combinations of the weighted residual and small parameters $\epsilon$ and $\delta$. In principle, all these terms can be made arbitrarily small by taking $v$ close enough to the exact solution $u$ of the problem with $A = A_0$ and $\ell = f_0$ and $y$ close enough to $A_0 \nabla u$. 
Inherent error

In contrast, the term $M_{22}(v, \beta; y)$ is always positive. This term contains the inherent part of the error, which does not depend on the accuracy of numerical approximations. Indeed, in all cases we have

$$M_{22}(v, \beta, y) \geq C^2(\Omega, A_0) \frac{c_1 \delta^2}{2(c_1 - \varepsilon)}.$$ 

This quantity does not depend on the choice of $v$, $\beta$, and $y$. It gives an idea of the accuracy limit that could be achieved within the framework of the worst-case scenario.
Computable upper bounds

Take \( \{Q_k^*\} \subset Q^* \). Then,

\[
e^2_\oplus(v, \Upsilon) \leq e^2_\oplus(k, \Upsilon) =
\]

\[
= \frac{1}{c_1} \left( 1 + \frac{\varepsilon}{c_1 - \varepsilon} \right) \inf_{y \in Q^*_k, \beta > 0} \left\{ \sum_{s,t=0}^{1} M_{st}(v, \beta, y) + M_{22}(v, \beta, y) \right\}.
\]

If \( Q_k^* \subset Q^*_{k+1} \), then the sequence \( \{e^2_\oplus(v, \Upsilon)\} \) monotonically decreases but may not tend to zero.
Lower bound of the error

To find a lower bound, we use the relation

\[ \frac{1}{2} \int_{\Omega} A \nabla (v - \tilde{u}) \cdot \nabla (v - u) \, dx = \sup_{w \in V_0} M_{\ominus}(v, w), \]

where

\[ M_{\ominus}(v, w) = - \int_{\Omega} \left( \frac{1}{2} A \nabla w \cdot \nabla w + A \nabla v \cdot \nabla w + f w \right) \, dx. \]

Recall that

\[ A \xi \cdot \xi \leq (\epsilon + c_2) |\xi|^2. \]

By this inequality we can estimate the left–hand side from below, so that for any pair

\[ (A, f) \in U_A \times U_f \]

that generates the respective solution \( \tilde{u} \) we have
\[
\frac{1}{2} \|\nabla (v - \tilde{u})\|^2 \geq \frac{1}{2(\varepsilon + c_2)} \int_{\Omega} A \nabla (v - \tilde{u}) \cdot \nabla (v - \tilde{u}) \, dx = \\
= \frac{1}{\varepsilon + c_2} \sup_{w \in V_0} M_{\ominus}(v, w).
\]

Therefore,

\[
e_{\text{min}}(v, \Upsilon) = \inf_{\tilde{u} \in \Upsilon} \frac{1}{2} \|\nabla (v - \tilde{u})\|^2 \geq \\
\geq \frac{1}{\varepsilon + c_2} \inf_{(A,f) \in U_A \times U_f} \sup_{w \in V_0} M_{\ominus}(v, w) \geq \\
\geq \frac{1}{\varepsilon + c_2} \sup_{w \in V_0} \inf_{(A,f) \in U_A \times U_f} M_{\ominus}(v, w).
\]
We have

$$e_{\min}(v, \gamma) \geq \frac{1}{\varepsilon + c_2} \sup_{w \in V_0} \left\{ - \int_{\Omega} \left( \frac{1}{2} A_0 \nabla w \cdot \nabla w + A_0 \nabla v \cdot \nabla w + f_0 w \right) \, dx + \inf_{E \in E, \phi \in F} \left( -\varepsilon \int_{\Omega} \left( \frac{1}{2} E \nabla w \cdot \nabla w + E \nabla v \cdot \nabla w \right) \, dx - \delta \int_{\Omega} w \phi \, dx \right) \right\}. $$
By the algebraic inequality

\[ E_a \cdot b + E_c \cdot b = E_{ij}(a_j b_j + c_j b_i) = E_{ij}(b_i(a_j + c_j)) = E : (b \otimes (a + c)) \]

we find that

\[
\inf_{E \in \mathcal{E}} \left\{ - \int_{\Omega} \left( \frac{1}{2} E \nabla w \cdot \nabla w + E \nabla v \cdot \nabla w \right) \, dx \right\} =
\]

\[
= - \int_{\Omega} \left| \left( \frac{1}{2} \nabla w + \nabla v \right) \otimes \nabla w \right| \, dx.
\]

It is easy to see that

\[
\inf_{\varphi \in \mathcal{F}} \left\{ - \int_{\Omega} w \varphi \, dx \right\} = -\|w\|.
\]
Now, we obtain

\[ e_{\min}^2(v, \mathcal{R}) \geq \frac{1}{\varepsilon + c_2} \sup_{w \in V_0} \left\{ - \int_{\Omega} \left( \frac{1}{2} A_0 \nabla w \cdot \nabla w + A_0 \nabla v \cdot \nabla w + f_0 w \right) \, dx - \varepsilon \int_{\Omega} \left| \left( \frac{1}{2} \nabla w + \nabla v \right) \otimes \nabla w \right| \, dx - \delta \| w \| \right\}. \]
Introduce the quantities

\[ m_{00}(v, w) = - \int_{\Omega} \left( \frac{1}{2} A_0 \nabla w \cdot \nabla w + A_0 \nabla v \cdot \nabla w + f_0 w \right) \, dx, \]

\[ m_{10}(v, w) = - \varepsilon \int_{\Omega} \left| \left( \frac{1}{2} \nabla w \otimes \nabla w + \nabla v \otimes \nabla w \right) \right| \, dx, \]

\[ m_{01}(w) = - \delta \| w \|. \]

Then, we represent the lower bound in the form

\[ e^2_\ominus(v, \mathcal{Y}) := \frac{1}{\varepsilon + c_2} \sup_{w \in \mathcal{V}_0} \{ m_{00}(v, w) + m_{01}(v, w) + m_{10}(w) \} \geq 0. \]

(8.12)

In this estimate, the term \( m_{00}(v, w) \) contains the major part of the approximation error. It vanishes if \( v \) is a solution of the ”mean” problem with \( A = A_0 \) and \( \ell = f_0 \). Two other terms reflect the influence of the small parameters \( \delta \) and \( \varepsilon \).
Computable lower bounds

Take a collection of finite-dimensional subspace $V_{0k}$ and solve the problems

$$e^2_{\min}(v, \Upsilon) \geq e^2_{k\ominus}(v, \Upsilon) =$$

$$= \frac{1}{\epsilon + c_2} \sup_{w \in V_{0k}} \{m_{00}(v, w) + m_{01}(v, w) + m_{10}(w)\}.$$

Now $e^2_{k\ominus}(v, \Upsilon)$ can be used to estimate the efficiency of further computational efforts within the framework of the best-case scenario.
To refine or not to refine? That is the question.

If $e^2_{k\otimes}(v, \mathcal{Y})$ are large, then **approximation errors are significant**. In this case, it is worth computing a new approximation on a finer mesh.

If for a certain $k$ the quantity $e^2_{k\otimes}(v, \mathcal{Y})$ is very small, then an approximate solution computed is already close to some $u \in \mathcal{Y}$. Since we do not know exactly the data (and, thus, have no way to select the proper $u$) **all further computations and mesh refinements are in a sense useless** because they cannot improve our presentation on the true solution.
Lecture 9.
A POSTERIORI ESTIMATES FOR MIXED METHODS
Mixed approximations. A glance from the minimax theory

Consider our basic problem

\[ \text{div} A \nabla u + f = 0 \quad \text{in} \; \Omega, \]
\[ u = u_0 \text{ on } \partial_1 \Omega, \]
\[ A \nabla u \cdot n = F \text{ on } \partial_2 \Omega, \]

\[ c_1^2 |\xi|^2 \leq A(x) \xi \cdot \xi \leq c_2^2 |\xi|^2 \quad \forall \xi \in \mathbb{R}^d, \text{ for a.e. } x \in \Omega, \]

where \( u_0 \in H^1(\Omega), \; f \in L^2(\Omega), \; F \in L^2(\partial_2 \Omega). \) Functional spaces

\[ V := H^1(\Omega), \; V_0 := \{v \in V \mid v = 0 \text{ on } \partial_1 \Omega\}, \; \hat{V} := L^2(\Omega), \]
\[ Q := L^2(\Omega; \mathbb{R}^d) \]
\[ \hat{Q}^+ := \{y \in \hat{Q} \mid y \cdot n|_{\partial_2 \Omega} \in L^2(\partial_2 \Omega)\}. \]
We recall that $\|q\|_{\text{div}}$ is the norm in $H(\Omega; \text{div})$:

$$\|q\|_{\text{div}} := (\|q\|^2 + \|\text{div}q\|^2)^{1/2} \quad \forall q \in Q$$

and

$$\|q\| := \left( \int_{\Omega} Aq \cdot q \, dx \right)^{1/2}, \quad q \in Q$$

$$\|q\|^* := \left( \int_{\Omega} A^{-1}q \cdot q \, dx \right)^{1/2}$$

Note that,

$$\bar{c}_1^2|\xi|^2 \leq A^{-1}(x)\xi \cdot \xi \leq \bar{c}_2^2|\xi|^2 \quad \forall \xi \in \mathbb{R}^d, \text{ for a.e. } x \in \Omega$$

with $\bar{c}_1 = 1/c_2$, $\bar{c}_2 = 1/c_1$. 

S. Repin  

RICAM, Special Radon Semester, Linz, 2005.
Generalized solution of the problem considered can be viewed as a saddle point of the Lagrangian

\[ L(v, q) := \int_{\Omega} \left( \nabla v \cdot q - \frac{1}{2} A^{-1} q \cdot q \right) \, dx - \ell(v), \]

where \( \ell(v) = \int_{\Omega} fv \, dx + \int_{\partial_2 \Omega} Fv \, ds. \)
In this formulation \((u, p) \in (V_0 + u_0) \times Q\) satisfies the relations

\[
\int_{\Omega} \left( A^{-1} p - \nabla u \right) \cdot q \, dx = 0 \quad \forall q \in Q, \tag{9.1}
\]

\[
\int_{\Omega} p \cdot \nabla w \, dx - \ell(w) = 0 \quad \forall w \in V_0. \tag{9.2}
\]

Here

\[p = A \nabla u, \quad \text{is satisfied in } L_2(\Omega) - \text{sense}\]

\[\text{div } p + f = 0 \quad \text{in } \Omega \text{ and}\]

\[p \cdot n = F \text{ on } \partial_2 \Omega \quad \text{are satisfied in a weak sense.}\]
As we have seen in previous lectures \( L \) generates two functionals

\[
J(v) := \sup_{q \in Q} L(v, q) = \frac{1}{2} \| \nabla v \|^2 - \ell(v)
\]

and

\[
I^*(q) := -\frac{1}{2} \| q \|^2 - \ell(u_0) + \int_{\Omega} \nabla u_0 \cdot q \, dx.
\]

Also, we know that

\[
\inf_{v \in V_0 + u_0} J(v) := \inf \mathcal{P} = L(u, p) = \sup \mathcal{P}^* := \sup_{q \in Q_\ell} I^*(q), \quad (9.3)
\]

where \( Q_\ell := \{ q \in Q \mid \int_{\Omega} q \cdot \nabla w \, dx = \ell(w) \quad \forall w \in V_0 \} \).
Primal Mixed Method (PMM)

Let $Q_h \subset Q$ and $V_{0h} \subset V_0$ are subspaces constructed by FE approximation, then a discrete analog of $(9.1)-(9.2)$ is the **Primal Mixed Finite Element Method**.


In PMM, we need to find a pair of functions 
$(u_h, p_h) \in (V_{0h} + u_0) \times Q_h$ such that

\[
\int_{\Omega} \left( A^{-1} p_h - \nabla u_h \right) \cdot q_h \, dx = 0 \quad \forall q_h \in Q_h, \tag{9.4}
\]

\[
\int_{\Omega} p_h \cdot \nabla w_h \, dx - \ell(w_h) = 0 \quad \forall w_h \in V_{0h}. \tag{9.5}
\]

In this formulation, $u_h$ can be constructed by means of the Courant-type elements and $p_h$ by piecewise constant functions.
Dual Mixed Method (DMM)

Another mixed formulation arises if we represent $L$ in a somewhat different form. First, we introduce the functional $g : (\mathbf{V}_0 + \mathbf{u}_0) \times \hat{Q} \rightarrow \mathbb{R}$ by the relation

$$g(v, q) := \int_{\Omega} (\nabla v \cdot q + v(\text{div}q)) \, dx. \quad (9.6)$$

We have

$$L(v, q) = \int_{\Omega} \left( \nabla v \cdot q - \frac{1}{2} A^{-1} q \cdot q \right) \, dx - \ell(v) =$$

$$= g(v, q) - \int_{\Omega} v(\text{div}q) \, dx - \frac{1}{2} \| q \|^2_* - \ell(v).$$
Introduce the set

$$\hat{Q}_F := \{ q \in \hat{Q} \mid g(w, q) = \int_{\partial_2 \Omega} Fw \ ds \ \forall w \in V_0 \}.$$ 

Note that for $q \in \hat{Q}_F$ we have

$$g(v, q) = g(w + u_0, q) = g(w, q) + g(u_0, q) = \int_{\partial_2 \Omega} Fw \ ds + g(u_0, q) \ \forall w \in V_0.$$ 

Therefore, if the variable $q$ is taken not from $Q$ but from the narrower set $\hat{Q}_F$, then the Lagrangian can be written as

$$\hat{L}(v, q) := -\frac{1}{2} \| q \|^2_* - \int_{\Omega} v(\text{div}q) \ dx - \int_{\Omega} fv \ dx - \int_{\partial_2 \Omega} Fu_0 \ ds + g(u_0, q).$$
We observe that the new Lagrangian $\hat{L}$ is defined on a wider set of primal functions $v \in \hat{V}$, but uses a narrower set $\hat{Q}_F$ for the fluxes.

The problem of finding $(\hat{u}, \hat{p}) \in \hat{V} \times \hat{Q}_F$ such that

$$\hat{L}(\hat{u}, q) \leq \hat{L}(\hat{u}, \hat{p}) \leq \hat{L}(\hat{v}, \hat{p}) \quad \forall q \in \hat{Q}_F, \forall \hat{v} \in \hat{V}$$

leads to the so-called **Dual Mixed Formulation** of the problem in question (see, e.g., F. Brezzi and M. Fortin).
From (9.7) we obtain the necessary conditions for the dual mixed formulation. Since
\[ \hat{L}(\hat{u}, \hat{q}) \leq \hat{L}(\hat{u}, \hat{p}) \quad \forall \hat{q} \in \hat{Q}_F, \]
we have
\[ -\frac{1}{2} \| \hat{p} + \lambda \eta \|^2_\ast - \int_\Omega \hat{u}(\text{div}(\hat{p} + \lambda \eta) - \hat{f}) \, dx - \int_\Omega \hat{F} u_0 \, ds + \hat{g}(u_0, \hat{p} + \lambda \eta) \leq \]
\[ -\frac{1}{2} \| \hat{p} \|^2_\ast - \int_\Omega \hat{u}(\text{div}\hat{p}) \, dx - \int_\Omega \hat{f} \, \hat{u} \, dx - \int_{\partial_2 \Omega} \hat{F} u_0 \, ds + \hat{g}(u_0, \hat{p}), \]
where \( \lambda \) is a real number and \( \eta \) is a function in \( \hat{Q}_0 := \hat{Q}_F \) with \( F = 0 \). Now, arrive at the relation
\[ -\lambda \int_\Omega (A^{-1}\hat{p} \cdot \eta + \hat{u}(\text{div}\eta)) \, dx + \lambda \hat{g}(u_0, \eta) \leq \frac{\lambda^2}{2} \int_\Omega A^{-1} \eta \cdot \eta \, dx. \]
Rewrite it as

\[
\int_{\Omega} (A^{-1}\hat{p} \cdot \eta + \hat{u}(\text{div}\eta)) \, dx - g(u_0, \eta) \geq \frac{\lambda}{2} \int_{\Omega} A^{-1} \eta \cdot \eta \, dx.
\]

Since \( \lambda > 0 \) can be taken arbitrarily small, the latter relation may hold only if

\[
\int_{\Omega} (A^{-1}\hat{p} \cdot \eta + \hat{u}\text{div}\eta) \, dx - g(u_0, \eta) \geq 0.
\]

But \( \eta \) is an arbitrary element of a linear manifold \( \hat{Q}_0 \), so that \( +\eta \) can be replaced by \( -\eta \) what leads to the conclusion that

\[
\int_{\Omega} (A^{-1}\hat{p} \cdot \eta + \hat{u}\text{div}\eta) \, dx - g(u_0, \eta) = 0 \quad \forall \eta \in \hat{Q}_0.
\]
From

\[ \hat{L}(\hat{u}, \hat{p}) \leq \hat{L}(\hat{u} + \hat{v}, \hat{p}) \quad \forall \hat{v} \in \hat{V} := L^2(\Omega) \]

we observe that the terms of \( \hat{L} \) linear with respect to the "pressure" must vanish. Namely, we obtain

\[ \int_{\Omega} (\hat{v} \text{div} \hat{p} + f\hat{v}) \, dx = 0 \]

Thus, we arrive at the system

\[ \int_{\Omega} \left( A^{-1} \hat{p} \cdot \hat{q} + (\text{div} \hat{q})\hat{u} \right) \, dx = g(u_0, \hat{q}) \quad \forall \hat{q} \in \hat{Q}_0, \quad (9.8) \]

\[ \int_{\Omega} (\text{div} \hat{p} + f)\hat{v} \, dx = 0 \quad \forall \hat{v} \in \hat{V}. \quad (9.9) \]
We observe that now the condition

$$\text{div}\hat{p} + f = 0$$

is satisfied in a ”strong” ($L_2$) sense, the Neumann type boundary condition is viewed as the essential boundary condition, and the relation

$$\hat{p} = A \nabla \hat{u}$$

and the Dirichlet type boundary condition are satisfied in a weak sense.

These properties of the DMM lead to that the respective finite dimensional formulations are better adapted to the satisfaction of the equilibrium type relations for the fluxes. This fact is important in many applications where a sharp satisfaction of the equilibrium relations is required.
The Lagrangian $\hat{L}$ also generates two functionals

$$\hat{J}(\hat{v}) := \sup_{\hat{q} \in \hat{Q}_F} \hat{L}(\hat{v}, \hat{q}) \quad \text{and} \quad \hat{I}^*(\hat{q}) := \inf_{\hat{v} \in \hat{V}} \hat{L}(\hat{v}, \hat{q}) .$$

The two corresponding variational problems are

$$\inf_{\hat{v} \in \hat{V}} \hat{J}(\hat{v}) \quad \text{and} \quad \sup_{\hat{q} \in \hat{Q}_F} \hat{I}^*(\hat{q}).$$

They are called Problems $\hat{P}$ and $\hat{P}^*$, respectively. Note that the functional $\hat{J}$ (unlike $J$) has no simple explicit form. However, we can prove the solvability of Problem $\hat{P}$ by the following Lemma.
Lemma

For any \( \hat{v} \in \hat{V} \) and \( F \in L_2(\partial_2 \Omega) \) there exists \( p^v \in \hat{Q}_F \) such that

\[
\text{div} p^v + \hat{v} = 0 \quad \text{in } \Omega, \quad \tag{9.10}
\]
\[
\| p^v \|_* \leq C_\Omega \left( \| \hat{v} \| + \| F \|_{\partial_2 \Omega} \right). \quad \tag{9.11}
\]

Proof. We know that the boundary-value problem

\[
\text{div} A \nabla u^v + \hat{v} = 0 \quad \text{in } \Omega, \\
\quad u^v = 0 \quad \text{on } \partial_1 \Omega, \\
\quad A \nabla u^v \cdot n = F \quad \text{on } \partial_2 \Omega
\]

possesses the unique solution \( u^v \in V_0 \).
For it and the energy estimate

\[ \| \nabla u^v \| \leq C_\Omega \left( \| \hat{v} \| + \| F \|_{\partial_2 \Omega} \right) \]

holds. Let \( p^v := A \nabla u^v \). We have

\[ \text{div} p^v + \hat{v} = 0. \]

Obviously, \( p^v \in \hat{Q}_F \) and, since

\[ \| p^v \|_{\ast}^2 = \int_{\Omega} A^{-1}(A \nabla u^v) \cdot (A \nabla u^v) \, dx = \| \nabla u^v \|^2, \]

we find that (9.11) also holds.

\( \square \)
By the Lemma we can easily prove the coercivity of \( \hat{J} \) on \( \hat{V} \).

Indeed,

\[
\hat{J}(\hat{v}) \geq \hat{L}(\hat{v}, \alpha p^v) = \\
- \frac{1}{2} \| \alpha p^v \|^2 - \alpha \int_{\Omega} \hat{v}(\text{div} p^v) \, dx - \int_{\Omega} f \hat{v} \, dx - \int_{\partial_2 \Omega} F u_0 \, ds + g(u_0, \alpha p^v) = \\
\left( - \frac{1}{2} \alpha^2 \right) \| p^v \|^2 + \| \hat{v} \|^2 - \| f \| \| \hat{v} \| + g(u_0, \alpha p^v) - \int_{\partial_2 \Omega} F u_0 \, ds.
\]

Here \( |g(u_0, \alpha p^v)| \leq \alpha \| p^v \|_{\text{div}} \| u_0 \|_{1,2,\Omega} \) and

\[
\| p^v \|_{\text{div}}^2 = \| p^v \|^2 + \| \text{div} p^v \|^2 \leq \frac{1}{c_1} \| p^v \|^2 + \| \hat{v} \|^2 \leq \\
\leq \frac{1}{c_1} C^2_\Omega \left( \| \hat{v} \| + \| F \|_{\partial_2 \Omega} \right)^2 + \| \hat{v} \|^2.
\]
Therefore

\[ \hat{J}(\hat{v}) \geq -\frac{1}{2} \alpha^2 C^2_\Omega \| \hat{v} \|^2 + \alpha \| \hat{v} \|^2 + \Theta(\| \hat{v} \|) + \Theta_0 , \]

where \( \Theta(\| \hat{v} \|) \) contains the terms linear with respect to \( \| \hat{v} \| \) and \( \Theta_0 \) does not depend on \( \hat{v} \). Take \( \alpha = 1/C^2_\Omega \). Then

\[ \hat{J}(\hat{v}) \geq \frac{1}{2C^2_\Omega} \| \hat{v} \|^2 + \Theta(\| \hat{v} \|) + \Theta_0 \rightarrow +\infty \quad \text{as} \quad \| \hat{v} \| \rightarrow \infty . \]

It is not difficult to prove that the functional \( \hat{J} \) is convex and lower semicontinuous. Therefore, Problem \( \hat{P} \) has a solution \( \hat{u} \).
Inf-Sup condition for the dual mixed formulation

Corollary

Lemma implies the \textit{inf-sup} condition

\[
\inf_{\phi \in L^2(\Omega)} \sup_{\psi \in L^2(\partial_2 \Omega)} \inf_{q \in \hat{Q}_F} \sup_{\Omega} \left( \int \phi \text{div} q \, dx + \int_{\partial_2 \Omega} \psi q \cdot n \, ds \right) \geq C_0 > 0.
\]
The Dual Problem with respect to the Lagrangian $\hat{L}$

Let us now construct the dual functional $\hat{I}^\ast$. It is easy to see that

$$\hat{I}^\ast(\hat{q}) = \inf_{\hat{v}} \hat{L}(\hat{v}, \hat{q}) =$$

$$= \inf_{\hat{v}} \left\{ -\frac{1}{2} \| \hat{q} \|_2^* - \int_{\Omega} \hat{v} (\text{div} \hat{q}) dx - \int_{\Omega} f v dx - \int_{\partial_2 \Omega} F u_0 ds + g(u_0, \hat{q}) \right\} =$$

$$= -\frac{1}{2} \| \hat{q} \|_2^* + g(u_0, \hat{q}) - \int_{\partial_2 \Omega} F u_0 ds$$

provided that $\text{div} \hat{q} + f = 0$ (in the $L_2$-sense). In all other cases $\hat{I}^\ast(\hat{q}) = -\infty$. 
Recalling that \( \mathsf{div} \hat{q} = -f \) (in \( L_2(\Omega) \)-sense), we find that the dual functional for such a case has the form

\[
\hat{l}^*(q) = -\frac{1}{2} \| \hat{q} \|_{\ast}^2 + \int_\Omega (\nabla u_0 \cdot \hat{q} - fu_0) \, dx - \int_{\partial_2 \Omega} Fu_0 \, ds
\]

\[
= \int_\Omega \nabla u_0 \cdot \hat{q} \, dx - \frac{1}{2} \| \hat{q} \|_{\ast}^2 - \ell(u_0),
\]

Since \( \hat{q} \in \hat{Q}_F \), we have

\[
\int_\Omega \nabla w \cdot \hat{q} \, dx = -\int_\Omega (\mathsf{div} \hat{q}) w \, dx + \int_{\partial_2 \Omega} Fw \, ds \quad \forall w \in V_0.
\]

we see that \( \hat{q} \) satisfies the relation

\[
\int_\Omega \nabla w \cdot \hat{q} \, dx = \ell(w) \quad \forall w \in V_0.
\]

In other cases, \( \hat{l}^*(\hat{q}) = -\infty \).
Thus, Problems $\mathcal{P}^*$ and $\hat{\mathcal{P}}^*$ coincide and are reduced to the maximization of $I^*$ on the set $Q_\ell$. This means that

$$\sup \mathcal{P}^* = \sup \hat{\mathcal{P}}^*.$$ 

Since the saddle point of $\hat{L}$ exists, we have

$$\hat{L}(\hat{u}, \hat{p}) = \inf \hat{\mathcal{P}} = \sup \hat{\mathcal{P}}^*,$$

but

$$\sup \hat{\mathcal{P}}^* = \sup \mathcal{P}^* = \inf \mathcal{P}.$$ 

Thus, we infer that

$$\inf \hat{\mathcal{P}} = \inf \mathcal{P}.$$
Thus, we conclude that \( u \in V_0 + u_0 \) (minimizer of \( P \)) also minimizes \( \hat{J} \) on \( \hat{V} \).

Analogously, if \( p \in Q_\ell \) is the maximizer of Problem \( P^* \), then

\[
\int_\Omega \nabla w \cdot p \, dx = \int_\Omega fw \, dx + \int_{\partial_2 \Omega} Fw \, ds \quad \forall w \in V_0. 
\]

From here we see that \( \text{div} p + f = 0 \) a.e. in \( \Omega \) and, hence,

\[
\int_\Omega (\nabla w \cdot p + (\text{div} p)w) \, dx = \int_{\partial_2 \Omega} Fw \, ds \quad \forall w \in V_0, 
\]

that is \( p \in \hat{Q}_F \). Thus, \( p \) is also the maximizer of Problem \( \hat{P}^* \).

The reverse statement that the solutions of \( \hat{P}, \hat{P}^* \) are also the solutions of \( P, P^* \) is not difficult to prove as well.
Hence, both mixed formulations have the same solution \((u, p)\) which is in fact the generalized solution of our problem.
Finite dimensional formulations

Let

\[ \hat{V}_h \subset \hat{V}, \quad \hat{Q}_{0h} \subset \hat{Q}_0 \quad \hat{Q}_{Fh} \subset \hat{Q}_F \]

A discrete analog of the dual mixed formulation is: Find \((\hat{u}_h, \hat{p}_h) \in \hat{V}_h \times \hat{Q}_{Fh}\) such that

\[
\int_{\Omega} \left( A^{-1} \hat{p}_h \cdot \hat{q}_h + \hat{u}_h \text{div} \hat{q}_h \right) dx = g(u_0, \hat{q}_h) \quad \forall \hat{q}_h \in \hat{Q}_{0h}, \tag{9.12}
\]

\[
\int_{\Omega} (\text{div} \hat{p}_h + f) \hat{v}_h dx = 0 \quad \forall \hat{v}_h \in \hat{V}_h. \tag{9.13}
\]
Error analysis for DMM

First we will obtain a priori error estimates for the dual mixed method and after that we will derive computable upper bounds for the quantities

\[ \| \nabla (u - u_h) \|, \| p - p_h \|_*, \| p - \hat{p}_h \|_{\text{div}}. \]
Below we will show a simple way of the derivation of projection type error estimates for the dual mixed method. By combining them with standard interpolation results, one can obtain known rate convergence estimates. A detailed exposition of this subject can be found in the above cited books.

Here, we present a simplified version, which, however contains the main ideas of the a priori error analysis for the dual mixed approximations.
For the sake of simplicity we will consider the case of uniform Dirichlet boundary conditions and a constant matrix $\mathbf{A}$. In this case, the basic system is as follows

$$
\int_{\Omega} \left( \mathbf{A}^{-1} \hat{\mathbf{p}} \cdot \hat{\mathbf{q}} + (\text{div}\, \hat{\mathbf{q}}) \hat{\mathbf{u}} \right) \, dx = 0 \quad \forall \hat{\mathbf{q}} \in \hat{\mathbf{Q}}_0,
$$

$$
\int_{\Omega} (\text{div}\, \hat{\mathbf{p}} + f) \hat{\mathbf{v}} \, dx = 0 \quad \forall \hat{\mathbf{v}} \in \hat{\mathbf{V}}.
$$

Since there is no Neumann part of the boundary, $\hat{\mathbf{Q}}_F$ and $\hat{\mathbf{Q}}_0$ coincides with $\hat{\mathbf{Q}} := H(\Omega, \text{div})$. 

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
In the considered case the system of DMM is as follows

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

$$\int_{\Omega} (\text{div} \hat{p}_h + f) \hat{v}_h dx = 0 \quad \forall \hat{v}_h \in \hat{V}_h.$$

Assumptions.

(a) $T_h$ is a regular triangulation of a polygonal domain $\Omega$.
(b) $\hat{V}_h = \{ v_h \in L^2 \mid v_h \in P^0(T) \ \forall T \in T_h \}$.
(c) $\hat{Q}_h = \{ q_h \in H(\Omega, \text{div}) \mid q_h \in RT^0(T) \ \forall T \in T_h \}$.
(d) $f \in P^0(T), \quad \forall T \in T_h$
Note that under the assumptions made

\[ \text{div} \bar{p}_h + f = 0 \quad \text{on any } T. \]

Indeed, this fact directly follows from the relation

\[ \int_{\Omega} (\text{div} \bar{p}_h + f) \hat{v}_h \, dx = 0 \quad \forall \hat{v}_h \in \hat{V}_h. \]

Therefore \( p_h \in Q_f \).
Compatibility and stability conditions

We need that one more condition be satisfied in order to provide the stability of the discrete DM formulation. We assume that a pair of finite dimensional spaces $\hat{V}_h, \hat{Q}_h$ satisfies the following condition:

For any $v_h \in \hat{V}_h$ exists $q_h^v \in \hat{Q}_h$ such that

\begin{align}
\text{div}q_h^v &= v_h \quad \text{(compatibility),} \\
\|q_h^v\| &\leq C\|v_h\| \quad \text{(stability).}
\end{align} 

(9.14) (9.15)
Discrete Inf-Sup condition

From (9.14) and (9.15), it follows that

\[ \inf_{v_h \in \hat{V}_h} \sup_{q_h \in \hat{Q}_h} \frac{\int_{\Omega} v_h \text{div} q_h \, dx}{\|v_h\| \|q_h\|_{\text{div}}} \geq C > 0 \]

Indeed,

\[ \sup_{q_h \in \hat{Q}_h} \frac{\int_{\Omega} v_h \text{div} q_h \, dx}{\|v_h\| \|q_h\|_{\text{div}}} \geq \frac{\int_{\Omega} v_h \text{div} q^\vee_h \, dx}{\|v_h\| \|q^\vee_h\|_{\text{div}}} = \frac{\|v_h\|}{\|q_h\|_{\text{div}}} \geq \frac{1}{\sqrt{1 + C^2}}. \]

Now, we refer to known results on the solvability of DMM, that can be summarized as follows: if the triangulations are "regular" and the discrete Inf-Sup condition holds, then the discrete formulation has a unique solution.
Projection type estimate for the dual problem

Since \( p \) is a maximizer, i.e.,
\[
-\frac{1}{2} \|q\|_*^2 \leq -\frac{1}{2} \|p\|_*^2 \quad \forall q \in Q_f,
\]
we find that
\[
\int_{\Omega} A^{-1}p \cdot q \, dx = 0 \quad \forall q \in Q_0,
\]
where \( Q_0 \) is the space of solenoidal functions. Therefore, for any \( q \in Q_f \),
\[
\frac{1}{2} \|q - p\|_*^2 = \frac{1}{2} \|q\|_*^2 - \frac{1}{2} \|p\|_*^2 + \int_{\Omega} A^{-1}p \cdot (p - q) \, dx =
\]
\[
= \frac{1}{2} \|q\|_*^2 - \frac{1}{2} \|p\|_*^2.
\]
Let $Q_{fh} = Q_f \cap \hat{Q}_h$. Note that $p_h \in Q_{fh}$ is also the maximizer of $-\frac{1}{2} ||q_{fh}||^2$ on $Q_{fh}$, so that

$$
\frac{1}{2} ||p_h - p||^2_* = \frac{1}{2} ||p_h||^2_* - \frac{1}{2} ||p||^2_* \leq \frac{1}{2} ||q_{fh}||^2_* - \frac{1}{2} ||p||^2_* = \frac{1}{2} ||q_{fh} - p||^2_* \quad \forall q_{fh} \in Q_{fh}.
$$

Thus, we arrive at the first projection estimate

$$
||p - p_h||_* \leq \inf_{q_{fh} \in Q_{fh}} ||p - q_{fh}||_*.
$$

(9.16)
However, this projection error estimate has an obvious drawback. It is applicable only for a very narrow class of approximations: conforming (internal) approximations of the set $Q_f$.

To obtain an estimate for a wider class, we first derive one auxiliary result.
A Modified DM problem

Take $\tilde{f} = \text{div}(\hat{q}_h - p)$ where $\hat{q}_h \in \hat{Q}_h$ and solve the modified DM problem

$$
\int_{\Omega} \left( A^{-1} \hat{p}_h \cdot \hat{q}_h + \hat{u}_h \text{div} \hat{q}_h \right) \, dx = 0 \quad \forall \hat{q}_h \in \hat{Q}_0 h, \quad (9.17)
$$

$$
\int_{\Omega} (\text{div} \hat{p}_h + \tilde{f}) \hat{v}_h \, dx = 0 \quad \forall \hat{v}_h \in \hat{V}_h. \quad (9.18)
$$

Under the assumptions made $\tilde{f} \in P^0(T)$, the above DM problem is solvable, and

$$
\| \hat{p}_h \|^2_2 + \int_{\Omega} \hat{u}_h \text{div} \hat{p}_h \, dx = 0,
$$

$$
\| \hat{p}_h \|^2_* + \| \hat{u}_h \|_{\text{div} \hat{p}_h} = \| \hat{u}_h \| \| \tilde{f} \|.
$$
From here, we observe that
\[
\bar{c}_1 \| \hat{p}_h \|^2 \leq \| \hat{p}_h \|^* \leq \| \hat{u}_h \| \| \hat{f} \|. \tag{9.19}
\]
By (9.14) and (9.15) we conclude that for \( \hat{u}_h \) we can find \( \bar{q}_h \) in \( \hat{Q}_h \) such that
\[
\text{div} \bar{q}_h + \hat{u}_h = 0 \quad \text{and} \quad \| \bar{q}_h \| \leq C \| \hat{u}_h \|
\]
Use \( \bar{q}_h \) in the first identity (9.17). We have,
\[
\int_{\Omega} \left( A^{-1} \hat{p}_h \cdot \bar{q}_h + \hat{u}_h \text{div} \bar{q}_h \right) \, dx = 0
\]
Thus,
\[
\| \hat{u}_h \|^2 = \int_{\Omega} \hat{u}_h \text{div} \bar{q}_h \leq \| \hat{p}_h \|^* \| \bar{q}_h \|^*
\]
\[
\leq \bar{c}_2 \| \hat{p}_h \|^* \| \bar{q}_h \| \leq \bar{c}_2 C \| \hat{p}_h \|^* \| \hat{u}_h \|.
\]
We observe that

$$
\| \hat{u}_h \| \leq \bar{c}_2 C \| \hat{p}_h \|_* .
$$

(9.20)

Now, we use (9.19) and obtain

$$
\| \hat{p}_h \|_*^2 \leq \| \hat{u}_h \| \| \tilde{f} \| \leq \bar{c}_2 C \| \hat{p}_h \|_* \| \tilde{f} \|. 
$$

so that

$$
\| \hat{p}_h \|_* \leq \bar{c}_2 C \| \tilde{f} \|. 
$$

(9.21)

Hence,

$$
\| \hat{p}_h \|^2 = \| \hat{p}_h \|^2 + \| \text{div} \hat{p}_h \|^2 \leq (1 + \frac{c^2}{c^2_1} C^2) \| \tilde{f} \|^2 .
$$

(9.22)
We note that the estimates (9.20), (9.21), and (9.22) show that the modified DM problem is stable, i.e. its solutions \((\hat{p}_h^f, \hat{u}_h^f)\) are bounded by the problem data uniformly with respect to \(h\).

If replace \(\tilde{f}\) by \(f\), then we can derive the same stability estimate for the functions \((\hat{p}_h, \hat{u}_h)\) that present an approximate solution of the original DM problem.
Now, we return to the projection error estimates. As we have seen

\[ \| p - p_h \| \leq \inf_{q_f \in Q_{f_h}} \| p - q_f \|. \]

This estimate did not satisfy us because the set \( Q_{f_h} \) is difficult to construct. To avoid this drawback, we apply the following procedure.

Let \( \eta_h = \hat{p}_h^f + \hat{q}_h \), where \( \hat{q}_h \) is an arbitrary element of \( \hat{Q}_h \).

We have,

\[ \text{div}\eta_h = \text{div}\hat{p}_h^f + \text{div}\hat{q}_h = -\tilde{f} + \text{div}\hat{q}_h = \]

\[ = \text{div}(p - \hat{q}_h) + \text{div}\hat{q}_h = \text{div}p = -f. \]

Therefore, \( \eta_h \in Q_f \)
Now, we recall the projection inequality and substitute in it $\eta_h$:

$$\| p - p_h \|_* \leq \| p - \eta_h \|_* = \| p - \hat{p}_h - \hat{q}_h \|_* \leq \| p - \hat{q}_h \|_* + \| \hat{p}_h \|_*$$

Note that in the case considered $\text{div}(p - p_h) = 0$, so that

$$\| p - p_h \|_{\text{div}} = \| p - p_h \| \leq \frac{1}{\bar{c}_1} \| p - p_h \|_*.$$

Therefore, by means of (9.21) we obtain

$$\| p - p_h \|_{\text{div}} \leq \frac{1}{\bar{c}_1}(\| p - \hat{q}_h \|_* + \| \hat{p}_h \|_*)$$

$$\leq \frac{1}{\bar{c}_1}(\| p - \hat{q}_h \|_* + \bar{c}_2 C \| \tilde{f} \|).$$

But $\tilde{f} = \text{div}(p - \hat{q}_h)$. 
Thus, we arrive at the estimate

\[ \| p - p_h \|_{\text{div}} \leq \frac{1}{\bar{c}_1} \left( \| p - \hat{q}_h \|_\ast + \bar{c}_2 C \| \text{div}(p - \hat{q}_h) \| \right) \quad \forall \hat{q}_h \in \hat{Q}_h. \]

and, therefore,

\[ \| p - p_h \|_{\text{div}} \leq \bar{C}_p \inf_{\hat{q}_h \in \hat{Q}_h} \left\{ \| p - \hat{q}_h \|_\ast + \| \text{div}(p - \hat{q}_h) \| \right\}. \quad (9.23) \]

where \( \bar{C}_p \) depends on \( C, \bar{c}_1, \) and \( \bar{c}_2 \) and does not depend on \( h \).
Projection type error estimates for $\hat{u} - \hat{u}_h$

We have

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

Since $\hat{Q}_h \subset Q$, we also have

$$\int_\Omega \left( A^{-1} p \cdot \hat{q}_h + u \text{div} \hat{q}_h \right) \, dx = 0.$$ 

From here, we observe that

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

\[ [u]_T = \frac{1}{|T|} \int_T u \, dx, \quad [u]_h(x) = [u]_{T_i} \text{ if } x \in T_i. \]

Since \( \text{div} \hat{q}_h \) is constant on each \( T_i \), we rewrite the relation as follows:

\[
\int_{\Omega} \left( A^{-1}(\hat{p}_h - p) \cdot \hat{q}_h + (\hat{u}_h - [u]_h) \text{div} \hat{q}_h \right) \, dx = 0 \quad \forall \hat{q}_h \in \hat{Q}_h.
\]

Note that \( [u]_h \in \hat{V}_h \) and \( \bar{u}_h := \hat{u}_h - [u]_h \in \hat{V}_h \) Now, we exploit the compatibility and stability conditions (9.14) and (9.15) again. For \( \bar{u}_h \) one can find \( q'_h \in \hat{Q}_h \) such that

\[
\text{div} q'_h + \bar{u}_h = 0 \quad \text{and} \quad \| q'_h \| \leq C \| \bar{u}_h \|.
\]
Let us use this function $q'_h$ in the integral relation. We have

$$
\int_\Omega \left( A^{-1}(\hat{p}_h - p) \cdot q'_h + \bar{u}_h \text{div} q'_h \right) dx = 0.
$$

From here, we conclude that

$$
\| \bar{u}_h \|^2 = \left| \int_\Omega A^{-1}(\hat{p}_h - p) \cdot q'_h \right| \leq

\leq \| \hat{p}_h - p \|_* \| q'_h \|_* \leq C \bar{c}_2 \| \hat{p}_h - p \|_* \| \bar{u}_h \|.
$$

Thus,

$$
\| \bar{u}_h \| = \|[u]_h - \hat{u}_h\| \leq C \bar{c}_2 \| \hat{p}_h - p \|_*.
$$
Since

\[ \| u - \hat{u}_h \| \leq \| u - [u]_h \| + \| [u]_h - \hat{u}_h \| \leq \]
\[ \leq \| u - [u]_h \| + C \bar{c}_2 \| \hat{p}_h - p \| ^* \]

Note that by the definition of \([u]_h\)

\[ \| u - [u]_h \| \leq \| u - v_h \| \quad \forall v_h \in \hat{V}_h. \]

From here, we observe that

\[ \| u - \hat{u}_h \| \leq C \bar{c}_2 \| \hat{p}_h - p \| ^* + \inf_{v_h \in \hat{V}_h} \| u - v_h \| \]

Recall that

\[ \| p - p_h \| ^* \leq \| p - \hat{q}_h \| ^* + \| \hat{p}_h^f \| ^* \leq \]
\[ \| p - \hat{q}_h \| ^* + \bar{c}_2 C \| \text{div}(p - \hat{q}_h) \|. \]
Then, we arrive at the projection type error estimate for the primal variable

\[ \| u - \hat{u}_h \| \leq \]

\[ \leq C_u \inf_{\hat{q}_h \in Q_h} \left\{ \| p - \hat{q}_h \|_\ast + \| \text{div}(p - \hat{q}_h) \| + \right. 
\]

\[ + \left. \inf_{v_h \in \hat{V}_h} \| u - v_h \| \right\}, \quad (9.24) \]

where \( C_u \) depends on \( C, \bar{c}_1, \) and \( \bar{c}_2 \) and does not depend on \( h \). Estimates \((9.23)\) and \((9.25)\) lead to a qualified a priori convergence estimates provided that the solution possesses proper regularity.
A posteriori estimates for the primal mixed formulation

Further analysis follows the lines of the paper
A posteriori estimates for the mixed formulation are based on the relation that we have already derived:

\[ \| p - q \|^2_\ast + \| \nabla (u - v) \|^2 = 2(J(v) - I^*(q)), \]  
\[ (9.25) \]

where \( q \in Q_\ell \) and \( v \in V_0 + u_0 \).
Since the difference of the functionals in the right–hand side can be estimated by the known way, we arrive at the estimate

$$
\| p - q \|_2^2 + \| \nabla (u - v) \|_2^2 \leq 2(1 + \beta)D(\nabla v, y) + \left(1 + \frac{1}{\beta}\right)C^2 \left(\| \text{div} y + f \|_2^2 + \| y \cdot n - F \|_{\partial_2 \Omega}^2 \right), \quad (9.26)
$$

where $y \in \widehat{Q}^+$, $q \in Q_\ell$ and $v \in V_0 + u_0$ are arbitrary functions and $\beta$ is any positive number.
Thus, for the error in the primal variable we have

\[ \| \nabla (u - u_h) \|^2 \leq 2(1 + \beta)D(\nabla u_h, y) \]
\[ + \left( 1 + \frac{1}{\beta} \right) C^2 \left( \| \text{div} y + f \|^2 + \| y \cdot n - F \|^2_{\partial_2 \Omega} \right). \quad (9.27) \]

where \( C \) is a constant in the inequality

\[ \| w \|^2 + \| w \|^2_{\partial_2 \Omega} \leq C^2 \| \nabla w \|^2 \quad \forall w \in V_0. \]
A posteriori estimate for the dual variable

By using the general estimate derived in Lecture 4, we find that

\[
\| p - p_h \| \leq \sqrt{2} D^{1/2} (\nabla v, y) + \| y - p_h \| + 2C \left( \| \text{div} y + f \|^2 + \| y \cdot n - F \|^2 \right)^{1/2}.
\] (9.28)

Here \( v \) is an arbitrary function from \( V_0 + u_0 \) and \( y \) is an arbitrary function from \( \hat{Q}^+ \). If \( y = A \nabla u \) and \( v = u \), then the right-hand side of (9.37) coincides with the left-hand side, i.e. is exact in the sense that there exist such ”free variables” that the inequality holds as the equality.
A directly computable upper bound of $\| p - p_h \|_*$ is given by (9.37), if we set

$$ v = u_h, \quad \text{and} \quad y = G_h p_h, $$

where $G_h : Q_h \to \hat{Q}^+$ is a certain projection operator (some examples such operators has been already discussed in the previous lectures).

We have

$$ \| p - p_h \|_* \leq \sqrt{2} D^{1/2} (\nabla u_h, G_h p_h) + \| G_h p_h - p_h \|_* $$

$$ + 2C \left( \| \text{div} G_h p_h + f \|^2 + \| G_h p_h \cdot n - F\|^2 \right)^{1/2}. $$
Projection from $Q_h$ onto $\hat{Q}^+$

If $p_h$ is a piecewise-constant vector field on a simplicial mesh $\mathcal{T}_h$, then, Raviart–Thomas elements (e.g., $RT^0$–elements) can be used in order to define the mapping $\mathcal{G}$.

Assume that the $\Omega$ has a polygonal boundary, and the latter is exactly matched by the triangulation $\mathcal{T}_h$. Let $T_i$ and $T_j$ be two neighboring simplexes with the common edge $E_{ij}$. Let $q_h$ be a piecewise constant vector-valued function that has the values $q_i$ and $q_j$ on $T_i$ and $T_j$ respectively. Let $E_{ij}$ be the common edge with the unit normal $n_{ij}$ oriented from $T_i$ to $T_j$ if $i > j$.

**How to define the common value $\tilde{q}_{ij} \cdot n_{ij}$ on $E_{ij}$?**
One possible option is as follows:

\[
\tilde{q}_{ij} \cdot n_{ij} = \frac{1}{2} (q_i + q_j) \cdot n_{ij},
\]

Another option is

\[
\tilde{q}_{ij} \cdot n_{ij} = \frac{|T_i|q_i + |T_j|q_j}{|T_i| + |T_j|} \cdot n_{ij},
\]

where \(|T_i|\) and \(|T_j|\) are the areas of \(T_i\) and \(T_j\). We repeat this procedure for all internal edges of \(T_h\).

If \(E_{i0} \in \partial_1 \Omega\), then we set \(\tilde{q}_{i0} \cdot n_{i0} = q_{i0} \cdot n_{i0}\). If \(E_{i0} \in \partial_2 \Omega\), then

\[
\tilde{q}_{i0} \cdot n_{i0} = \frac{1}{|E_{i0}|} \int_{E_{i0}} F \, ds.
\]

Here \(|E_{i0}|\) is the length of the edge \(E_{i0}\).
Thus, all the normal components $\tilde{q}_{ij} \cdot n_{ij}$ on internal and external edges are defined. By prolongation inside all $T_i$, with the help of $RT_0$-approximations we obtain the function a piecewise affine function, which has continuous normal components at all the edges and piecewise constant normal components on $\partial \Omega$. Therefore, we, in fact, have constructed a mapping $q_h \to \tilde{q}_h$ such that

$$\tilde{q}_h = G_h q_h \in \hat{Q}^+.$$
An a posteriori estimate for the flux $\hat{p}_h$ readily follows from the general estimate

$$\frac{1}{2} \| y - p \|_*^2 \leq (1 + \gamma) \left( 1 + \frac{1}{\gamma} + \frac{1}{\beta \gamma} \right) \| \ell + \Lambda^* y \|^2 +$$

$$+ (1 + \beta) \left( 1 + \frac{1}{\gamma} \right) D(\Lambda v, y).$$

that we have derived in Lecture 5. We set $y = \hat{p}_h \in \hat{Q}^+$. Since $\hat{p}_h$ is a piecewise polynomial function, it has a summable trace on $\partial_2 \Omega$. Then, we estimate $\| \ell + \Lambda^* y \|$ from above in the same way we did it in Lecture 6. Minimization with respect to $\gamma$ and $\beta$ leads to the estimate
\[
\| p - \hat{p}_h \|_* \leq \sqrt{2}D^{1/2}(\nabla v, \hat{p}_h) + \\
+ 2C \left( \| \text{div}\hat{p}_h + f \|^2 + \| \hat{p}_h \cdot n - F \|_{\partial_2 \Omega}^2 \right)^{1/2},
\]

where \( v \) is an arbitrary function from \( V_0 + u_0 \).
For the sake of simplicity we assume that $\Omega$ is a polygonal domain decomposed into a regular collection of simplexes. If $\hat{p}_h$ is constructed by means of $RT_0$-elements, then

$$
\int_{\Omega} (\text{div}\hat{p}_h + f)w_h \, dx = 0 \quad \forall w_h \in \hat{V}_h \subset \hat{V},
$$

(9.30)

where the subspace $\hat{V}_h$ contains piecewise constant functions. Therefore, on each element $T_i$

$$
\text{div}\hat{p}_h = -\frac{1}{|T_i|} \int_{T_i} f \, dx.
$$

(9.31)

Let us define by $[f]$ the function that belongs to $\hat{V}_h$ and whose values on $T_i$ coincide with the mean values of $f$ on $T_i$. Then, we have

$$
\text{div}\hat{p}_h = -[f] \quad \text{on every} \quad T_i.
$$
Remark. We observe that estimate (9.30) is valid for any approximate flux $\hat{p}_h$ from $\hat{Q}^+$. If $\hat{p}_h$ were in the narrower set $\hat{Q}_F$ (as it is supposed to be in the discrete dual mixed method) the last norm in (9.30) would be identically zero. It cannot, however, be expected, when $\hat{p}_h$ is constructed in the space $RT_0$, unless the function $F$ is a constant on $\partial_2\Omega$. 
The problem of taking into account the essential boundary condition for the flux variable

$$\hat{p} \cdot n = F \quad \text{on} \quad \partial_2 \Omega$$


Since (9.30) still works for such approximations of the flux, we propose a simple modification of the discrete dual method, particularly suited for the lowest-order Raviart-Thomas approximation.
Namely, instead of requiring \( \hat{p}_h \in \hat{Q}_F \), we impose a weaker condition
\[
\hat{p}_h \cdot n \bigg|_{E_{i0}} = \frac{1}{|E_{i0}|} \int_{E_{i0}} F \, ds
\]
(9.32)
on every edge \( E_{i0} \in \partial_2 \Omega \). The space of test functions \( \hat{Q}_{0h} \subset \hat{Q}_0 \) will obviously consist of the \( \text{RT}_0 \)-approximations \( \hat{q}_h \) such that \( \hat{q}_h \cdot n = 0 \) on each edge \( E_{i0} \in \partial_2 \Omega \).

If now we denote by \([F]\) the piecewise constant function defined on the set of edges forming \( \partial_2 \Omega \) and whose value on every edge \( E_{i0} \in \partial_2 \Omega \) is equal to the mean value of \( F \) on that edge, we can write that \( \hat{p}_h \cdot n = [F] \) for all \( E_{i0} \in \partial_2 \Omega \).

As a result, we obtain from (9.30)
\[
\| p - \hat{p}_h \|_* \leq \sqrt{2D}^{1/2} (\nabla \nu, \hat{p}_h) + 2C \left( \| f - [f] \|^2 + \| F - [F] \|^2_{\partial_2 \Omega} \right)^{1/2}.
\]
(9.33)
The question that now arises is how to choose in (9.33) the function $v \in V_0 + u_0$. The simplest way is to use the function $\hat{u}_h \in \hat{V}_h$ available from the solution of the discrete dual mixed problem and to construct a suitable projection operator $P_h : \hat{V}_h \rightarrow V_0 + u_0$. Again, the projection can be easily accomplished with a simple averaging.

**Projection from $\hat{V}_h$ onto $V_0 + u_0$.**

In order to find $v \in V_0 + u_0$, it is sufficient to find $w \in V_0$ in the representation $v = w + u_0$ (the function $u_0$ is given). Using the computed piecewise-constant function $\hat{u}_h$, we define $w_h \in V_0$ as follows.
We set

$$w_h(x_k) = \frac{\sum_{s=1}^{N_k} |T_s^{(k)}| \cdot \hat{u}_h|_{T_s^{(k)}}}{\sum_{s=1}^{N_k} |T_s^{(k)}|} - u_0(x_k)$$

(9.34)

for any internal node $x_k$ and when $x_k \in \partial_2 \Omega$. Here $T_s^{(k)}$, $s = 1, N_k$, are the elements containing the vertex $x_k$, and we have assumed that the function $u_0$ has a sufficient regularity, so that its point values are defined.

If the node $x_k \in \partial_1 \Omega$, we simply set $w_h(x_k) = 0$.

Thus, using the nodal values of $w_h$ and the piecewise-linear continuous finite element approximation on the mesh $T_h$ we define the function

$$w_h + u_0 = P_h \hat{u}_h \in V_0 + u_0.$$
Hence, from (9.33) one obtains

\[ \bar{c}_1 \left\| \mathbf{p} - \hat{\mathbf{p}}_h \right\| \leq \left\| \mathbf{p} - \hat{\mathbf{p}}_h \right\|_{\ast} \leq \sqrt{2D^{1/2} \left( \nabla (P_h \hat{\mathbf{u}}_h), \hat{\mathbf{p}}_h \right) + 2C \left( \left\| \mathbf{f} - [\mathbf{f}] \right\|^2 + \left\| \mathbf{F} - [\mathbf{F}] \right\|_{\partial_2 \Omega}^2 \right)^{1/2}}, \]

(9.35)

which, together with the obvious relation

\[ \left\| \text{div} (\hat{\mathbf{p}} - \hat{\mathbf{p}}_h) \right\| = \left\| - \mathbf{f} - \text{div} \hat{\mathbf{p}}_h \right\| = \left\| \mathbf{f} - [\mathbf{f}] \right\| \]

leads to the upper bound for \( \left\| \hat{\mathbf{p}} - \hat{\mathbf{p}}_h \right\|_{\text{div}} \):
Let \((\hat{u}, \hat{p}) \in \hat{V} \times \hat{Q}_F\) be the exact solution of the dual mixed problem and \((\hat{u}_h, \hat{p}_h) \in \hat{V}_h \times \hat{Q}_{Fh}\) the solution of the discrete dual mixed problem with \(\hat{Q}_{Fh}\) being the Raviart-Thomas space \(\text{RT}^0\).

Then, the following estimate holds true:

\[
\|\hat{p} - \hat{p}_h\|_{\text{div}} \leq \|A \nabla (\mathcal{P}_h \hat{u}_h) - \hat{p}_h\|_{\ast} + (2C + 1)\|f - [f]\| + 2C\|F - [F]\|_{\partial^2 \Omega},
\]  

(9.36)

where \(\mathcal{P}_h : \hat{V}_h \rightarrow V_0 + u_0\) is the projection (averaging) operator introduced above and \([f]\) and \([F]\) are the averaged functions.

**Remark.** The first and the second terms in (9.36), being computed elementwise, can serve as local error indicators.
A sharper estimate can be obtained by the minimization of the Majorant with respect to $\mathbf{v}$. Here, we can restrict ourselves to certain subspace $V_h$, i.e.,

$$\| \hat{p} - \hat{p}_h \|_{\text{div}} \leq \inf_{v_h \in V_h} \| A \nabla (v_h) - \hat{p}_h \|_{*} + (2C + 1) \| f - [f] \| + 2C \| F - [F] \|_{\partial_2 \Omega}.$$  

(9.37)
By (9.28) we can also the squared norm of the error of the averaged solution \( P_h \hat{u}_h \) using the computed flux approximation \( \hat{p}_h \):

\[
\| \nabla (u - P_h \hat{u}_h) \|^2 \leq 2(1 + \beta) D(\nabla(P_h \hat{u}_h), \hat{p}_h) \\
+ \left( 1 + \frac{1}{\beta} \right) C^2 (\| f - [f] \|^2 + \| F - [F] \|^2_{\partial^2 \Omega}), \quad (9.38)
\]

where \( \beta > 0 \) is an arbitrary number that can be used to minimize the right-hand side of (9.38) and to obtain the estimate for the norm of the error.

A sharper estimate may be obtained, if one spends some time on the minimization of the right-hand side of (9.38) with respect to the dual variable \( y \) over some finite-dimensional subspace of \( \hat{Q}^+ \).
Remark.
If one has the solutions of both the primal and the dual mixed problems, the flux approximation $\hat{p}_h$ can be substituted into (9.28) to immediately yield the error estimate for the primal variable (which is the most important in the primal mixed method), while the approximation $u_h$ can be used in (9.36) to bring the error estimate for the dual variable (which is the most important in the dual mixed method).
Lecture 10.
A POSTERIORI ERROR ESTIMATES FOR ITERATION METHODS
Lecture plan

- Banach fixed point theorem;
- Two-sided error estimates by A. Ostrovski;
- Advanced two-sided estimates;
- Applications to matrix equations;
- Positivity methods and a posteriori error bounds.
- Applications to integral equations;
- Applications to ordinary differential equations.
Fixed point theorem

Consider a Banach space \((X, d)\) and a continuous operator

\[ \mathcal{T} : X \to X. \]

**Definition**

A point \(x \circ\) is called a fixed point of \(\mathcal{T}\) if

\[ x \circ = \mathcal{T}x \circ. \]  \hfill (10.1)

Approximations of a fixed point are usually constructed by the iteration sequence

\[ x_i = \mathcal{T}x_{i-1} \quad i = 1, 2, \ldots \]  \hfill (10.2)
Contractive mappings

Two basic tasks:
(a) find the conditions that guarantee convergence of \( x_i \) to \( x_\oplus \),
(b) find computable estimates of the error \( e_i = d(x_i, x_\oplus) \).

These problems possess solutions provided, that \( \mathfrak{T} \) is subject to the following additional condition.

**Definition**

An operator \( \mathfrak{T} : X \rightarrow X \) is called \( q \)-contractive on a set \( S \subset X \) if there exists a positive real number \( q \) such that the inequality

\[
d(\mathfrak{T}x, \mathfrak{T}y) \leq q d(x, y) \tag{10.3}
\]

holds for any elements \( x \) and \( y \) of the set \( S \).
Theorem (S. Banach)

Let \( \mathcal{F} \) be a \( q \)-contractive mapping of a closed nonempty set \( S \subset X \) to itself with \( q < 1 \). Then, \( \mathcal{F} \) has a unique fixed point in \( S \) and the sequence \( x_i \) obtained by (10.2) converges to this point.
Proof. It is easy to see that

\[ d(x_{i+1}, x_i) = d(\mathcal{T}x_i, \mathcal{T}x_{i-1}) \leq qd(x_i, x_{i-1}) \leq \ldots \leq q^i d(x_1, x_0). \]

Therefore, for any \( m > 1 \) we have

\[
d(x_{i+m}, x_i) \leq \]

\[
\leq d(x_{i+m}, x_{i+m-1}) + d(x_{i+m-1}, x_{i+m-2}) + \ldots + d(x_{i+1}, x_i) \leq \]

\[
\leq q^i(q^{m-1} + q^{m-2} + \ldots + 1)d(x_1, x_0). \quad (10.4)
\]
Since

\[ \sum_{k=0}^{m-1} q^k \leq \frac{1}{1 - q}, \]

(10.4) implies the estimate

\[ d(x_{i+m}, x_i) \leq \frac{q^i}{1 - q} d(x_1, x_0). \]  

(10.5)

Let \( i \to \infty \), then the right-hand side of (10.5) tends to zero, so that \( \{x_i\} \) is a Cauchy sequence. It has a limit in \( y \in X \).
Then, $d(x_i, y) \to 0$ and

$$d(\mathcal{T}x_i, \mathcal{T}y) \leq qd(x_i, y) \to 0$$

so that $d(\mathcal{T}x_i, \mathcal{T}y) \to 0$ and $\mathcal{T}x_i \to \mathcal{T}y$. Pass to the limit in (10.2) as $i \to +\infty$. We observe that

$$\mathcal{T}y = y.$$

Hence, any limit of such a sequence is a fixed point.
It is easy to prove that a fixed point is unique. Assume that there are two different fixed points $x^1$ and $x^2$, i.e.

$$x^k = x^k, \quad k = 1, 2.$$ 

Therefore,

$$d(x^1, x^2) = d(Tx^1, Tx^2) \leq qd(x^1, x^2).$$ 

But $q < 1$, and thus such an inequality cannot be true.
A priori convergence estimate

Let

\[ e_j = d(x_j, x_\circ) \]

denote the error on the \( j \)-th step. Then

\[ e_j = d(\mathcal{T}x_{j-1}, \mathcal{T}x_\circ) \leq qe_{j-1} \leq q^je_0. \]

This estimate gives a certain presentation on that how the error decreases. However, as we will see later, this a priori upper bound may be rather coarse.
A posteriori estimates

The proposition below furnishes upper and lower estimates of \( e_j \), which are easy to compute provided, that the number \( q \) (or a good estimate of it) is known.

**Theorem (A. Ostrowski)**

Let \( \{x_j\}_{j=0}^{\infty} \) be a sequence obtained by the iteration process (10.2) with a mapping \( \mathcal{T} \) satisfying the condition (10.3). Then, for any \( x_j, j > 1 \), the following estimate holds:

\[
M_j^\ominus := \frac{1}{1+q} d(x_{j+1}, x_j) \leq e_j \leq M_j^\oplus := \frac{q}{1-q} d(x_j, x_{j-1})
\]

(10.6)
Proof. The upper estimate in (10.6) follows from (10.5). Indeed, put \(i = 1\) in this relation. We have

\[
d(x_{1+m}, x_1) \leq \frac{q}{1 - q}d(x_1, x_0).
\]

Since \(x_{1+m} \to x_\circ\) as \(m \to +\infty\), we pass to the limit with respect to \(m\) and obtain

\[
d(x_\circ, x_1) \leq \frac{q}{1 - q}d(x_1, x_0).
\]

We may view \(x_{j-1}\) as the starting point of the sequence. Then, in the above relation \(x_0 = x_{j-1}\) and \(x_1 = x_j\) and we arrive at the following upper bound of the error:

\[
d(x_\circ, x_j) \leq \frac{q}{1 - q}d(x_j, x_{j-1}).
\]
The lower bound of the error follows from the relation

\[ d(x_j, x_{j-1}) \leq d(x_j, x_\circ) + d(x_{j-1}, x_\circ) \leq (1 + q)d(x_{j-1}, x_\circ), \]

which shows that

\[ d(x_{j-1}, x_\circ) \geq \frac{1}{1 + q}d(x_j, x_{j-1}). \]

Note that

\[ \frac{M_j^\oplus}{M_j^\ominus} = \frac{q(1 + q)}{1 - q} \frac{d(x_j, x_{j-1})}{d(x_{j+1}, x_j)} \geq \frac{1 + q}{1 - q}, \]

we see that the efficiency of the upper and lower bounds given by (10.6) deteriorates as \( q \to 1 \).
Remark. If $X$ is a normed space, then

$$d(x_{j+1}, x_j) = \| R(x_j) \|,$$

where

$$R(x_j) := \Xi x_j - x_j$$

is the residual of the basic equation (10.1). Thus, the upper and lower estimates of errors are expressed in terms of the residuals of the respective iteration equation computed for two neighbor steps:

$$\frac{1}{1 + q} \| R(x_j) \| \leq e_j = d(x_j, x_\circ) \leq \frac{q}{1 - q} \| R(x_{j-1}) \|.$$
In the iteration methods, it is often easier to analyze the operator

\[ \mathcal{T} = T^n := \underbrace{TT\ldots T}_{n \text{ times}} \]

where \( T \) is a certain mapping.

**Proposition (1)**

*Let \( T : S \rightarrow S \) be a continuous mapping such that \( \mathcal{T} \) is a \( q \)-contractive mapping with \( q \in (0, 1) \). Then, the equations

\[ x = Tx \quad \text{and} \quad x = \mathcal{T}x \]

have one and the same fixed point, which is unique and can be found by the above described iteration procedure.*
Proof. By the Banach Theorem, we observe that the operator $\mathcal{T}$ has a unique fixed point $\xi\odot$.
Let us show that $\xi\odot$ is a fixed point of $T$, we note that

$$T\xi\odot = T(\mathcal{T}\xi\odot) = T\mathcal{T}^2\xi\odot = \ldots$$

$$= T\mathcal{T}^i\xi\odot = T^{(1+i)}\xi\odot = T^{\text{in}}T\xi\odot. \quad (10.7)$$

Denote $x_0 = T\xi\odot$. By (10.7) we conclude that for any $i$

$$T\xi\odot = \mathcal{T}^i x_0. \quad (10.8)$$

Passing to the limit on the right-hand side in (10.8), we arrive at the relation $T\xi\odot = \xi\odot$, which means that $\xi\odot$ is a fixed point of the operator $T$. 

S. Repin

RICAM, Special Radon Semester, Linz, 2005.
Let $\tilde{x}_\circ$ be a fixed point of $T$. Then,

$$\tilde{x}_\circ = T^2\tilde{x}_\circ = ... = T^n\tilde{x}_\circ = \mathcal{S}\tilde{x}_\circ$$

and we observe that $\tilde{x}_\circ$ is a fixed point of $T$. Since the saddle point of $\mathcal{S}$ exists and is unique, we conclude that

$$x_\circ = \tilde{x}_\circ.$$

**Remark.** This assertion may be practically useful if it is not possible to prove that $T$ is $q$–contractive, but this fact can be established for a certain power of $T$. 

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
We can derive more accurate bounds of errors if we use more terms of the sequence \( \{x_j\} \).

Indeed,

\[
d(x_j, x_\odot) \leq d(x_j, x_{j+1}) + d(x_{j+1}, x_\odot) \leq d(x_j, x_{j+1}) + \frac{q}{1-q}d(x_j, x_{j+1}),
\]

and we obtain another upper bound

\[
d(x_j, x_\odot) \leq \frac{1}{1-q}d(x_j, x_{j+1}). \tag{10.9}
\]

It estimates the error on \( j \)-th step by \( x_j \) and \( x_{j+1} \).
Which bound is sharper: \((10.9)\) or \(M^j_\oplus\)?

Since

\[
\frac{1}{1 - q} d(x_j, x_{j+1}) \leq \frac{q}{1 - q} d(x_{j-1}, x_j),
\]

we observe that this bound is sharper than \(M^j_\oplus\).

Obviously, \((10.9)\) can also be applied to any subsequence of \(\{x_j\}\). For example, we can take \(\{x_{\ell s}\}, s = 0, 1, 2...\) with some fixed \(\ell\). In this case, we obtain various upper bounds of \(d(x_j, x_\circ)\) computed on the basis of some terms of the sequence \(\{x_j\}\):

\[
d(x_j, x_\circ) \leq M^j_\oplus \ell := \frac{1}{1 - q^\ell} d(x_j, x_{j+\ell}).
\]

Note that the right-hand side of this estimate tends to \(d(x_j, x_\circ)\) as \(\ell \to +\infty\). Thus, for a sufficiently large \(\ell\) the bound will be accurate even if \(q\) is close to 1.
The lower estimates can be improved by similar arguments. We have the estimate

\[ d(x_j, x_\circ) \geq M^j,\ell := \frac{1}{1 + q^\ell} d(x_j, x_{j+\ell}) \]

whose right-hand side also tends to the exact value of the error as \( \ell \to +\infty \).

Let \( L \) be a given number that indicates the number of successive elements used in the evaluation of error bounds for \( x_j \). Compute the quantities

\[
\bar{M}^j,L := \sup_{\ell=1,2,...,L} \left\{ \frac{1}{1 + q^\ell} d(x_j, x_{j+\ell}) \right\}, \\
\tilde{M}^j,L := \inf_{\ell=1,2,...,L} \left\{ \frac{1}{1 - q^\ell} d(x_j, x_{j+\ell}) \right\}. \tag{10.10} \tag{10.11}
\]

These upper and lower bounds are the sharper, the greater is \( L \).
Another sequence of upper bounds follows from the relation

\[ d(x_j, x_{\odot}) \leq d(x_j, x_{j+1}) + d(x_{j+1}, x_{\odot}) \leq \]

\[ \leq \mathcal{M}^{j,2}_{\oplus}(x_j, x_{j+1}, x_{j+2}) := d(x_j, x_{j+1}) + \frac{1}{1 - q} d(x_{j+1}, x_{j+2}). \quad (10.12) \]

Note that

\[ \mathcal{M}^{j,2}_{\oplus} \leq d(x_j, x_{j+1}) + \frac{q}{1 - q} d(x_j, x_{j+1}) = \]

\[ = \frac{1}{1 - q} d(x_j, x_{j+1}) := \mathcal{M}^{j,1}_{\oplus}. \]

Similarly, we can obtain lower bounds of the error computed by \( x_j, x_{j+1}, \) and \( x_{j+2}. \)
Consider a bounded linear operator $\mathcal{L} : \mathbf{X} \rightarrow \mathbf{X}$, where $\mathbf{X}$ is a Banach space. Given $b \in \mathbf{X}$, the iteration process is defined by the relation

$$x_j = \mathcal{L} x_{j-1} + b.$$  

(10.13)

Let $x_\circ$ be a fixed point of (10.13) and

$$\|\mathcal{L}\| = q < 1.$$
By applying the Banach Theorem it is easy to show that

$$\{x_j\} \rightarrow x_{\odot}.$$ 

Indeed, let $$\bar{x}_j = x_j - x_{\odot}$$. Then

$$\bar{x}_j = \mathcal{L}x_{j-1} + b - x_{\odot} = \mathcal{L}(x_{j-1} - x_{\odot}) = \mathcal{L}\bar{x}_{j-1} \quad (10.14)$$

Since

$$0_X = \mathcal{L}0_X,$$

we note that the zero element $$0_X$$ is a unique fixed point of the operator $$\mathcal{L}.$$
Therefore, we have an *a priori* estimate

\[ \| x_j - x_\circ \|_X = \| \bar{x}_j - 0x \|_X \leq \]

\[ \leq \frac{q^j}{1 - q} \| \bar{x}_1 - \bar{x}_0 \|_X = \frac{q^j}{1 - q} \| R(x_0) \|_X \quad (10.15) \]

and the *a posteriori* one

\[ \| x_j - x_\circ \|_X \leq \frac{q}{1 - q} \| R(x_{j-1}) \|_X , \quad (10.16) \]

where \( R(z) = Lz + b - z \) is the *residual* of the functional equation considered.
By applying the general theory, we also obtain a lower bound of the error

\[ \|x_j - x_\circ\|_X \geq \frac{1}{1+q} \|x_{j+1} - x_j\|_X = \frac{1}{1+q} \|R(x_j)\|_X. \quad (10.17) \]

Hence, we arrive at the following estimates for the error in the linear operator equation:

\[ \frac{1 - q}{q} \|x_j - x_\circ\|_X \leq \|R(x_{j-1})\|_X \leq (1 + q) \|x_{j-1} - x_\circ\|_X. \]

Advanced estimates that provide sharper bounds can be easily obtained by applying (10.10) and (10.11).
Important applications of the above results are associated with systems of linear simultaneous equations and other algebraic problems. Set $X = \mathbb{R}^n$ and assume that $\mathcal{L}$ is defined by a nondegenerate matrix $A \in \mathbb{M}^{n \times n}$ decomposed into three matrixes

$$A = A_\ell + A_d + A_r,$$

where $A_\ell$, $A_r$, and $A_d$ are certain lower, upper, and diagonal matrices, respectively.
Iteration methods for systems of linear simultaneous equations associated with $A$ are often represented in the form

$$B \frac{x_i - x_{i-1}}{\tau} + A x_{i-1} = f.$$  \hspace{1cm} (10.18)

In (10.18), the matrix $B$ and the parameter $\tau$ may be taken in various ways (depending on the properties of $A$). We consider three frequently encountered cases:

(a) $B = A_d$,

(b) $B = A_d + A_\ell$,

(c) $B = A_d + \omega A_\ell$, $\tau = \omega$.

For $\tau = 1$, (a) and (b) lead to the methods of Jacobi and Zeidel, respectively. In (c), the parameter $\omega$ must be in the interval $(0, 2)$. If $\omega > 1$, we have the so-called "upper relaxation method", and $\omega < 1$ corresponds to the "lower relaxation method",
The method (10.18) is reduced to (10.13) if we set

$$\mathcal{L} = \mathbb{I} - \tau \mathbf{B}^{-1} \mathbf{A} \quad \text{and} \quad \mathbf{b} = \tau \mathbf{B}^{-1} \mathbf{f}, \quad (10.19)$$

where $\mathbb{I}$ is the unit matrix. It is known that $\mathbf{x}_i$ converges to $\mathbf{x}_\circ$ that is a solution of the system

$$\mathbf{A} \mathbf{x}_\circ = \mathbf{f} \quad (10.20)$$

if an only if all the eigenvalues of $\mathcal{L}$ are less than one. Obviously, $\mathbf{B}$ and $\tau$ should be taken in such a way that they guarantee the fulfillment of this condition.
Assume that $\|\mathcal{L}\| \leq q < 1$. In view of (10.15)-(10.17), the quantities

$$
M_i^{\oplus} = q(1 - q)^{-1} \| R(x_{i-1}) \| ,
$$

$$
M_0^{\oplus} = q^i(1 - q)^{-1} \| R(x_0) \| ,
$$

$$
M_i^{\ominus} = (1 + q)^{-1} \| R(x_i) \|.
$$

furnish upper and lower bounds of the error for the vector $x_i$. The validity of them is demonstrated with an example below. It is worth noting that from the practical viewpoint finding an upper bound for $\|\mathcal{L}\|$ and proving that it is less than 1 presents a special and often not easy task.
Remark. If \( q \) is very close to 1, then the convergence of an iteration process may be very slow. As we have seen, in this case, the quality of error estimates is also degraded. A well–accepted way for accelerating the convergence consists of using a modified system obtained from the original one by means of a suitable preconditioner \( P^{-1} \) and solving the system

\[
(P^{-1}A) \, x = P^{-1}f
\]

with a smaller condition number. Of cause, the best preconditioner is the unknown matrix \( A^{-1} \). Therefore, a preconditioner is often constructed from the parts of \( A \) that are not difficult to invert (e.g., in the simplest case it is taken as the matrix inverse to the diagonal part of \( A \)). This iteration technique is well presented in the literature: see, e.g., O. Axelsson. *Iterative solution methods*. Cambridge University Press, Cambridge, 1994.
Examples

Consider the problem $Ax = f$ for a symmetric matrix $A$ with coefficients $a_{ij} = 0.8/ij$ if $i \neq j$ and $a_{ii} = i$. The system is solved by the method

$$x_{i+1} = (I - \tau B^{-1}A)x_i + \tau B^{-1}F$$

with $B = A_D$ and $x_0 = \{0, 0, \ldots, 0\}$.

In this example $n = 200$, $q = 0.662$, and $\tau = 0.760$. The values of the error and the estimates are presented below.
### Table:

<table>
<thead>
<tr>
<th>i</th>
<th>$M_i^\ominus$</th>
<th>$|e|$</th>
<th>$M_i^\oplus$</th>
<th>$M_0^\ominus$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.187145E+03</td>
<td>4.12471E+03</td>
<td>2.45893E+04</td>
<td>2.45893E+04</td>
</tr>
<tr>
<td>2</td>
<td>.452820E+02</td>
<td>1.04019E+03</td>
<td>6.10732E+03</td>
<td>1.62904E+04</td>
</tr>
<tr>
<td>3</td>
<td>.123433E+02</td>
<td>3.11517E+02</td>
<td>1.47774E+03</td>
<td>1.07924E+04</td>
</tr>
<tr>
<td>4</td>
<td>.405504E+01</td>
<td>1.16679E+02</td>
<td>4.02813E+02</td>
<td>7.14995E+03</td>
</tr>
<tr>
<td>5</td>
<td>.166633E+01</td>
<td>5.17711E+01</td>
<td>1.32333E+02</td>
<td>4.73684E+03</td>
</tr>
<tr>
<td>6</td>
<td>.767379E+00</td>
<td>2.44532E+01</td>
<td>5.43792E+01</td>
<td>3.13815E+03</td>
</tr>
<tr>
<td>7</td>
<td>.366283E+00</td>
<td>1.17450E+01</td>
<td>2.50428E+01</td>
<td>2.07902E+03</td>
</tr>
<tr>
<td>8</td>
<td>.176340E+00</td>
<td>5.66166E+00</td>
<td>1.19533E+01</td>
<td>1.37735E+03</td>
</tr>
<tr>
<td>16</td>
<td>.515722E-03</td>
<td>.165576E-02</td>
<td>.349042E-02</td>
<td>.511127E+01</td>
</tr>
<tr>
<td>17</td>
<td>.248671E-03</td>
<td>.798371E-03</td>
<td>.168302E-02</td>
<td>.338621E+01</td>
</tr>
<tr>
<td>18</td>
<td>.119903E-03</td>
<td>.384956E-03</td>
<td>.811515E-03</td>
<td>.224336E+01</td>
</tr>
<tr>
<td>19</td>
<td>.578146E-04</td>
<td>.185617E-03</td>
<td>.391295E-03</td>
<td>.148623E+01</td>
</tr>
<tr>
<td>20</td>
<td>.278769E-04</td>
<td>.895001E-04</td>
<td>.188673E-03</td>
<td>.984624E+00</td>
</tr>
</tbody>
</table>
Figure: A priori and a posteriori estimates for an iteration process:
1 – $M_i^\ominus$, 2 – $\|e\|$, 3 – $M_i^\oplus$, 4 – $M_0^\oplus$. 
Positivity methods and a posteriori error bounds.

In some cases, one can obtain two-sided estimates for each component of a solution. The respective methods can be viewed as a simplest example of the so-called positivity methods widely used in the analysis of differential equations.

Let $x_\circ$ be a solution of the system of linear simultaneous equations

$$x_\circ = Ax_\circ + f,$$

where

$$A = A^\oplus - A^\ominus$$

and

$$A^\ominus = \{a^\ominus_{ij}\} \in \mathbb{M}^{n \times n}, \quad a^\ominus_{ij} \geq 0,$$

$$A^\oplus = \{a^\oplus_{ij}\} \in \mathbb{M}^{n \times n}, \quad a^\oplus_{ij} \geq 0.$$
We may *partially order* the space $\mathbb{R}^n$ by saying that $\mathbf{x} \leq \mathbf{y}$ if and only if $x_i \leq y_i$ for $i = 1, 2, \ldots, n$.

Assume that the vectors $\mathbf{x}_0^\ominus$ and $\mathbf{x}_0^\oplus$ are ordered such that

$$\mathbf{x}_0^\ominus \leq \mathbf{x} \leq \mathbf{x}_0^\oplus.$$ 

The vectors $\mathbf{x}_0^\ominus$ and $\mathbf{x}_0^\oplus$ are considered as the *initial guesses* for the bounds of the solution components.

Compute $\mathbf{x}_1^\ominus$ and $\mathbf{x}_1^\oplus$ by the relations

$$\mathbf{x}_1^\ominus = A^\ominus \mathbf{x}_0^\ominus - A^\ominus \mathbf{x}_0^\oplus + \mathbf{f},$$

$$\mathbf{x}_1^\oplus = A^\oplus \mathbf{x}_0^\oplus - A^\oplus \mathbf{x}_0^\ominus + \mathbf{f}.$$
It is easy to see that

\[ x_1^\ominus - x_\oslash = A^\oplus (x_0^\ominus - x_\oslash) - A^\ominus (x_0^\oplus - x_\oslash) \leq 0, \]
\[ x_1^\oplus - x_\oslash = A^\oplus (x_0^\oplus - x_\oslash) - A^\ominus (x_0^\ominus - x_\oslash) \geq 0. \]

Hence,

\[ x_1^\ominus \leq x_\oslash \leq x_1^\oplus. \]

and we observe that \( x_1^\ominus \) and \( x_1^\oplus \) also give componentwise bounds for the exact solution.
Quite similarly, we observe that the subsequent elements of the iteration process

$$x_{k+1}^\ominus = A^\oplus x_k^\ominus - A^\ominus x_k^\oplus + f$$

$$x_{k+1}^\oplus = A^\ominus x_k^\oplus - A^\ominus x_k^\oplus + f,$$

possess the same properties. Therefore, for the $i$th component we find the following two-sided bounds:

$$\max_{j=0,1,...,k+1} (x_j^\ominus)_i \leq (x_i^\varotimes)_i \leq \min_{j=0,1,...,k+1} (x_j^\oplus)_i.$$
Similar methods can be applied to *functional equations*, provided that the operator $A$ is presented as the sum of

$$A^\oplus \quad \text{and} \quad (-A^\ominus)$$

which are certain monotone operators defined on a partially ordered space:

see, e.g.,

Applications to integral equations

Many problems in science and engineering can be stated in terms of integral equations. One of the most typical cases is to find a function \( x \odot (t) \in C[a, b] \) such that

\[
x \odot (t) = \lambda \int_a^b K(t, s) x \odot (s) \, ds + f(t),
\]

(10.24)

where \( \lambda \geq 0 \), \( K \) (the kernel) is a continuous function for

\[
(x, t) \in Q := \{a \leq s \leq b, \ a \leq t \leq b\}
\]

and

\[
|K(t, s)| \leq M, \quad \forall (t, s) \in Q.
\]

Also, we assume that \( f \in C[a, b] \).
Let us define the operator $\mathcal{I}$ as follows:

$$
y(t) := \mathcal{I}x(t) := \lambda \int_{a}^{b} K(t, x)x(s)\,ds + f(t) \quad (10.25)
$$

and show that $\mathcal{I}$ maps continuous functions to continuous ones. Let $t_{0}$ and $t_{0} + \Delta t$ belong to $[a, b]$. Then,

$$
|y(t_{0} + \Delta t) - y(t_{0})| \leq \lambda \int_{a}^{b} |K(t_{0} + \Delta t, s) - K(t_{0}, s)||x(s)|\,ds + |f(t_{0} + \Delta t) - f(t_{0})|.
$$

Since $K$ and $f$ are continuous on the compact sets $Q$ and $[a, b]$, respectively, they are uniformly continuous on these sets.
Therefore, for any given $\varepsilon$ one can find a small number $\delta$ such that

$$|f(t_0 + \Delta t) - f(t_0)| < \varepsilon$$

and

$$|K(t_0 + \Delta t, s) - K(t_0, s)| < \varepsilon,$$

provided that $|\Delta t| < \delta$.

Thus, we have

$$|y(t_0 + \Delta t) - y(t_0)| \leq \lambda \varepsilon (|x| |b - a| \max_{s \in [a,b]} |x(s)| + 1) = C\varepsilon,$$

and, consequently, $y(t_0 + \Delta t)$ tends to $y(t_0)$ as $|\Delta t| \to 0$. 

S. Repin

RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
\( \mathcal{T} : C[a, b] \to C[a, b] \) is a **contractive mapping**. Indeed,

\[
\begin{align*}
d(\mathcal{T}x, \mathcal{T}y) &= \max_{a \leq t \leq b} |\mathcal{T}x(t) - \mathcal{T}y(t)| = \\
&= \max_{a \leq t \leq b} \left| \lambda \int_{a}^{b} K(t, s)(x(s) - y(s)) \, ds \right| \leq \\
&\leq |\lambda| M(b - a) \max_{a \leq s \leq b} |x(s) - y(s)| = |\lambda| M(b - a) d(x, y),
\end{align*}
\]

so that \( \mathcal{T} \) is a \( q \)-contractive operator with

\[
q = |\lambda| M(b - a), \quad (10.26)
\]

provided that

\[
|\lambda| < \frac{1}{M(b - a)}. \quad (10.27)
\]
Numerical procedure

An approximate solution of (10.24) can be found by the iteration method

\[ x_{i+1}(t) = \lambda \int_a^b K(t, s)x_i(s) \, ds + f(t). \]  

(10.28)

If (10.27) holds, then from the Banach theorem it follows that the sequence \( \{x_i\} \) converges to the exact solution. We apply the theory exposed above and find that the accuracy of \( x_i \) is subject to the estimate

\[
\frac{1}{1 + q} \int_a^b K(t, s)(x_{i+1}(s) - x_i(s)) \, ds \leq \max_{a \leq t \leq b} |x_i(t) - x_\circ(t)| \leq \frac{q}{1 - q} \int_a^b K(t, s)(x_i(s) - x_{i-1}(s)) \, ds.
\]  

(10.29)
Applications to Volterra type equations

Consider the fixed point problem

\[ x \odot (t) = \lambda \int_a^t K(t, s) x \odot (s) \, ds + f(t), \quad (10.30) \]

where

\[ |K(t, s)| \leq M, \quad \forall (t, s) \in Q \]

and \( f \in C[a, b] \).

Define the operator \( T \) as follows:

\[ T x(t) = \lambda \int_a^t K(t, s) x(s) \, ds + f(t). \]

Similarly, to the previous case we establish that

\[ d(Tx, Ty) \leq |\lambda| M(t - a) d(x, y). \]
By the same arguments we find that

$$d(T^n x, T^n y) \leq |\lambda|^n M^n (t - a)^n \frac{n!}{n!} d(x, y),$$

Thus, the operator $\mathcal{T} := T^n$ is $q$-contractive with a certain $q < 1$, provided that $n$ is large enough.

In view of Proposition 1, we conclude that the iteration method converges to $x_\circ$ and the errors are controlled by the two–sided error estimates.
Applications to ordinary differential equations

Let \( u \) be a solution of the simplest initial boundary-value problem

\[
\frac{du}{dt} = \varphi(t, u(t)), \quad u(t_0) = a,
\]

where the solution \( u(t) \) is to be found on the interval \([t_0, t_1]\). Assume that the function \( \varphi(t, p) \) is continuous on the set

\[
Q = \{ t_0 \leq t \leq t_1, \ a - \Delta \leq p \leq a + \Delta \}
\]

and

\[
|\varphi(t, p_1) - \varphi(t, p_2)| \leq L|p_1 - p_2|, \quad \forall (t, p) \in Q.
\]

(10.31)
Problem (10.31) can be reduced to the integral equation

\[ u(t) = \int_{t_0}^{t} \varphi(s, u(s)) \, ds + a \]  

(10.33)

and it is natural to solve the latter problem by the iteration method

\[ u_j(t) = \int_{t_0}^{t} \varphi(s, u_{j-1}(s)) \, ds + a. \]  

(10.34)

To justify this procedure, we must verify that the operator

\[ T u := \int_{t_0}^{t} \varphi(s, u(s)) \, ds + a \]

is q-contractive with respect to the norm

\[ \|u\| := \max_{t \in [t_0, t_1]} |u(t)|. \]  

(10.35)
We have

\[
\|\mathcal{T} z - \mathcal{T} y \| = \max_{t \in [t_0, t_1]} \left| \int_{t_0}^{t} (\varphi(s, z(s)) - \varphi(s, y(s)) \, ds \right| \leq \\
\leq \max_{t \in [t_0, t_1]} L \int_{t_0}^{t} |z(s) - y(s)| \, ds \leq L \int_{t_0}^{t_1} |z(s) - y(s)| \, ds \leq \\
\leq L(t_1 - t_0) \max_{s \in [t_0, t_1]} |z(s) - y(s)| = L(t_1 - t_0) \|z - y\|.
\]

We see that if

\[
t_1 < t_0 + L^{-1},
\]

then the operator \(\mathcal{T}\) is \(q\)-contractive with

\[
q := L(t_1 - t_0) < 1.
\]
Therefore, if the interval $[t_0, t_1]$ is small enough (i.e., it satisfies the condition (10.36)), then the existence and uniqueness of a continuous solution $u(t)$ follows from the Banach theorem. In this case, the solution can be found by the iteration procedure whose accuracy is explicitly controlled by the two–sided error estimates.

For a more detailed investigation of the fixed point methods for integral and differential equations see


Lecture 11.
A POSTERIORI ESTIMATES FOR VARIATIONAL INEQUALITIES
Lecture plan

- Variational inequalities. Background;
- Deviation estimates for variational inequalities;
- Obstacle problem;
- Functional type a posteriori estimates for problems with two obstacles;
- Examples;
- Elasto-plastic torsion problem;
Variational inequalities

Variational inequalities provide a mathematical description of a wide spectrum of nonlinear boundary-value problems that arise in various applications (see, e.g., G. Duvant and J.-L. Lions. *Les inéquations en mécanique et en physique*, Dunod, Paris, 1972.)

**First we establish the relationship between variational inequalities and certain variational problems.** Consider the functional

\[ J(v) = J_0(v) + j(v), \]

where \( J_0 : V \rightarrow \mathbb{R} \) is a convex, continuous, and Gateaux-differentiable functional and \( j(v) : V \rightarrow \mathbb{R} \) is a convex and continuous functional.
Let $K$ be a convex closed subset of a reflexive Banach space $V$. Consider the following problem: find $u \in K$ such that

$$J(u) = \inf_{v \in K} J(v), \quad J(v) = J_0(v) + j(v).$$  \hfill (11.1)

Hereafter, we assume that $J$ is coercive on $V$, so that the above problem has a solution $u$. Moreover, the minimizer satisfies the relation

$$\langle J'_0(u), u - v \rangle + j(u) - j(v) \leq 0 \quad \forall v \in K$$ \hfill (11.2)

**Theorem (1)**

*Relations (11.1) and (11.2) are equivalent.*
Proof

1. Let (11.1) holds, i.e.

\[ J_0(v) + j(v) \geq J_0(u) + j(u) \quad \forall v \in K. \]

Take \( v = u + \lambda (w - u), \ w \in K, \lambda \in [0, 1]. \) Then

\[ J_0 (u + \lambda (w - u)) - J_0(u) + j(v) - j(u) \geq 0 \quad \forall u \in K. \]

By the convexity of \( j \) we have

\[ j(v) = j(u + \lambda (w - u)) = j(\lambda w + (1 - \lambda)u) \leq \lambda j(w) + (1 - \lambda)j(u). \]
Thus, for any $w \in K$ we have

\[
J_0 (u + \lambda(w - u)) - J_0(u) + \lambda j(w) \\
+ (1 - \lambda)j(u) - j(u) \geq 0 \quad \forall w \in K,
\]

\[
\frac{1}{\lambda} (J_0 (u + \lambda(w - u)) - J_0(u)) \\
+ j(w) - j(u) \geq 0 \quad \forall u \in K.
\]

Passing to the limit as $\lambda \to 0$ we obtain

\[
(J'_0(u), w - u) + j(w) - j(u) \geq 0 \quad \forall w \in K.
\]
2. Assume now that (11.2) holds For a convex functional $J_0$ we have the relation

$$J_0(v) \geq J_0(u) + (J'_0(u), v - u).$$

Since

$$(J'_0(u), u - v) + j(u) - j(v) \leq 0 \quad \forall v \in K$$

and

$$(J'_0(u), u - v) \geq J_0(u) - J_0(v)$$

we find that

$$-J_0(v) + J_0(u) + j(u) - j(v) \leq 0 \quad \forall v \in K,$$

what means that

$$J(u) \leq J(v) \quad \forall v \in K.$$
Variational inequalities can be regarded as Euler’s equations to certain convex variational problems with nondifferentiable functionals defined on convex subsets. If the nondifferentiable part of such a functional vanishes and the set coincide with the whole space, then the respective variational inequality converts to a variational equality (integral identity). However, in many practically interesting problems it is impossible to define a minimizer throughout an integral identity. This fact stimulated the development of the theory of variational inequalities and their numerical analysis (see, e.g.,

Obstacle problem. Introduction

Let $\Omega$ be a bounded domain in $\mathbb{R}^n$ ($n = 1, 2$) with $L$–continuous boundary $\partial \Omega$ and

$$f \in L_2(\Omega),$$
$$\varphi \in H^2(\Omega), \quad \varphi(x) \leq 0 \text{ on } \partial \Omega.$$

"Admissible" functions belong to the set

$$K_\varphi := \{ v \in V \mid v(x) \geq \varphi(x) \text{ a. e. in } \Omega \},$$

where

$$V := \{ v \in H^1(\Omega) \mid v = 0 \text{ on } \partial \Omega \}.$$
Let

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

$$(u, v) := \int_{\Omega} u v \, dx.$$ 

Then, the problem has a variational form

**Problem \( \mathcal{P} \).** Find \( u \in K_{\varphi} \) such that

$$J(u) = \inf_{v \in K_{\varphi}} J(v),$$

$$J(v) = \frac{1}{2} a(v, v) - (f, v).$$
Physical interpretation

This problem can be interpreted as the one for an elastic membrane deformed at the neighborhood of an obstacle $\varphi(x)$. 
Existence of a minimizer

**Theorem (Lions – Stampacchìa)**

*Under the above assumptions Problem $\mathcal{P}$ possesses a unique solution $u$.***

Problem $\mathcal{P}$ is, in fact, a **free boundary problem**:

$$
\overline{\Omega} = \Omega_\varphi \cup \Omega_0 \quad \text{coincidenceset}
$$

where

$$
\Omega_\varphi := \{ x \in \Omega \mid u(x) = \varphi(x) \}
$$

and

$$
\Omega_0 := \{ x \in \Omega \mid u(x) > \varphi(x) \}
$$

Minimizer $u$ satisfies **the variational inequality**

$$
a(u, v - u) \geq (f, v - u) \quad \forall v \in K_\varphi.
$$
If \( u \) is sufficiently regular, then directly from the variational inequality we derive the following relations that must hold for the solution:

\[
\Delta u + f = 0 \quad \text{on} \quad \Omega_0,
\]
\[
\Delta u + f \leq 0 \quad u \geq \varphi \quad \text{a. e. in} \quad \Omega,
\]
\[
(\Delta u + f)(u - \varphi) = 0 \quad \text{a. e. in} \quad \Omega,
\]
Regularity estimates for obstacle problems

In the papers by H. Brezis, D. Kinderlehrer, H. Lewy, G. Stampacchia and others, it was shown that

\[
\text{If } f \in L_2 \text{ and } \varphi \in H^2(\Omega) \text{ then } u \in H^2(\Omega). 
\]

Moreover, if \( f \in C^1(\Omega) \), \( \Omega \) is a bounded domain with smooth boundary, and \( \varphi \in C^2 \) than the respective solution possesses second derivatives bounded in \( L_\infty \).
Coincidence set

If $\Omega \subset \mathbb{R}^2$ is strictly convex with smooth boundary and if $\varphi \in C^2$ is strictly concave, then the coincidence set is connected and its boundary is smooth and homeomorphic to the unit circle.

In general, for any $\Omega$ one can point out such an obstacle $\varphi$ that $\Omega_\varphi$ has any number of disjoint subsets.
Problem $\mathcal{P}$ is related to a variational inequality.

The coincidence set $\Omega_\varphi$ is unknown a priori, so that a solution has a free boundary.

Solutions of Problem $\mathcal{P}$ have a bounded regularity even for smooth external data (in the best case scenario the second derivatives are summable, but the third ones are only distributions).
A priori convergence estimates for problems with obstacles were derived in:


It was shown that for regular FE approximations of $u \in H^2$:

$$\|\nabla (u - u_h)\|_{\Omega} \leq C(u, f, \varphi) h$$
Two methods usually applied for linear PDEs, namely *residual method* and *gradient averaging methods* are difficult to directly apply because:

- There is no differential equation whose "residual" could control the error in the sense of residual method.
- The applicability of averaging (post-processing) is based on higher regularity of exact solutions that implies the *superconvergence* phenomenon. Typically, solutions of variational inequalities have bounded regularity and, therefore, we cannot await such type effects.
Below we show that by the functional method it is possible to derive a posteriori estimates of the difference between the exact solution of an obstacle problem and any conforming approximation. This estimate does not require a priori knowledge on the configuration of the coincidence set.
Basic deviation estimate for variational inequalities

Let $a : V \times V \rightarrow \mathbb{R}$ be a bilinear $V$-elliptic form and $j : V \rightarrow \mathbb{R}$ be a given convex continuous functional.

Consider the following problem: find $u \in K$ such that the inequality

$$ a(u, w - u) + j(w) - j(u) \geq \langle f, w - u \rangle \quad (11.3) $$

holds for any $w \in K$, where $K$ is a convex closed subset of $V$ and $f \in V^*$. 
The solution $u$ of (11.3) is a minimizer of the following variational problem $\mathcal{P}$: Find $u \in K$ such that

$$J(u) = \inf_{w \in K} J(w),$$

(11.4)

$$J(w) = \frac{1}{2}a(w, w) + j(w) - \langle f, w \rangle.$$

Our aim is to derive a computable upper bound for the quantity $\frac{1}{2}a(u - v, u - v)$ where $v$ is any element of the set $K$. 

S. Repin  

RICAM, Special Radon Semester, Linz, 2005. 

LECTURES ON A POSTERIORI ERROR CONTROL
Further analysis follows the lines of the paper


First, we use (11.3) to obtain the inequality

\[ J(v) - J(u) = \frac{1}{2} a(v - u, v - u) + a(u, v - u) - \langle f, v - u \rangle + j(v) - j(u) \geq \frac{1}{2} a(v - u, v - u), \]

which implies the **basic deviation estimate**.

\[
\frac{1}{2} \| v - u \|^2 \leq J(v) - J(u), \quad \forall v \in K, \tag{11.5}
\]

where \( \| w \|^2 := a(v, v) \).
For linear problems we have derived deviation estimates by means of the inequality

\[
\frac{1}{2} \| v - u \|^2 \leq J(v) - J(u) = J(v) - l^*(p^*).
\]

In Lectures 4 and 5 we have shown how to find a directly computable and physically meaningful upper bound of

\[
J(v) - l^*(q^*).
\]

For variational inequalities, deviation estimates are obtained in a similar way, but with some complications caused by the fact that the problem dual to \( P \) has a more cumbersome form. Below, we show how we can circumvent this difficulty by using the so-called perturbed functionals.
Consider a bilinear form $a : V_0 \times V_0 \to \mathbb{R}$ defined by the relation

$$a(v, w) := \int_{\Omega} A \nabla v \cdot \nabla w \, dx,$$

where $\Omega$ is a bounded domain in $\mathbb{R}^2$ with Lipschitz continuous boundary $\partial \Omega$, $V_0 := \overset{\circ}{H}^1(\Omega)$, and $A = \{a_{ij}\}$ is a symmetric matrix satisfying the conditions

$$\nu_1 |\xi|^2 \leq A \xi \cdot \xi \leq \nu_2 |\xi|^2, \quad \forall \xi \in \mathbb{R}^n, \quad \nu_2 \geq \nu_1 > 0.$$
Let \( K = K_{fp} := \{ v \in V_0 \mid \varphi(x) \leq v(x) \leq \psi(x) \ \text{a.e. in } \Omega \} \), where \( \varphi, \psi \in H^2(\Omega) \) are two given functions such that

\[
\varphi(x) \leq \psi(x), \quad \forall x \in \Omega,
\]

Set in the general setting

\[
j \equiv 0 \quad \text{and} \quad \langle f, v \rangle = \int_{\Omega} fv \, dx.
\]

Then Problem \( \mathcal{P} \) is the classical obstacle problem. A solution \( u \) minimizes the functional

\[
J(v) = \int_{\Omega} A \nabla v \cdot \nabla v \, dx - \int_{\Omega} fv \, dx \quad \text{on } K_{fp}.
\]
In general, $\Omega$ is divided into three sets:

$$\Omega_u^\oplus := \{ x \in \Omega \mid u(x) = \psi(x) \}$$  (upper coincidence set),

$$\Omega_u^\ominus := \{ x \in \Omega \mid u(x) = \varphi(x) \}$$  (lower coincidence set),

$$\Omega_0^u := \{ x \in \Omega \mid \varphi(x) < u(x) < \psi(x) \}.$$

Here, $\Omega_0^u$ is an open set, where a solution satisfies the differential equation. Thus, we see that this problem involves free boundaries, which are *unknown a priori*.

Differentiability properties of solutions to linear and quasilinear problems with obstacles were investigated by many authors. In particular, it was proved that, under natural assumptions on external data $u \in H^2(\Omega)$ and even for very smooth data, solutions have a limited regularity (which is $W^{2,\infty}$). We assume that these assumptions are fulfilled and the solutions are $H^2$-regular.
Perturbed problem

To estimate the difference $J(v) - J(u)$ we introduce the *perturbed functional*

$$J_\lambda(v) := J(v) - \int_\Omega \lambda \cdot (v - \Phi) \, dx,$$

where $\Phi = (\varphi, -\psi)$ and $v = (v, -v)$,

$$\lambda \in \mathbb{R}_+^2 := \left\{ (\lambda_1, \lambda_2) \mid \lambda_i \in L^2(\Omega), \lambda_i(x) \geq 0 \text{ a.e. in } \Omega, \ i = 1, 2 \right\}.$$

It is easy to see that

$$\sup_{\lambda \in \mathbb{R}_+^2} J_\lambda(v) = J(v) - \inf_{\lambda \in \mathbb{R}_+^2} \int_\Omega \lambda \cdot (v - \Phi) \, dx$$

$$= \begin{cases} J(v) & \text{if } v \in K_{fp}, \\ +\infty & \text{if } v \notin K_{fp} \end{cases} \quad (11.7)$$
With $J_\lambda$ we associate a *perturbed variational problem*. Problem $P_\lambda$. Find $u_\lambda \in V_0$ such that

$$J_\lambda(u_\lambda) = \inf_{v \in V_0} J_\lambda(v) : = \inf P_\lambda.$$  \hspace{1cm} (11.8)

Since

$$\inf_{v \in V_0} J_\lambda(v) \leq \inf_{v \in K_{fp}} J_\lambda(v) = \inf_{v \in K_{fp}} J(v) = \inf P,$$

we see that

$$\inf P_\lambda \leq \inf P, \hspace{0.5cm} \forall \lambda \in \mathbb{N}_\oplus.$$  \hspace{1cm} (11.9)
It follows from (11.5) and (11.9) that

\[ \frac{1}{2} \| v - u \|^2 \leq J(v) - \inf \mathcal{P}_\lambda, \quad \lambda \in \mathbb{N}_\oplus. \] (11.10)

To estimate the right-hand side of (11.9), we introduce a dual counterpart of Problem $\mathcal{P}_\lambda$. 
Dual perturbed problem

By the Lagrangian

\[ L(v, \tau, \lambda) := \int_{\Omega} \left( \tau \cdot \nabla v - \frac{1}{2} A^{-1} \tau \cdot \tau - fv - \lambda \cdot (v - \Phi) \right) \, dx, \]

we define the perturbed functional as follows:

\[ J_\lambda(v) = \sup_{\tau \in Y^*} L(v, \tau, \lambda), \quad Y^* := L^2(\Omega, \mathbb{R}^2). \]

**Problem \( P_\lambda^* \).** Find \( \tau_\lambda \) such that

\[ J_\lambda^*(\tau_\lambda) = \sup_{q \in Q_{f\lambda}^*} J_\lambda^*(q), \]

where \( J_\lambda^*(q) = \int_{\Omega} \left( -\frac{1}{2} A^{-1} q \cdot q + \lambda \cdot \Phi \right) \, dx \) and

\[ Q_{f\lambda}^* := \left\{ q \in Y^* \middle| \int_{\Omega} q \cdot \nabla v \, dx = \int_{\Omega} (fv + \lambda \cdot v) \, dx \; \forall v \in V_0 \right\}. \]
$Q_{f\lambda}^*$ is a closed affine manifold in $Y^*$ and the functional $-J^*_\lambda$ is convex and continuous on $Y^*$.
Therefore, Problem $P^*_\lambda$ has a solution and

$$\inf P_\lambda = \sup P^*_\lambda.$$  \hfill (11.11)
Estimates of the deviation

By means of (11.5) and (11.11) we obtain

\[ \frac{1}{2} \| v - u \|^2 \leq J(v) - \sup \mathcal{P}_\lambda^* \leq J(v) - J_\lambda^*(q) \tag{11.12} \]

Here

\[ v \in K_{fp}, \quad q \in Q_{f,\lambda}^*, \quad \lambda \in \mathbb{N}_+. \]

Rewrite \( J(v) - J_\lambda^*(q) \) in a more transparent form. We have

\[ \frac{1}{2} \| v - u \|^2 \leq \int_\Omega \left( \frac{1}{2} A \nabla v \cdot \nabla v - fv \right) \, dx + \int_\Omega \frac{1}{2} A^{-1} \tau \cdot \tau \, dx \]

\[ + \frac{1}{2} \int_\Omega \left( A^{-1} q \cdot q - A^{-1} \tau \cdot \tau \right) \, dx - \int_\Omega \lambda \cdot \Phi \, dx. \]
In view of the relation

\[ A^{-1}a \cdot a - A^{-1}b \cdot b = A^{-1}(a - b) \cdot (a - b) + 2A^{-1}b \cdot (a - b) \]

and the integral identity

\[ \int_{\Omega} fv \, dx = \int_{\Omega} q \cdot \nabla v \, dx - \int_{\Omega} \lambda \cdot v \, dx, \quad \forall q \in Q^{*}_{f\lambda}, \]

we obtain

\[
\frac{1}{2} \| v - u \|^2 \leq \frac{1}{2} \int_{\Omega} (A \nabla v - \tau) \cdot (\nabla v - A^{-1} \tau) \, dx \\
+ \int_{\Omega} \lambda \cdot (v - \Phi) \, dx + \frac{1}{2} \int_{\Omega} A^{-1} (q - \tau) \cdot (q - \tau) \\
+ \int_{\Omega} (\nabla v - A^{-1} \tau) \cdot (\tau - q) \, dx.
\]
The last integral can be estimated as follows:

\[
\int_{\Omega} (\nabla v - A^{-1}\tau) \cdot (\tau - q) \, dx \\
\leq \frac{\beta}{2} \int_{\Omega} A(\nabla v - A^{-1}\tau) \cdot (\nabla v - A^{-1}\tau) \, dx \\
+ \frac{1}{2\beta} \int_{\Omega} A^{-1}(q - \tau) \cdot (q - \tau) \, dx,
\]

where \(\beta\) is any positive number.

Introduce the quantity

\[
d^2(\tau, Q^*_{f\lambda}) := \inf_{q \in Q^*_{f\lambda}} \int_{\Omega} A^{-1}(q - \tau) \cdot (q - \tau) \, dx,
\]

which is the distance between \(\tau\) and the set \(Q^*_{f\lambda}\).
Now, we rewrite the estimate as follows:

\[
\| \mathbf{v} - \mathbf{u} \|^2 \leq \]

\[
\leq (1 + \beta)d^2(\tau, Q^*_f) + \left(1 + \frac{1}{\beta}\right) \| \nabla \mathbf{v} - A^{-1}\tau \|^2 + 
\]

\[
+ 2 \int_{\Omega} \lambda \cdot (\mathbf{v} - \Phi) \, dx. \quad (11.13)
\]
Recall some relations that has been established in Lecture 5. We have proved that

\[ d(y, Q^*_\ell) = \| \ell + \Lambda^*y \| := \sup_{w \in V_0} \frac{\langle \ell + \Lambda^*y, w \rangle}{\| \Lambda w \|}, \]

where \( Q^*_\ell := \{ y \in Y^* \mid (y, \Lambda w) + \langle \ell, w \rangle = 0, \quad \forall w \in V_0 \} \).

In our case, \( y = \tau, \Lambda = \nabla, \Lambda^* = -\text{div}, \) and \( Q^*_\ell = Q^*_f \lambda \) if set

\[ \langle \ell, w \rangle = -\int_\Omega (fw + \lambda \cdot \nu)dx. \]

Therefore,

\[ d(y, Q^*_\ell) = \sup_{w \in V_0} \frac{\int_\Omega (-f - \lambda_1 + \lambda_2 - \text{div} y)wdx}{\| w \|}. \]
Assume that $\tau \in Q^* := H(\Omega, \text{div})$. Then,

$$\int_{\Omega} (-f - \lambda_1 + \lambda_2 - \text{div} y) w \, dx \leq C_{\Omega,A} \| f + \lambda_1 - \lambda_2 + \text{div} y \| \| w \|$$

and we obtain

$$d(\tau, Q^*) \leq C_{\Omega,A} \| \text{div} \tau + f + \lambda_1 - \lambda_2 \|, \quad (11.14)$$

where $C_{\Omega,A}$ is a constant in the inequality

$$\| w \|_{2,\Omega} \leq C_{\Omega,A} \| w \|, \quad \forall w \in V_0.$$
Thus, for $y^* \in Q^*$ we obtain the estimate

$$\|v - u\|^2 \leq C_{\Omega,A}^2 (1 + \beta) \|\text{div}\, \tau + f + \lambda_1 - \lambda_2\|^2$$

$$+ \left(1 + \frac{1}{\beta}\right) \|\nabla v - A^{-1}\tau\|^2 + 2 \int_{\Omega} \lambda \cdot (v - \Phi) \, dx. \quad (11.15)$$

In this estimate, $\lambda$ is a ”free” vector–valued function. We use this freedom to obtain the most accurate upper bound for the deviation $\|v - u\|$.

Below we consider two options that lead to two different a posteriori error estimates for the obstacle problem.
The first option is as follows. Let \( v \in V \) be an approximate solution. For almost all points of \( \Omega \) the function \( v(x) \) is either equal to \( \varphi \), or \( \psi \) or lies between these two values. Thus, almost all points of \( \Omega \) can be referred to one of the three sets:

\[
\begin{align*}
\Omega^v_0 & := \{ x \in \Omega \mid \varphi(x) < v(x) < \psi(x) \}, \\
\Omega^v_\ominus & := \{ x \in \Omega \mid v(x) = \varphi(x) \}, \\
\Omega^v_\oplus & := \{ x \in \Omega \mid v(x) = \psi(x) \}.
\end{align*}
\]

Now, we can choose \( \lambda \) as follows:

\[
\begin{align*}
\lambda_1 &= \lambda_2 = 0 \quad \text{a.e. in } \Omega^v_0, \\
\lambda_1 &= -\langle \text{div} \, \tau + f \rangle_\ominus, \quad \lambda_2 = 0 \quad \text{a.e. in } \Omega^v_\ominus, \\
\lambda_1 &= 0, \quad \lambda_2 = \langle \text{div} \, \tau + f \rangle_\oplus, \quad \text{a.e. in } \Omega^v_\oplus.
\end{align*}
\]

Here \( \langle z \rangle_\oplus \) is zero if \( z \leq 0 \) and \( z \) if \( z > 0 \).
As a result of such a choice of $\lambda$, we obtain the estimate

$$
\| v - u \|^2 
\leq 
M_1(v, \tau, \beta) := \left( 1 + \frac{1}{\beta} \right) \| \nabla v - A^{-1} \tau \|^2 
+ C_{\Omega,A}^2 (1 + \beta) \left[ \int_{\Omega_0^v} |r(\tau)|^2 \, dx + \int_{\Omega_\oplus} \langle r(\tau) \rangle_\ominus^2 \, dx + \int_{\Omega_\ominus} \langle r(\tau) \rangle_\oplus^2 \, dx \right],
$$

(11.16)

where

$$
r(\tau) = \text{div} \, \tau + f
$$

and $\langle \rangle_{\ominus}$ and $\langle \rangle_{\oplus}$ denote the negative and positive parts of a quantity, respectively.
What is the meaning of the four terms of the Majorant?

The first term $\| \nabla v - A^{-1} \tau \|^2$ penalizes the error in the duality relation

$$\nabla v = A^{-1} \tau.$$ 

Other terms penalize "improper" behavior of $r(\tau)$ on the sets $\Omega^v_0$, $\Omega^v_\oplus$, and $\Omega^v_\ominus$, respectively. Indeed, on $\Omega^v_0$ the differential equation must be satisfied. Therefore, the term

$$\int_{\Omega^v_0} |r(\tau)|^2 \, dx$$

can be viewed as a penalty, which is nonzero if the variable $\tau$ (flux image) does not satisfy the differential equation.
By the necessary conditions for the obstacle problem, we find that

\[
\text{div } A \nabla u(x) + f(x) \leq 0, \quad \text{for a.e. } x \in \Omega^\nabla, \\
\text{div } A \nabla u + f(x) \geq 0 \quad \text{for a.e. } x \in \Omega^\nabla.
\]

Thus, the terms

\[
\int_{\Omega^\nabla} \langle r(\tau) \rangle^2_\ominus dx \quad \text{and} \quad \int_{\Omega^\nabla} \langle r(\tau) \rangle^2_\oplus dx
\]

are certain penalties for the violation of above conditions.
We see that the majorant $M_1$ is a nonnegative functional, which vanishes if and only if

$$\nabla v(x) = A^{-1}\tau(x) \quad \text{for a. e. } x \in \Omega,$$  \hspace{1cm} (11.17)

$$\text{div} \tau(x) + f(x) \leq 0, \quad \text{for a. e. } x \in \Omega^\vee_\ominus,$$  \hspace{1cm} (11.18)

$$\text{div} \tau(x) + f(x) = 0, \quad \text{for a. e. } x \in \Omega^\vee_0,$$  \hspace{1cm} (11.19)

$$\text{div} \tau(x) + f(x) \geq 0 \quad \text{for a. e. } x \in \Omega^\vee_\oplus.$$  \hspace{1cm} (11.20)

Let us show that in this case $v = u$ and $\tau = A\nabla u$. 
Assume that \((11.17) - (11.20)\) hold. Then for any \(w \in K_{fp}\), we have

\[
\int_{\Omega} A \nabla v \cdot \nabla (w - v) \, dx - \int_{\Omega} f(w - v) \, dx
\]

\[
= \int_{\Omega} (\text{div} \, \tau + f)(v - w) \, dx = \int_{\Omega_{\cap}} (\text{div} \, \tau + f)(\varphi - w) \, dx
\]

\[
+ \int_{\Omega_{\cup}} (\text{div} \, \tau + f)(v - w) \, dx + \int_{\Omega_{\cup}} (\text{div} \, \tau + f)(\psi - w) \, dx \geq 0.
\]

This inequality means that

\[
a(v, w - v) \geq \int_{\Omega} f(w - v) \, dx, \quad \forall w \in K_{fp},
\]

so that \(v\) coincides with the exact solution \(u\) (which is unique!).

S. Repin

\textit{RICAM, Special Radon Semester, Linz, 2005.}

\textbf{LECTURES ON A POSTERIORI ERROR CONTROL}
All said above can be summarized as follows:

**Theorem**

For any $\beta > 0$, $M_1(v, \tau, \beta)$ is a nonnegative functional that majorizes $\| v - u \|^2$ and vanishes if and only if

$$v = u \quad \text{and} \quad \tau = A \nabla u,$$

where $u$ is a solution of the variational inequality

$$a(u, w - u) \geq \int_{\Omega} f(w - u) \, dx, \quad \forall w \in K_{fp},$$
To obtain a more rigorous upper bound of $\| v - u \|$, we should find $\lambda$ by minimizing the right-hand side of the estimate

$$
\| v - u \|^2 \leq C^2_{\Omega,A}(1 + \beta) \| \text{div} \tau + f + \lambda_1 - \lambda_2 \|^2 \\
+ \left( 1 + \frac{1}{\beta} \right) \| \nabla v - A^{-1} \tau \|^2 + 2 \int_{\Omega} \lambda \cdot (v - \Phi) \, dx.
$$

Note that it leads to a quadratic type minimization problem in $L^2$ that can be solved analytically. On this way, we arrive at the estimate

$$
\| v - u \|^2 \leq M_2(v, \tau, \beta) := \left( 1 + \frac{1}{\beta} \right) \| \nabla v - A^{-1} \tau \|^2 \\
+ \int_{\Omega} R(v, \text{div} \tau + f, \beta) \, dx,
$$

where

\[ (11.21) \]
\[ R(v, r, \beta) = \begin{cases} 
- \frac{(\varphi - v)^2}{c_\beta} + 2r(\varphi - v) & \text{if } c_\beta r + v \leq \varphi, \\
 c_\beta r^2 & \text{if } \varphi < c_\beta r + v < \psi, \\
- \frac{(\psi - v)^2}{c_\beta} + 2r(\psi - v) & \text{if } c_\beta r + v \geq \psi
\end{cases} \]

and \( c_\beta = C^2_{\Omega_A}(1 + \beta) \).
Let us show that the term $R(v, r, \beta)$ is equal to zero in the following three cases:

(I) $r = 0$,  (II) $v = \varphi$ and $r < 0$,  (III) $v = \psi$ and $r > 0$.

Assume that $r = 0$. If $\varphi < v < \psi$, then the second branch is realized and we see that $R = 0$. If $\varphi = v$ (or $\psi = v$), then the first (third) branch is realized and also $R = 0$.

Let $r > 0$ (the case $r < 0$ is considered quite similar). Then the first branch is impossible. On the second one we have only positive values. For the third branch we have

$$r \geq \frac{\psi - v}{c_{\beta}} \quad \text{and, therefore,} \quad -\frac{(\psi - v)^2}{c_{\beta}} + 2r(\psi - v) \geq \frac{(\psi - v)^2}{c_{\beta}}.$$

We see that this quantity can be zero if and only if $v = \psi$. 

S. Repin
This behavior of $R(v, r, \beta)$ is clearly observed on the figure below, where $\varphi = 0$ and $\psi = 1$. 
The functional $M_2$ is defined for any $v \in K_{fp}$, $\tau \in Q^*$, and $\beta > 0$. It is clear that

$$M_2(v, \tau, \beta) \leq M_1(v, \tau, \beta).$$

This fact immediately implies the following assertion.

**Theorem**

For any $\beta > 0$, $M_2(v, \tau, \beta)$ is a nonnegative functional that majorizes $\|v - u\|^2$ and vanishes if and only if

$$v = u \quad \text{and} \quad \tau = A\nabla u,$$

where $u$ is a solution of the obstacle problem.
Approximative properties

It is not difficult to prove that for any $\beta > 0$ the functional $M_2(v, \tau, \beta)$ possesses necessary continuity properties with respect to the first and second arguments. Namely,

$$M_2(v_k, \tau_k, \beta) \rightarrow 0$$

if

$$v_k \rightarrow u \quad \text{in} \quad V_0$$

and

$$\tau_k \rightarrow A\nabla u \quad \text{in} \quad Q^*.$$
If the problem contains only one obstacle (e.g., if $\psi = +\infty$), then the function $R$ has a more compact form:

$$R(v, \tau, \beta) = c_\beta \left[ |r|^2 - \left\langle \frac{v - \varphi}{c_\beta} + r \right\rangle^2 \right].$$

For a membrane problem, this case was analyzed in

Numerical tests. Example 1

We start with simple 1D tests, where the equation is $u'' = f$ on $(0, 1)$ and the boundary conditions are homogeneous. An approximate solution was computed for a uniform mesh with 60 intervals. In this example,

$$f = -2.0, \quad \varphi(x) = -0.16,$$

and the coincidence set is $[0.400, 0.600]$. The minimal value of the functional is $-0.149$. 
Exact solution, obstacle and approximate solution
In this case, the error is 0.000118. 
$M_\oplus$ computed for $y^* = G_h(\nabla u_h)$ (when the dual variable $y^*$ is computed by a simple gradient averaging procedure) gives the first upper bound 0.000647. Thus, without noticeable additional expenditures, we obtain an estimate with

$$I_{\text{eff}} = 5.473.$$ 

In this case, two parts of the Majorant have the following values: 0.000164 (duality term) and 0.000483 (generalized residual term).
Then, $M_\oplus$ was minimized with respect to the dual variable. In the table, we present values of the Majorant obtained in this process. Computational expenditures are measured by the "time unit", which is the time required for computing the approximate solution.

Table:

<table>
<thead>
<tr>
<th>Iteration</th>
<th>The majorant</th>
<th>$I_{\text{eff}}$</th>
<th>Expenditures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.000214</td>
<td>1.804802</td>
<td>.448</td>
</tr>
<tr>
<td>2</td>
<td>.000163</td>
<td>1.379279</td>
<td>.660</td>
</tr>
<tr>
<td>3</td>
<td>.000152</td>
<td>1.281185</td>
<td>.787</td>
</tr>
<tr>
<td>4</td>
<td>.000146</td>
<td>1.232812</td>
<td>.881</td>
</tr>
<tr>
<td>5</td>
<td>.000143</td>
<td>1.209496</td>
<td>.977</td>
</tr>
<tr>
<td>6</td>
<td>.000140</td>
<td>1.185637</td>
<td>1.073</td>
</tr>
<tr>
<td>7</td>
<td>.000137</td>
<td>1.158350</td>
<td>1.169</td>
</tr>
<tr>
<td>8</td>
<td>.000136</td>
<td>1.148871</td>
<td>1.243</td>
</tr>
<tr>
<td>9</td>
<td>.000134</td>
<td>1.134931</td>
<td>1.336</td>
</tr>
</tbody>
</table>
This process is depicted below

THE EFFECTIVITY INDEX

S. Repin

RICAM, Special Radon Semester, Linz, 2005.
Distributions of subinterval errors and errors computed by the Majorant are depicted on the next picture. We see that $M_\oplus$ provides a good representation of the error distribution.
Numerical tests. Example 2

Take the same problem with $f = -2.0$ and

$$\varphi(x) = -0.3x^2 - 0.06.$$ 

In this case $\Omega_{\varphi} = [0.215, 0.474]$ and the lower bound of the primal variational problem is equal to $-1.125$. An approximate solution was computed for the uniform mesh with 60 subintervals.
Figure: Example 2. Approximate solution of an obstacle problem
S. Repin
RICAM, Special Radon Semester, Linz, 2005.

LECTURES ON A POSTERIORI ERROR CONTROL
In this example, the error is 0.000158. The value of $M_\oplus$ computed for $y^* = G_h(\nabla u_h)$ gives the first (rough) upper bound of the error 0.000861. Thus, without serious additional expenditures, we obtain an estimate with

$$I_{\text{eff}} = 5.457.$$ 

Two parts of the Majorant are as follows: 0.000156 (duality term) and 0.000704 (generalized residual term).
Then, the Majorant was minimized with respect to $y$. The respective results are presented below.

Table:

<table>
<thead>
<tr>
<th>Iteration</th>
<th>The majorant</th>
<th>$I_{\text{eff}}$</th>
<th>Expenditures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.000296</td>
<td>1.877766</td>
<td>.625</td>
</tr>
<tr>
<td>2</td>
<td>.000245</td>
<td>1.551655</td>
<td>.842</td>
</tr>
<tr>
<td>3</td>
<td>.000240</td>
<td>1.521251</td>
<td>.938</td>
</tr>
<tr>
<td>4</td>
<td>.000235</td>
<td>1.492785</td>
<td>1.033</td>
</tr>
<tr>
<td>5</td>
<td>.000227</td>
<td>1.439252</td>
<td>1.165</td>
</tr>
<tr>
<td>6</td>
<td>.000218</td>
<td>1.379380</td>
<td>1.287</td>
</tr>
<tr>
<td>7</td>
<td>.000212</td>
<td>1.343913</td>
<td>1.383</td>
</tr>
<tr>
<td>8</td>
<td>.000210</td>
<td>1.334278</td>
<td>1.457</td>
</tr>
<tr>
<td>9</td>
<td>.000209</td>
<td>1.322333</td>
<td>1.542</td>
</tr>
</tbody>
</table>
In the next figure, we show the distribution of actual errors on the intervals and those computed by the Majorant.
Numerical tests. Example 3

Now we consider a 2-dimensional obstacle problem with a plane obstacle and take $\Omega$ as a unit square. In the figure below, we present an approximate solution computed by the finite element method on a uniform mesh. The elements that belong to $\Omega_\varphi$ are colored black.
Below it is shown the distribution of local (elementwise) errors and those given by the integrand of the Majorant (for $t = 1$).
If we spend more efforts on the minimization of the Majorant ($t = 3$), then the computed error distribution is practically the same as the true one.
Below we show the dependence of the effectivity index with respect to the CPU time used for the minimization of the Majorant
The elasto-plastic torsion problem

Let $\Omega$ be a bounded domain in $\mathbb{R}^2$ with Lipschitz continuous boundary $\partial \Omega$. Consider the torsion problem for a long elasto-plastic bar whose cross-section is the domain $\Omega$. If such a bar is made of an isotropic material, then the torsion problem is reduced to the following variational inequality: find $u \in K$ such that

$$\int_{\Omega} \nabla u \cdot \nabla (v - u) \, dx \geq \mu \int_{\Omega} (v - u) \, dx, \quad \forall v \in K,$$

where $\mu$ is a positive parameter,

$$K := \{ v \in V_0 \mid |\nabla v| \leq 1 \text{ a.e. in } \Omega \},$$

Elasto–plastic torsion problem has a unique solution with a free boundary that separates the sets

\[ \Omega_e := \{ x \in \Omega \mid \| \nabla u \| < 1 \} \]

and

\[ \Omega_p := \{ x \in \Omega \mid \| \nabla u \| = 1 \} , \]

which are called elastic and plastic sets, respectively.
If $\Omega$ is a 1-connected domain, then $u$ coincides with a solution of the following obstacle problem (see, e.g., A. Friedman. *Variational principles and free-boundary problems*. Wiley, NY, 1982.).

**Problem.** Find $u \in K_d$ such that

$$J(u) = \inf_{v \in K_d} J(v), \quad J(v) = \frac{1}{2} \int_{\Omega} (|\nabla v|^2 - \mu v) \, dx,$$

where $K_d := \{ v \in V_0 \mid |v| \leq d(x, \partial \Omega) \text{ for a.e. } x \in \Omega \}$

and $d(x, \partial \Omega)$ denotes the distance between $x$ and $\partial \Omega$. 
It is easy to see that we arrived at a special type obstacle problem. Now, we can use the estimates (11.16) or (11.21) with
\[ \psi = d(x, \partial \Omega) \quad \phi = -d(x, \partial \Omega). \]

In particular, if \( v \) has a fixed sign in \( \Omega \) (e.g., \( v \geq 0 \)), then (11.16) implies the estimate
\[
\| v - u \|^2 \leq \left( 1 + \frac{1}{\beta} \right) \| \nabla v - \tau \|^2 + C^2_{\Omega}(1 + \beta) \left[ \int_{\Omega_e(v)} (\text{div} \tau + \mu)^2 \, dx + \int_{\Omega_p(v)} \langle \text{div} \tau + \mu \rangle^2 \, dx \right].
\]

(11.23)

In (11.23), \( \Omega_e(v) \) and \( \Omega_p(v) \) are the elastic and plastic sets defined by the approximate solution \( v \in K \) and \( C_{\Omega} \) is a constant in the Friedrichs–Poincaré inequality.
We end up this lecture with several pictures that show some results for the elasto–plastic torsion problem.

Figure: Elastic and plastic zones for $f = 5$ computed on two different meshes
Below it is shown the distribution of local error and the distribution computed by the integrand of the Majorant for $t = 1$. 

![Graph showing the distribution of local error and the distribution computed by the integrand of the Majorant for $t = 1$.]
On this figure, it is shown the distribution of local error and the distribution computed by the integrand of the Majorant for $t = 3$. 
On this picture, we present the dependence of the effectivity index with respect to CPU time.
Lecture 12.
FUNCTIONAL A POSTERIORI ESTIMATES FOR NONLINEAR VARIATIONAL PROBLEMS
The objective of this lecture is to introduce a general scheme for deriving a posteriori error estimates by using duality theory of the calculus of variations. We consider variational problems of the form

\[ \inf_{v \in V} \{ F(v) + G(\Lambda v) \}, \]

where \( F : V \rightarrow \mathbb{R} \) is a convex lower semicontinuous functional, \( G : Y \rightarrow \mathbb{R} \) is a uniformly convex functional, \( V \) and \( Y \) are reflexive Banach spaces and \( \Lambda : V \rightarrow Y \) is a bounded linear operator.
General variational problem

Consider the general variational problem: find $u$ in a Banach space $V$ such that

$$J(u, \Lambda u) = \inf_{v \in V} J(v, \Lambda v),$$

(12.1)

where $J(v) = F(v) + G(\Lambda v)$, $F$ is a convex, lower semicontinuous functional, $G$ is a uniformly convex functional and $\Lambda : V \rightarrow Y$ is a bounded linear operator. $V$ and $Y$ are reflexive Banach spaces endowed with the norms $\| \cdot \|_V$ and $\| \cdot \|_Y$, respectively.
Dual spaces are denoted by $V^*$ and $Y^*$ with duality pairings $\langle ., . \rangle$ and $\langle\langle ., . \rangle\rangle$, respectively. The spaces $Y$ and $Y^*$ are endowed with the norms $\| . \|$ and $\| . \|^*$. We assume that

$$\| \Lambda w \| \geq c_0 \| w \|_V \quad \forall w \in V,$$  \hspace{1cm} (12.2)

where $c_0$ is a positive constant independent of $w$. In addition to $\Lambda$, we introduce its conjugate $\Lambda^* : Y^* \rightarrow V^*$. This amounts to say that

$$\langle\langle y^*, \Lambda v \rangle\rangle = \langle \Lambda^* y^*, v \rangle \quad \forall y^* \in Y^*, \ v \in V.$$ \hspace{1cm} (12.3)

$J(v, \Lambda v) := F(v) + G(\Lambda v)$ is assumed to be coercive on $V$, i.e.

$$J(v, \Lambda v) \rightarrow +\infty \quad \text{if} \quad \| v \|_V \rightarrow +\infty.$$
Primal and Dual Problems

**Problem \( \mathcal{P} \).** Find \( u \in V \) such that

\[
J(u, \Lambda u) = \inf \mathcal{P} := \inf_{v \in V} J(v, \Lambda v). \tag{12.4}
\]

The problem dual to \( \mathcal{P} \) is (see e.g. I. Ekeland and R. Temam *Convex analysis and variational problems.* North-Holland, Amsterdam, 1976.)

**Problem \( \mathcal{P}^* \).** Find \( p^* \in Y^* \) such that

\[
- J^*(\Lambda^* p^*, -p^*) = \sup \mathcal{P}^* := \sup_{y^* \in Y^*} -J^*(\Lambda^* y^*, -y^*) \tag{12.5}
\]

\[
J^*(\Lambda^* y^*, -y^*) := F^*(\Lambda^* y^*) + G^*(-y^*),
\]

where \( F^* \) and \( G^* \) are the functionals conjugate of \( F \) and \( G \), respectively.
Theorem (1)

If the functional $F$ is finite at some $u_0 \in V$ and the functional $G$ is continuous and finite at $\Lambda u_0 \in Y$, then there exists a minimizer $u$ to Problem $\mathcal{P}$ and a maximizer $p^*$ to Problem $\mathcal{P}^*$. Besides,

$$\inf \mathcal{P} = \sup \mathcal{P}^* \quad (12.6)$$

and the following duality relations hold

(i) $F(u) + F^*(\Lambda^* p^*) - \langle \Lambda^* p^*, u \rangle = 0,$

(ii) $G(\Lambda u) + G^*(-p^*) + \langle p^*, \Lambda u \rangle = 0. \quad (12.7)$

Above relations are equivalent to

(i) $\Lambda^* p^* \in \partial F(u), \quad$ (ii) $-p^* \in \partial G(\Lambda u).$
Problems with uniformly convex functionals

We recall (see Lecture 4) that a continuous functional \( G : Y \rightarrow \mathbb{R} \) is uniformly convex in a ball \( B(0, \delta) := \{ y \in Y \mid \|y\| < \delta \} \) if there exists a continuous functional \( \Phi_\delta : Y \rightarrow \mathbb{R}_+ \) such that \( \Phi_\delta(y) = 0 \) only if \( y = 0 \) is and

\[
G\left(\frac{y_1 + y_2}{2}\right) + \Phi_\delta(y_2 - y_1) \leq \frac{1}{2} (G(y_1) + G(y_2)) \quad \forall y_1, y_2 \in B(0, \delta).
\]

Usually, \( \Phi_\delta \) is given by a continuous strictly increasing function of the norm \( \|y\| \).

General form of a posteriori estimates for uniformly convex variational problems was established in

General form of the functional a posteriori estimate

**Theorem (2)**

Assume that the above conditions on $\mathbf{F}$ and $\mathbf{G}$ are satisfied and

(i) $\mathbf{G}$ is uniformly convex on a ball $B(0, \delta)$,

(ii) the solution $\mathbf{u}$ of Problem $\mathcal{P}$ and an element $\mathbf{v} \in \mathbf{V}$ are such, that $\Lambda \mathbf{u}$, $\Lambda \mathbf{v} \in B(0, \delta)$.

Then, for any $\mathbf{y}^* \in \mathbf{Y}^*$

$$
\Phi_{\delta}(\Lambda(\mathbf{v} - \mathbf{u})) \leq \mathbf{M}_\oplus(\mathbf{v}, \mathbf{y}^*) := \mathbf{D}_F(\Lambda^* \mathbf{y}^*, \mathbf{v}) + \mathbf{D}_G(\mathbf{y}^*, \Lambda \mathbf{v}) \quad (12.8)
$$

where

$$
\mathbf{D}_F(\Lambda^* \mathbf{y}^*, \mathbf{v}) := \frac{1}{2} \left( \mathbf{F}(\mathbf{v}) + \mathbf{F}^*(\Lambda^* \mathbf{y}^*) - \langle \Lambda^* \mathbf{y}^*, \mathbf{v} \rangle \right),
$$

$$
\mathbf{D}_G(\mathbf{y}^*, \Lambda \mathbf{v}) := \frac{1}{2} \left( \mathbf{G}(\Lambda \mathbf{v}) + \mathbf{G}^*(-\mathbf{y}^*) + \langle \mathbf{y}^*, \Lambda \mathbf{v} \rangle \right).
$$
Proof

Since $F$ is convex and $G$ is uniformly convex we obtain

$$
\Phi_\delta (\Lambda(v - u)) + G(\Lambda(\frac{v+u}{2})) + F(\frac{v+u}{2}) \leq 
\frac{1}{2} \left[ (F(v) + G(\Lambda v)) + (F(u) + G(\Lambda u)) \right].
$$

The element $u$ is a minimizer, therefore

$$
G(\Lambda u) + F(u) = J(u) \leq G(\Lambda \left( \frac{u+v}{2} \right)) + F(\frac{u+v}{2}),
$$

and we have

$$
\Phi_\delta (\Lambda(v - u)) + G(\Lambda u) + F(u) \leq 
\frac{1}{2} \left[ (F(v) + G(\Lambda v)) + (F(u) + G(\Lambda u)) \right].
$$
From the above we observe that

$$\Phi_\delta(\Lambda e) \leq \frac{1}{2} \left[ (F(v) + G(\Lambda v)) - (F(u) + G(\Lambda u)) \right] =$$

$$= \frac{1}{2} (J(v, \Lambda v) - J(u, \Lambda u)) \quad \forall v \in B(0, \delta).$$

In view of Theorem 1,

$$J(u, \Lambda u) = \inf \mathcal{P} = \sup \mathcal{P}^* = -F^*(\Lambda^* p^*) - G^*(-p^*).$$

Since $p^*$ is a solution of the dual problem, we have

$$-J^*(\Lambda^* p^*, -p^*) \geq -J^*(\Lambda^* y^*, -y^*) \quad \forall y^* \in Y^*,$$

so that

$$J(u, \Lambda u) \geq -F^*(\Lambda^* y^*) - G^*(-y^*).$$
Therefore

\[ \Phi_\delta (\Lambda e) \leq \frac{1}{2} (F(v) + G(\Lambda v) + F^*(\Lambda^* p^*) + G^*(-p^*)) \leq \frac{1}{2} (F(v) + G(\Lambda v) + F^*(\Lambda^* y^*) + G^*(-y^*)) . \]

However, by (12.3) we observe that

\[ \langle y^*, \Lambda v \rangle - \langle \Lambda^* y^*, v \rangle = 0 \ \forall y^* \in Y^*, \ v \in V . \]

We add this zero term to the above relation and obtain the required estimate.
The right-hand side of (12.8) is the sum of two compound functionals

\[ M_F : V^* \times V \to \mathbb{R} \quad \text{and} \quad M_G : Y^* \times Y \to \mathbb{R}. \]

They are nonnegative and vanishes if and only if \( v \) and \( y^* \) satisfy the relations (12.7)(i)–(ii).

Therefore, \( M_\oplus(v, y^*) \) is, in fact, a measure of the error in the duality relations for the pair \((v, y^*)\).
It vanishes if and only if \( v = u \) and \( y^* = p^* \).
Let the functional $F$ be uniformly convex on $V$ with a forcing functional $\varphi_\delta$. Then the "forcing functional" has the form we have

$$\Phi_\delta (\Lambda e) + \varphi_\delta (e) \leq \frac{1}{2} (J(v, \Lambda v) - J(u, \Lambda u))$$  \hspace{1cm} (12.9)$$

and, as a result, (12.8) is replaced by the strengthened estimate

$$\Phi_\delta (\Lambda e) + \varphi_\delta (e) \leq M_\oplus (v, y^*) \forall y^* \in Y^*.$$  \hspace{1cm} (12.10)
It is not difficult to verify that

\[ M_\oplus(v, y^*) - M_\oplus(v, p^*) = \]
\[ = \frac{1}{2} (F(v) + F^*(\Lambda^*y^*) - \langle \Lambda^*y^*, v \rangle + G(\Lambda v) + G^*(-y^*) + \langle y^*, \Lambda v \rangle) - \]
\[ \frac{1}{2} (F(v) + F^*(\Lambda^*p^*) - \langle \Lambda^*p^*, v \rangle + G(\Lambda v) + G^*(-p^*) + \langle p^*, \Lambda v \rangle) = \]
\[ = J^*(\Lambda^*y^*, -y^*) - J^*(\Lambda^*p^*, -p^*) \geq 0. \]

Therefore, for any \( v \) the right-hand side of (12.8) is minimal if \( y^* = p^* \). Consequently, to make the estimate effective we have to find some \( y^* \) close to \( p^* \) in \( Y^* \). A simple way to obtain a function "close" to \( p^* \) it to use duality relations. To this end, we set \( y^* = \sigma^*(v) \), where

\[ -\sigma^*(v) \in \partial G(\Lambda v) \subset Y^*. \]
In this case,

\[ M_G(\sigma^*(v), \Lambda v) = 0 \]

and we get the estimate

\[ \Phi_{\delta}(\Lambda e) \leq M_F(\Lambda^* \sigma^*(v), v) \] (12.11)

whose right-hand side depends on \( v \) only. However, the estimate (12.11) cannot be directly applied in one practically important case which we consider below.
Problems with linear functional $F$

Let

$$F(v) = \langle \ell^*, v \rangle, \quad \ell^* \in V^*. \quad (12.12)$$

Since

$$F^*(v^*) = \sup_{v \in V} \langle v^* - \ell^*, v \rangle = \begin{cases} 0 & \text{if } v^* = \ell^*, \\ +\infty & \text{if } v^* \neq \ell^* \end{cases}$$

we see that

$$M_F = \langle \ell^*, v \rangle + F^*(\Lambda^* y^*) - \langle \Lambda^* y^*, v \rangle =$$

$$= \langle \ell^* - \Lambda^* y^*, v \rangle + F^*(\Lambda^* y^*).$$
\[ M_F(\Lambda^* y^*, v) = F^*(\Lambda^* y^*) + \langle \ell^* - \Lambda^* y^*, v \rangle = \begin{cases} 0 & \text{if } y^* \in Q^*_\ell, \\ +\infty & \text{if } y^* \notin Q^*_\ell, \end{cases} \]

where
\[ Q^*_\ell := \{ y^* \in Y^* \mid \langle \Lambda^* y^*, w \rangle = \langle \ell^*, w \rangle \ \forall w \in V \} . \]

In general, above defined \( \sigma^* \) does not belong to \( Q^*_\ell \), so that the right hand side of (12.11) can become infinite. Therefore, the aim of our subsequent analysis is to obtain a modified error majorant \( \widehat{M}_\oplus(v, y^*) \) which is finite for all \( v \in V \) and all \( y^* \in Y^* \).
Let $\Pi(y) \geq 0$ for all $y \quad \Pi(0) = 0$. By $\Pi^*: Y^* \to \mathbb{R}_+$ we denote the functional conjugate of $\Pi$. For this pair the Joung–Fenchel inequality

$$\langle \xi^*, \xi \rangle \leq \Pi^*(\xi^*) + \Pi(\xi) \quad \forall \xi \in Y, \quad \xi^* \in Y^*$$

holds.

For the sake of simplicity, we set

$$\Pi(y) = \pi(\|y\|) \quad \text{and} \quad \Pi^*(y^*) = \pi^*(\|y^*\|_*).$$
General form of the Deviation Majorant

\[ \Phi_\delta (\Lambda(v - u)) \leq M_D(y^*, \Lambda v) + M_R(y^*), \quad (12.13) \]

where

\[ M_D(y^*, \Lambda v) = D_G(y^*, \Lambda v) + \frac{1}{2} \pi \left( \|G'(y^*) - \Lambda v\| \right), \quad (12.14) \]

\[ M_R(y^*) = \inf_{q^* \in Q^*} \pi^* (\|q^* - y^*\|_*). \quad (12.15) \]
Set $\Lambda v := \nabla v$ and consider variational problems for the functional

$$J(v, \nabla v) = \int_\Omega (g(\nabla v) + f(v)) \, dx.$$ 

Now $G$ and $F$ are integral functionals whose integrands $g : \mathbb{R}^d \to \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$ are convex differentiable functions. Denote their conjugate functions $g^*$ and $f^*$, respectively. The spaces $Y$ and $Y^*$ we identify with the Lebesque spaces $L^\alpha(\Omega, \mathbb{R}^d)$ and $L^{\alpha*}(\Omega, \mathbb{R}^d)$, where $\alpha_* = \frac{\alpha}{\alpha - 1}$, $\alpha > 1$ is taken such that the above integral has sense. In the considered case,

$$\langle y^*, y \rangle := \int_\Omega y^* \cdot y \, dx \quad \text{and} \quad \Lambda^* y^* := -\text{div} y^* \in V^*.$$
Example 1.

Let \( g(y) = \frac{1}{2} A y \cdot y \), where \( A \) is a symmetric real matrix satisfying the conditions

\[
\nu_1 |\eta|^2 \leq A \eta \cdot \eta \leq \nu_2 |\eta|^2 \quad \forall \eta \in \mathbb{R}^d,
\]

for some \( \nu_2 \geq \nu_1 > 0 \). It is straightforward to check that the functional \( G \) is uniformly convex on any ball. The two parts of the error majorant \( M_\oplus \) (cf. (12.8)) are given by the relations

\[
D_G(y^*, \Lambda v) = \frac{1}{4} \int_\Omega \left( A \nabla v \cdot \nabla v + A^{-1} y^* \cdot y^* - \nabla v \cdot y^* \right) dx,
\]

\[
D_F(\Lambda^* y^*, v) = \frac{1}{2} \int_\Omega (f(v) - y^* \cdot \nabla v) dx + \frac{1}{2} \sup_{w \in V} \int_\Omega (y^* \cdot \nabla w - f(w)) dx.
\]
If the function \( f^*(\mathbf{v}) \) is summable then we arrive at a more symmetric expression

\[
\mathbf{D}_F(\mathbf{v}, \mathbf{y}^*) \leq \frac{1}{2} \int_\Omega (f(\mathbf{v}) + f^*(-\text{div}\mathbf{y}^*) - \mathbf{y}^* \cdot \nabla \mathbf{v}) \, d\mathbf{x}.
\]

In particular, if \( f(\mathbf{v}) = \frac{\lambda}{2} \mathbf{v}^2 + \mu \mathbf{v} \), where \( \mu \in \mathbb{R} \) and \( \lambda \in \mathbb{R}_+ \), then

\[
f^*(\mathbf{v}^*) = \frac{1}{2\lambda} (\mathbf{v}^* - \mu)^2.
\]

We note that this case is related to the equation

\[
\text{div}\mathbf{A} \nabla u - \lambda u + \mu = 0.
\]
In this case, $\alpha = 2$ and for any

$$\mathbf{y}^* \in H(\Omega, \text{div})$$

we obtain

$$M_F(v, \mathbf{y}^*) \leq \frac{1}{4\lambda} \| \lambda v + \text{div}\mathbf{y}^* + \mu \|_{\Omega}^2,$$

Both functionals $G$ and $F$ are uniformly convex and we can take

$$\Phi(\nabla e) = \frac{1}{4} \int_{\Omega} A \nabla e \cdot \nabla e \, dx,$$

$$\varphi(e) = \frac{\lambda}{4} \int_{\Omega} |e|^2 \, dx.$$
Thus, we arrive at the estimate of deviation in the following form:

\[
\int_{\Omega} A \nabla (v - u) \cdot \nabla (v - u) \, dx + \lambda \|v - u\|_{\Omega}^2 \leq \int_{\Omega} (A^{-1} y^* + \nabla v) \cdot (y^* + A \nabla v) \, dx + \frac{1}{\lambda} \|\lambda v + \text{div} y^* + \mu\|_{\Omega}^2.
\]
Example 2

Consider the problem with

\[ G(y) = \frac{1}{2} (A y, y) + \Psi(y), \quad F(v) = \langle \ell, v \rangle, \]

where \( \Psi : Y \to \mathbb{R} \) is a convex continuous functional. Note that if \( \Psi \equiv 0 \), then

\[
D_G(\Lambda v, p^*) = \frac{1}{2} (A \Lambda v, \Lambda v) + \frac{1}{2} (A \Lambda u, \Lambda u) - (\Lambda v, A \Lambda u)
\]

\[
= \frac{1}{2} \| \Lambda (v - u) \|^2.
\]

We will measure the error in terms of the above norm generated by the operator \( A \).
In this case, the deviation estimate is as follows:

\[
\frac{1}{2} \| \Lambda (v - u) \|^2 \leq (1 + \beta) D_G(\Lambda v, y^*) + \\
+ \left(1 + \frac{1}{\beta}\right) C_\Omega^2 \| \Lambda^* y^* + \ell \|^2,
\]

where \( C_\Omega \) depends on \( \Omega \) and \( A \).
A consequent exposition functional type a posteriori error estimates for nonlinear variational problems can be found in the papers


and in the book

We end up this lecture course with concise exposition of two important problems closely related with functional type a posteriori estimates.

- Evaluation of errors in terms of local quantities;
- Evaluation of modeling errors.
Indication of local errors

Integrand of the Majorant is a good error indicator.

\[ M_\oplus = \int_\Omega \mu(x) \, dx. \]
It is proved that if $M_\oplus \to \| u - v \|$, then

$$\mu(x) \to e(x) := |\nabla (u - v)(x)|$$

in the sense of measures

This means that for any $\delta > 0$

$$\text{meas} E_\delta \to 0$$

where $E_\delta := \{ x \in \Omega : |\mu(x) - e(x)| > \delta \}$. 
Guaranteed upper bounds of local errors

GENERAL PRINCIPLE: Guaranteed upper bounds for

\[ \Phi(u - v) = \|u - v\|_\omega \quad \text{and} \quad \Phi(u - v) = (\ell, u - v), \quad \ell \in V^* \]

are obtained by projection of the functional a posteriori estimate onto a certain subspace.


Example. Local estimate for diffusion problem

Let \( \omega \subset \Omega \) and \( V_\omega := \{ H^1_0(\Omega) \mid v = \text{const in } \omega \} \). An upper bound of the local error is given by the estimate

\[
\| \nabla (u - v) \|^2_\omega \leq M_{\oplus \omega} := \\
\inf_{w \in V_\omega} \left\{ (1 + \beta) \| \nabla (v - w) - y \|^2 + \frac{1 + \beta}{\beta} C^2_\Omega \| \text{div} y + f \|^2 \right\}.
\]

(12.16)

Here \( \beta > 0 \) and \( y \in H(\text{div}, \Omega) \).
Quality of the estimate

\[ \| \nabla (u - v) \|_\omega^2 \leq M_\omega^\oplus(v) \leq \| \nabla (u - v) \|_\omega^2 + I_\omega(v), \]

where

\[ I_\omega(v) := \inf_{\phi \in V_\omega} \| \nabla (u - v - \phi) \|_{\Omega \setminus \omega}^2 \]

If \( v - u = \mu = \text{const} \) on \( \partial \omega \), then the function

\[ \phi := \begin{cases} u - v & \text{in } \Omega \setminus \omega; \\ \mu & \text{in } \omega \end{cases} \]

belongs to \( V_\omega \) and, therefore, \( I_\omega(v) = 0 \).
Errors in terms of goal oriented quantities

In the two above cited papers a guaranteed upper bounds for goal–oriented errors were also derived. Basic observation

\[ |\langle \ell, u - v \rangle| = |\langle \ell, u - v + \varphi \rangle| \quad \forall \varphi \in \mathbf{V}_{0\ell}(\Omega), \]

where \( \mathbf{V}_{0\ell}(\Omega) := \{ \varphi \in \mathbf{H}^1_0(\Omega) | \langle \ell, \varphi \rangle = 0 \} \). Therefore,

\[ |\langle \ell, u - v \rangle| \leq \| \ell \| \inf_{\varphi \in \mathbf{V}_{0\ell}} \| u - v + \varphi \|. \]

Since \( v - \varphi \) can be viewed as a certain approximation, we apply the functional error estimate to the left hand side and obtain a guaranteed bound for the goal–oriented quantity.
Since there are no "absolutely exact" mathematical models, modeling errors always exist in real life mathematical modeling. How to estimate their influence?
Let us shortly consider this question in connection with one type of modeling errors that arise in \textbf{dimension reduction models}.

\[ \Omega = \hat{\Omega} \times (-d, +d), \]

\( \hat{\Omega} \in \mathbb{R}^2 \) with boundary \( \gamma \),

\[ d \ll \text{diam}(\hat{\Omega}) := \sup_{(x_1, x_2) \in \hat{\Omega}} |x_1 - x_2|. \]
A more detailed exposition can be found in


Key idea

We can consider a solution of a $d - 1$-dimensional model as an approximate solution of the $d$-dimensional one. Since deviation estimates are valid for all conforming approximations in the energy space, we may somehow project $d - 1$-dimensional solution to the energy space of the $d$-dimensional problem and use the Deviation Estimate for the estimation of the respective error.
Example. Plain stress problem as a model of 3D linear elasticity one

Here, 3D solution \((u, \sigma)\) is approximated by the 2D one \((\hat{u}, \hat{\sigma})\), where \(\hat{u} = (\hat{u}_1, \hat{u}_2)\) and \(\hat{\sigma}\) is a \(2 \times 2\) tensor. Let

\[
\tilde{u} = (\hat{u}_1, \hat{u}_2, \phi(x_1, x_2, x_3)); \quad \tilde{\sigma}_{\alpha\beta} = \hat{\sigma}_{\alpha\beta}, \quad \tilde{\sigma}_{3\alpha} = 0,
\]

where \(\varphi \in H^1(\Omega)\) and meets boundary conditions \((u_{03})\) on the Dirichlet part of \(\partial \Omega\). Then, we have
An estimate of the dimension reduction error

\[ C_\varepsilon \| \varepsilon (\tilde{u} - u) \|^2_\Omega + C_\tau \| \tilde{\sigma} - \sigma \|^2_\Omega \leq \]

\[ \leq \left( \frac{K_0}{2} + \frac{2\mu}{3} \right) \int_\Omega \left( \rho (\hat{u}_{1,1} + \hat{u}_{2,2}) + \varphi,3 \right)^2 dx + \frac{\mu}{2} \int_\Omega \left( \varphi'^2 + \varphi'^2 \right) dx \]

See the proof in

Here \( \mu \) and \( K_0 \) are elasticity coefficients \( \rho = \frac{3K_0 - 2\mu}{3K_0 + 4\mu} \),

\[ C_\varepsilon = \min \left\{ \frac{1}{\mu}, \frac{2}{3\mu} \right\}, \; C_\tau = \min \left\{ 4\mu, 6K_0 \right\}. \]
General diffusion type equations. Non-plane domains

computable estimates for dimension reduction models we derived for general diffusion type problem

\[
\text{div}(A \nabla u) + f = 0, \quad u = u_0 \partial \Omega
\]

for ”thin” domains of the form \( \Omega = \Gamma \times [-d, d] \), where \( \Gamma \) is a certain surface in 3D.
Below is the list of publications related to the topics discussed in the Lectures. In brackets, it shown the number of a lecture related to a particular publication.

S. Repin

RICAM, Special Radon Semester, Linz, 2005.


M. Ainsworth and J. T. Oden, A posteriori error estimation in finite element analysis, Wiley and Sons, New York, 2000. (L2)


I. Babuška. Courant element: before and after, in *Fifty years of Courant element*, M. Křížek, P. Neittaanmäki and R. Stenberg (Eds.), Marcel Dekker, 1994, 37-51. (L1,2)


I. Babuška, T. Strouboulis, S. K. Gangaraj and C. S. Upadhyay. Pollution-error in the $h$-version of the FEM and the local quality of


adaptation minimization by means of an integral of $\int \nabla (u - u_h)$ where $u$ is replaced by a polynomial interpolation)


*Funktionanalysis und numerische mathematik.* Springer-Verlag, Berlin, 1964. (L1,10)


E. Gorshkova and S. Repin. Error control of the approximate solution to the Stokes equation using a posteriori error estimates of functional type.


**LECTURES ON A POSTERIORI ERROR CONTROL**


M. Křížek and P. Neittaanmäki. *Finite Element Approximations of Variational Problems and Applications*. Wiley and Sons, New York, 1990. (L1,2)


S. G. Mikhlin. *Constants in some inequalities of analysis.* Wiley and Sons, Chister–New York, 1986. (L1,3)


P. Neittaanmäki and M. Křížek. On $O(h^4)$ superconvergence of piecewise bilinear FE-approximations. In Teubner-Texte Math. 107, Teubner, 1988, 250-255. (L2)


S. Repin. A posteriori error estimation for variational problems with power growth functionals based on duality theory, Zapiski Nauchnych Seminarov POMI, 249(1997), 244-255. (L12)


W. Ritz. Über eine neue Methode zur Lözing gewisser Variationsprobleme der Mathematischen Physics, *J. Reine Angew. Math.*, 135(1909), 1-61. (L1)


(L2, mathematical style presentation of the residual and averaging methods)


S. Zaremba. Sur un problème toujours possible comprenant, à titre de cas particuliers, le problème de Dirichlet et celui de Neumann, *J. math. pures et appl.* 6(1927), 2, 127-163. (L2)


