UNIVERSITÄT JYVÄSKYLÄ
MATHEMATISCHES INSTITUT
BERICHT 74

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICS
REPORT 74

VARIATIONAL CRIMES AND EQUILIBRIUM FINITE ELEMENTS IN THREE-DIMENSIONAL SPACE

SERGEY KOROTOV

JYVÄSKYLÄ
1997

UNIVERSITÄT JYVÄSKYLÄ
MATHEMATISCHES INSTITUT
BERICHT 74

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICS
REPORT 74

VARIATIONAL CRIMES AND EQUILIBRIUM FINITE ELEMENTS IN THREE-DIMENSIONAL SPACE

SERGEY KOROTOV

To be presented, with the permission of the Faculty of Mathematics and Natural Sciences of the University of Jyväskylä, for public criticism in Auditorium Blomstedt, Villa Rana, on October 25th, 1997, at 12 o'clock noon.

JYVÄSKYLÄ
1997

Editor: Pertti Mattila
University of Jyväskylä
Department of Mathematics
P.O. Box 35
FIN–40351 Jyväskylä
Finland

URN:ISBN:978-952-86-0195-1
ISBN 978-952-86-0195-1 (PDF)
ISSN 0075-4641

University of Jyväskylä, 2024

ISBN 951-39-0075-4
ISSN 0075-4641

University Printing House
Jyväskylä 1997

Acknowledgements

First of all, I would like to express my sincere gratitude to Professor Pekka Neittaanmäki for inviting me in autumn of 1995 to the Laboratory of Scientific Computing in Jyväskylä, Finland, in order to begin my PhD studies, for his kind advice, continuous attention, support and encouragement during the countless steps of writing this work. I am also deeply indebted to Professor Michal Křížek from Academy of Sciences of the Czech Republic for his inspiring and excellent guidance during the entire course of my work and invitations to make part of the researches in his institute.

On this occasion I wish to thank Professor Timo Tiihonen (University of Jyväskylä, Finland) and Professor Ivan Hlaváček (Academy of Sciences of the Czech Republic) for many various fruitful discussions during the preparation of the papers of the author. I am thankful to Doctor Barbara Wohlmuth from Mathematical Institute in Augsburg, Germany, and Professor István Faragó from Eötvös Loránd University in Budapest, Hungary, for their comments and suggestions on the structure and the contents of the work.

Taking this chance, I thank my first teachers of mathematics: Professor Gennadi V. Demidenko from Institute of Mathematics (Novosibirsk, Russia) for excellently teaching me the partial differential equations and Sobolev spaces theory, Professors Valeri P. Il'in and Yuri I. Kuznetsov from Computing Center in Novosibirsk, Russia, and Professor Owe Axelsson from Nijmegen, the Netherlands, for introducing me to the Numerical Mathematics field.

I am also indebted to my colleague Doctor Liu Liping for carefully reading of the preliminary version of the manuscript and for valuable comments. Special thanks to Mrs. Heidi Laaksonen and Doctor Vesa Ruuska for helping me to solve many practical problems during my stay in Finland, to Ms. Marja-Leena Rantalainen for her kind advice during the typing of the work, to Mrs. Tuula Blåfield for her linguistic comments.

This work was carried out at the Laboratory of Scientific Computing, Department of Mathematics, University of Jyväskylä. For financial support, I would like to acknowledge COMAS Graduate School, Jyväskylä, Finland, and also the Grant no. 201/97/0217 of the Grant Agency of the Czech Republic, which supported partly my visits to Mathematical Institute in Prague.

Finally, I would like to express my deepest appreciation and sincere feelings of gratitude to my wife Tatiana and my daughter Alina for their patience and understanding. I am also grateful to all my friends for their continuous support.

Jyväskylä, June 1997

Sergey Korotov

To my wife Tatiana and daughter Alina

Contents

Preface

The theory of the finite element method (FEM) has been developed during the last fifty years. The discovery of the FEM is usually attributed to Prof. Richard Courant in 1943 (see [Courant, p. 21]). Nevertheless, in [Ciarlet, Lions, p. 5] there are some older references to finite element-like methods. The first monograph on the FEM belongs to Synge [Synge] in 1957.

The notion *element* was introduced in the 1950's by aerospace engineers performing elasticity computations. They divided a continuum into smaller pieces – called elements. The notion *finite element* was introduced by mathematicians later, in the 1960's.

At the beginning the finite element (FE) analysis arose essentially as a discipline for solving problems in the structural engineering, and its role in that field is still of fundamental importance today. It was found soon, however, that the method had implications far beyond those originally considered and that it presented a very general and powerful technique for the numerical solution of differential equations. This aspect of the FE analysis has been developed in resent years, with the result that at the present time, it is probably as important as the traditional applications. References to the growing literature on the FEM are collected, e.g., in [Mackerle], [Noor].

The main advantage of the FEM is that it enables us to simulate many physical processes on a computer. Thus, it can substitute for the creation of expensive technical models (prototypes) or for the performance of complicated measurements. For instance, in designing and manufacturing electrical devices we can compute their electromagnetic, temperature or stress fields by the FEM, see, e.g., the monographs [Křížek, Neittaanmäki, 1990] and [Křížek, Neittaanmäki, 1996].

A great deal of progress has been made in FEM software. The whole computational process can be essentially automated; including the following steps:

1. preprocessing of input data,
2. generation of triangulations,
3. assembling FE-matrices,
4. solving discrete problems,
5. postprocessing of output data,
6. a posteriori error estimates,
7. graphical illustration of results.

Nevertheless, many theoretical questions, related to the foundations of the method and being born by practical problems and their needs, are still open. Moreover, e.g., three-dimensional problems are usually very difficult to solve and always present a challenge to the researcher. In this dissertation the author presents his own works devoted to the following special questions of the FE analysis: an analysis of variational crimes for the nonlinear heat conduction equation in 3D, a construction of the three-dimensional equilibrium finite elements and a construction of a strongly regular family of triangulations for planar domains with piecewise smooth boundaries.

First, we briefly describe the main structure of the dissertation, then outline each of the mentioned above topics and the own contributions to the fields.

The work consists of four parts, which are divided into 10 chapters.

In Part I, consisting of two chapters, we present modern formulations of correct boundary value problems, being convenient for using numerical methods, introduce the main ideas of Rayleigh–Ritz, Bubnov–Galerkin and finite element methods. Necessary theorems and definitions from the Sobolev space theory and the functional analysis are given there as well.

Three works by the author ([Korotov, Křížek], [Korotov, 1997a] and [Korotov, 1997b]) form the basis of Parts II, III and IV, respectively.

In Part II we generalize the FE analysis of variational crimes, which arise in solving a nonlinear heat conduction problem in planar domains with a piecewise curved boundary (see [Feistauer, Křížek, Sobotíková]) to the three-dimensional domains with curved smooth boundaries. In particular, we approximate a smooth boundary by a polyhedral one, we use appropriate numerical quadrature formulae to evaluate all integrals and, finally, we interpolate boundary conditions. Doing these approximations, we commit three kinds of so-called *variational crimes*, in virtue of which the used finite element method becomes nonconforming.

A detailed analysis of variational crimes for linear boundary value problems can be found, e.g., in [Ciarlet], [Ciarlet, Raviart], [Strang]. Its extension to a class of nonlinear plane elliptic problems of monotone type was first done in [Feistauer, Ženíšek, 1987]. This analysis was later generalized into several directions: to pseudomonotone operators [Feistauer, Křížek, Sobotíková], [Feistauer, Ženíšek, 1988], to nonlinear elliptic problems with discontinuous coefficients [Feistauer, Sobotíková, 1990], [Ženíšek, 1990a], to nonlinear boundary conditions [Feistauer, Felcman, Rokyta, Vlášek], [Feistauer, Najzar], to polyhedral domains [Feistauer, Křížek, Sobotíková], [Křížek, Lin], etc.

Some of these results are also surveyed in the monographs [Feistauer, 1993], [Křížek, Neittaanmäki, 1996], [Ženíšek, 1990b]. However, the above mentioned references do not contain any finite element analysis of a nonlinear elliptic problem in three-dimensional domains with curved boundaries.

We suppose that $\Omega \subset \mathbf{R}^3$ is a bounded domain with a smooth boundary. There are several approaches how to treat the curved boundary $\partial\Omega$. The first one is to employ isoparametric elements. However, they do not have a simple form in $\mathbf{R}^3$ (see [Bernardi] and [Ciarlet]). The ansatz shape functions are nonpolynomial; in favorite cases rational, but, in general, they are quite complicated. Note that for a small discretization parameter h it is not possible to decompose $\overline{\Omega}$ into tetrahedral elements having at most one face curved (cf. [Křížek, Neittaanmäki, 1990, p. 76]). Thus, isoparametric elements with at least two curved faces are used (see [Lenoir]).

Another approach is to approximate Ω by a polyhedron Ω_h and then decompose $\overline{\Omega}_h$ into straight (non-curved) elements. Since we usually have no a priori information about the regularity of the true solution of a nonlinear problem, lower order finite elements on Ω_h have to be applied. This manner is often used in practical calculations, but for a theoretical finite element analysis the entire domain Ω should be taken into account.

In [Knobloch], the shape of tetrahedral elements near $\partial\Omega_h$ are slightly changed in such a way that Ω is completely covered by them.

We will present a different approach. We assume for simplicity that Ω is convex and $\Omega_h \subset \Omega$. The set $\overline{\Omega} \setminus \Omega_h$ is decomposed into two kinds of special elements–hat and slice elements, see Chapter 6. They are not applied for computer implementation, but only to prove the convergence of approximate solutions on polyhedral domains in the W_2^1-norm to the true solution, that is the main result of Part II.

Also, at the beginning of this part, in Chapter 3, we present a survey of the other main results obtained for the FE analysis of nonlinear heat conduction equations of the same kind.

The result of Part III is a construction of finite element subspaces of the spaces of divergence-free functions. Such a problem is frequently met when we treat numerically some phenomena in continuum mechanics, electromagnetism, heat and fluid flow problems, etc.

Namely, we shall describe an internal finite element approximation of the following space which appears in variational formulations of a considerable number of problems, see, e.g., [Girault, Raviart], [Hlaváček, Křížek, 1984], [Křížek, Neittaanmäki, 1990], [Nedelec], [Temam]:

$$H_0(\mathrm{div}^0;\Omega) = \left\{\vec{q} \in [L^2(\Omega)]^d \mid (\vec{q}, \nabla z)_0 = 0 \ \forall z \in H^1(\Omega)\right\}, \quad d = 2, 3.$$

We will deal only with the three-dimensional case. Note that in 2D the mapping $\mathrm{curl} : H_0^1(\Omega) \to H_0(\mathrm{div}^0;\Omega)$ is bijective for simply connected planar domains with Lipschitz continuous boundaries (see [Girault, Raviart]), but in 3D the situation is different: the mapping curl acts on the vector-functions already. Nevertheless, under certain conditions on the given domain and on the construction of the finite element subspaces we again prove the bijectivity of this mapping. This simplifies the procedure of a construction of the finite element basis functions in $H_0(\mathrm{div}^0;\Omega)$.

The question of Part IV is considered mainly due to pure theoretical needs. We prove the existence of a strongly regular family of decompositions into triangles for planar domains with piecewise smooth boundaries. This regularity property (see Section 2.7 for several equivalent formulations) implies getting a priory error estimates and is usually committed in the literature as existing one. Several constructive proofs can be found in Russian literature, see [Korneev, 1977], [Korneev, 1979b] for planar and space domains with C^2-smooth boundaries. These constructions are based on the rectangular meshes. Note that we widely use such regularity property in Part II of the dissertation.

Another approach based on triangulations of a circle and a ball is developed in [Matsokin, 1975a] and [Matsokin, 1975b]. Nevertheless, according to the knowledge of the author there is no strict proof of an existence of such triangulations for planar domains with piecewise smooth boundaries. Moreover the algorithm proposed by the author is different from the mentioned ones.

List of symbols

	equality by the definition
Φ	energy functional
$\arg\min\ \Phi(v)$	set of functions v which give the functional Φ the smallest value
sup	supremum
inf	infimum
∞	infinity
H	Hilbert space
d	dimension
$C,\ \widehat{C},\ C_i, \ldots$	generic constants (different at each occurrence)
$\mathbf{R}^d$	d-dimensional Euclidean space
	scalar product in $\mathbf{R}^d$
(a, b)	open interval in $\mathbf{R}^1$
$[a, b]$	closed interval in $\mathbf{R}^1$
Ω	problem domain (bounded connected open set in $\mathbf{R}^d$)
$\mathrm{meas}_d\, \Omega$	d-dimensional Lebesgue measure of Ω
$\overline{\Omega}$	closure of Ω
$\partial\Omega$	boundary of Ω
n	outward unit normal to $\partial\Omega$
$\mathcal{L}$	set of all bounded domains with a Lipschitz continuous boundary
$\Gamma_0,\ \Gamma_1,\ \Gamma_2$	parts of the boundary $\partial\Omega$
Γ_0	part of boundary where the Dirichlet condition is given
Ω_h	polyhedral approximation of the domain Ω
dist	distance
$v\vert_K$	restriction of function v to set K
$n \cdot w$	normal component of vector function w on $\partial\Omega$
$n \wedge w$	tangential component of vector function w on $\partial\Omega$
$\partial_n v$	normal derivative of v on $\partial\Omega$
$D^m v$	m-th generalized derivative of v (m–multi-index)
$\partial_j v$	first generalized derivative of v $(j \in \{1, ..., d\})$
Δ	Laplace operator
grad	gradient
div	divergence
rot	rotation of vector function for $d = 2, 3$
curl	rotation of scalar (vector) function for $d = 2$ $(d = 3)$
span	linear span
$\det A$	determinant of matrix A

$\imath$	imaginary unit
δ_{ij}	Kronecker's symbol: $\delta_{ij} = 1$ for $i = j$, $\delta_{ij} = 0$ otherwise
$\exists$	there exist(s)
$\forall$	for all
a.a.	almost all
a.e.	almost everywhere
$P_k(\Omega)$	space of polynomials of degree at most k defined on Ω
$L^p(\Omega)$, $p \in [1, \infty)$	Lebesgue space of measurable functions v defined on Ω for which $\int_\Omega \lvert v(x)\rvert^p \, dx$ is finite
$L^p(\partial\Omega)$, $p \in [1, \infty)$	Lebesgue space of measurable functions u defined on $\partial\Omega$ for which $\int_{\partial\Omega} \lvert u(s)\rvert^p \, ds$ is finite
$L^\infty(\Omega)$	Lebesgue space of measurable essentially bounded functions defined on Ω
$C(\overline{\Omega})$	space of functions continuous in $\overline{\Omega}$
$C^k(\overline{\Omega})$	space of functions whose all the classical derivatives up to order k belong to $C(\overline{\Omega})$
$C^\infty(\Omega)$	space of infinitely differentiable functions in Ω
$C_0^\infty(\Omega)$	space of infinitely differentiable functions with compact support in Ω
$H^k(\Omega)$	Sobolev space of functions whose generalized derivatives up to order k belong to $L^2(\Omega)$
$H_0^1(\Omega)$	space of functions from $H^1(\Omega)$ whose traces vanish on $\partial\Omega$
$H^{1/2}(\partial\Omega)$	space of traces of all functions from $H^1(\Omega)$
$W_p^k(\Omega)$	Sobolev space of functions whose generalized derivatives up to order k belong to $L^p(\Omega)$
$H(\mathrm{div}^0)$	space of divergence-free functions
$H(\mathrm{rot}^0)$	space of rotation-free functions
V	Banach space of test functions
V^*	space of linear continuous functionals on V
$\rightharpoonup$	weak convergence
$\rightarrow$	strong convergence
$\langle b, v\rangle$	scalar product or value of functional $b \in V^*$ at point $v \in V$
v	test function
u	classical or weak (variational, generalized) solution
$\lVert \cdot \rVert_V$	norm in V
$(\cdot\,,\cdot)_V$	scalar product in V
$\lVert \cdot \rVert_{k,\Omega}$	norm in $H^k(\Omega)$
$\lvert \cdot \rvert_{k,\Omega}$	seminorm in $H^k(\Omega)$
$(\cdot\,,\cdot)_{k,\Omega}$	scalar product in $H^k(\Omega)$
$\lVert \cdot \rVert_{k,p,\Omega}$	norm in $W_p^k(\Omega)$
$\langle\cdot\,,\cdot\rangle_{0,\partial\Omega}$	scalar product in $L^2(\partial\Omega)$
$\lvert \cdot \rvert$	absolute value
$\lVert \cdot \rVert$	Euclidean norm

$(\cdot\,,\cdot)$	Euclidean scalar product
$a(\cdot\,,\cdot)$	bilinear or sesquilinear form
$F(\cdot)$	linear form
K	element, simplex
P_K	space of shape functions
$\dim P_K$	dimension of P_K
h_K	diameter of K (diam K)
h	discretization parameter
$\mathcal{T}_h,\ \mathcal{T}^{(i)}$	triangulation (partition, decomposition)
V_h	finite element space
$\pi_h v$	V_h-interpolant of v
$\operatorname{supp} v$	support of function v
$\{v^i\}_{i=1}^N$	basis functions in V_h
$\mathcal{O}(\cdot)$	Landau's symbol: $f(\alpha) = \mathcal{O}(g(\alpha))$, if $\lvert f(\alpha)\rvert \le C\lvert g(\alpha)\rvert$ as $\alpha \to 0$ or $\alpha \to \infty$
$\emptyset$	empty set
$x \in A$	element x belongs to set A
$x \notin A$	element x does not belong to set A
$\{x \in A \mid \mathcal{P}(x)\}$	set of all elements x from A which possess property $\mathcal{P}(x)$
$A \subset B$	A is subset of set B
$A \cap B$	intersection of sets A and B
$A \cup B$	union of sets A and B
$A \backslash B$	subtraction of B from A
$f\colon A \to B$	function f from A to B
$x \mapsto f(x)$	function which assigns value $f(x)$ to x
$\square$	Halmos symbol

PART I

Auxiliary Results from Sobolev Space Theory, Functional Analysis and Finite Element Analysis

On contents of Part I

Part I consists of two chapters. In Chapter 1 we present modern formulations of correct boundary value problems, main results from the Sobolev space theory and several auxiliary theorems from the functional analysis.

The finite element method is briefly discribed in Chapter 2. Also several special questions of the FE analysis, being topics of the main parts of the dissertation, are discussed there.

The material of this part is auxiliary for Parts II, III and IV.

Chapter 1
Modern formulations of correct boundary value problems

1.1. Variational principles of mechanics[1]

Consider as an example the following homogeneous Dirichlet problem for the Poisson equation in a bounded planar domain Ω with a boundary $\partial\Omega$:

$$-\Delta u \equiv -u_{x_1x_1} - u_{x_2x_2} = f(x_1, x_2) \quad \text{in } \Omega \tag{1.1.1}$$

$$u = 0 \qquad \text{on } \partial\Omega. \tag{1.1.2}$$

It is well-known that the classical solution of the above problem has to be continuous in $\overline{\Omega}$ function, satisfying the boundary condition (1.1.2) and have continuous in Ω derivatives $u_{x_1x_1}$ and $u_{x_2x_2}$ such that (1.1.1) holds.

The equations (1.1.1)–(1.1.2) present a mathematical model of the following physical problem: find the equilibrium position of a thin elastic homogeneous membrane, fixed along the boundary and subject to the load $f(x)$.

Unfortunately, this problem has no solution even when f is a continuous in $\overline{\Omega}$ function, see, e.g., [Mikhailov, p. 246]. Unconveniency of such a model forced mathematicians to recall that the equation (1.1.1) has been obtained while minimizing the following energy functional:

$$\begin{aligned}\Phi(v) &\equiv \int_\Omega \{v_{x_1}^2 + v_{x_2}^2 - 2vf\}\, dx \\ &\equiv (|\nabla v|^2 - 2vf, 1)_{0,\Omega},\end{aligned} \tag{1.1.3}$$

i.e., derived from the variational problem:

$$u = \ \arg\min \Phi(v), \ \ v \in V, \tag{1.1.4}$$

where $V = \{v \in C^1(\overline{\Omega}) \mid v = 0 \text{ on } \partial\Omega\}$.

To study variational problems similar to (1.1.4), we will define in $V \times V$ a scalar product:

$$(u, v)_{1,\Omega} \equiv (u_{x_1}v_{x_1} + u_{x_2}v_{x_2}, 1)_{0,\Omega} \equiv (u, v)_V. \tag{1.1.5}$$

We recall here the axioms of the real scalar product $(\cdot,\cdot)$:

$$\begin{aligned}(u, v) &= (v, u), \\ (\alpha_1 u_1 + \alpha_2 u_2, v) &= \alpha_1(u_1, v) + \alpha_2(u_2, v), \\ (u, u) &\equiv \|u\|^2 \geq 0, \\ \|u\| = 0 &\iff u = 0,\end{aligned} \tag{1.1.6}$$

[1] While introducing a material of this section we give simultaneously several definitions and notations. Some of them, more important ones, will be formulated in strict mathematical form later.

where α_1, α_2 are any scalars (by a scalar we mean a real or a complex number).

The following important properties of the scalar product are valid

$$|(u,v)| \leq \|u\|\,\|v\| \quad \text{(Cauchy–Schwarz–Bunyakovskii inequality)}, \tag{1.1.7}$$

$$\|u+v\| \leq \|u\| + \|v\| \quad \text{(triangle inequality)}, \tag{1.1.8}$$

$$(u,v) = 0 \iff \|u+v\|^2 = \|u\|^2 + \|v\|^2 \quad \text{(Pythagoras rule)}, \tag{1.1.9}$$

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2) \quad \text{(parallelogram rule)}, \tag{1.1.10}$$

i.e., the mapping $\|\cdot\|$ is, in fact, a norm.

For the defined above space V, the *Poincaré–Steklov inequality* is valid (see, e.g., [Mikhailov, 1978]):

$$\|u\|_{0,\Omega} \leq C|u|_{1,\Omega}, \tag{1.1.11}$$

where $u \in V$, $\|u\|^2_{0,\Omega} \equiv \|u\|^2_{L^2(\Omega)} \equiv (u^2, 1)_{0,\Omega}$, $|u|_{1,\Omega} = \|u\|_V$ and C is an independent constant.

By V^* we will denote a linear normed space of linear continuous functionals defined on V (so-called a dual space of V). Then, if $\|f\|_{0,\Omega} < \infty$ and defining

$$l(v) \equiv (f,v)_{0,\Omega}, \tag{1.1.12}$$

we have $l \in V^*$, $\|l\| \equiv \sup\limits_{v\in V} \left\{ \dfrac{|l(v)|}{\|v\|_V} \right\} \leq C\|f\|_{0,\Omega}$ and

$$\Phi(v) = \|v\|^2_V - 2l(v) \geq \inf \Phi(v) \equiv \Phi_0 > -\infty, \tag{1.1.13}$$

since l is bounded.

Now we suppose that the value Φ_0 can be reached and let u be the solution of (1.1.4), for conditions providing the existence of the solution we refer to [Mikhlin, 1970]. Taking $\varphi(t) \equiv \Phi(u+tv)$, we have

$$\varphi(t) \geq \varphi(0),$$

$$\varphi'\big|_{t=0} = \left.\frac{d\Phi}{dv}\right|_u = 0.$$

This leads to the relations:

$$(u,v)_V = l(v),$$

$$(u_{x_1}, v_{x_1})_{0,\Omega} + (u_{x_2}, v_{x_2})_{0,\Omega} = (f,v)_{0,\Omega} \quad \forall v \in V.$$

If we suppose additionally (cf. the Green formula) that

$$(u_{x_i}, v_{x_i})_{0,\Omega} = -(u_{x_i x_i}, v)_{0,\Omega}, \quad i = 1, 2,$$

then for such u we immediately obtain (1.1.1). The equation (1.1.1) is called the *Euler equation* for the variational problem (1.1.4).

The understanding of the importance of such variational problems as mathematical models of physical phenomena is a great scientific achievement. Most mathematicians of the XIX century (Dirichlet, Riemann) had an opinion that the problems of the type (1.1.4) are always solvable. But Weierstrass' criticism showed that it is not true, in general. Later (in 1901), Hilbert included "a proof of validity of variational principles" into the famous list of the most important problems in mathematics. Moreover, the works of Rayleigh and Ritz had pointed out the importance of these principles for the construction of numerical methods (see Chapter 2).

The difficulty of the problem posed by Hilbert can be illustrated by the following example: If $\Phi(u^n) \to \Phi_0$ as $n \to \infty$, $u^n \in V$, then the sequence $\{u^n\}$ is fundamental only in a weak metric of V.

Indeed, let us have $\Phi(u^n) < \Phi_0 + \varepsilon/4$ for some fixed $\varepsilon > 0$, $n > N(\varepsilon)$. Then, for such n and $p \geq 1$, by (1.1.10), we get

$$\|u^{n+p} - u^n\|_L^2 = 2[\Phi(u^{n+p}) + \Phi(u^n)] - 4\Phi((u^{n+p} + u^n)/2) < \varepsilon,$$

i.e., $\{u^n\}$ is fundamental in V.

It means that if $\{u^n\}$ is convergent, then its limit would be the unique solution of (1.1.4). But $\{u^n\}$ is fundamental in a weak metric of V only; hence we cannot say that $\{u^n\}$ converges to $u \in V$.

1.2. Variational problems in Hilbert spaces

From now on we will usually omit subindices for all norms when it does not lead to misunderstanding.

Definition 1.2.1. *A linear space V with a scalar product that is complete in the norm induced by this product (i.e., when it is a Banach space in this norm) is called Hilbert space.*

Hilbert spaces will be usually denoted by the symbol H, sometimes with subindices.

Remark 1.2.2. All Hilbert spaces considered in the dissertation are separable, i.e., they contain everywhere dense countable subsets.

Theorem 1.2.3. *(Riesz theorem) Let H be a Hilbert space with the scalar product denoted by $(\cdot\,,\cdot)$. Then for any continuous linear functional l defined on H there exists exactly one element $h \in H$ such that*

$$l(v) = (h, v) \quad \forall v \in H \tag{1.2.1}$$

and $\|l\| = \|h\|$.

Proof. See [Taylor, p. 245]. □

Theorem 1.2.4. *Let H be a Hilbert space and the functional $l \in H^*$, where H^* is dual to H. Then the problem in the form of (1.1.4) with $\Phi(v) \equiv \|v\|^2 - 2l(v)$ and $V = H$ has a unique solution u.*

Proof. In view of the Riesz Theorem, any $l \in H^*$ can be uniquely represented as $l(v) = (h, v)$, and $\|h\| = \|l\|$. Then $\Phi(v) = \|v - h\|^2 - \|h\|^2$ and $u = h$. □

In the same manner the following two theorems can be proved.

Theorem 1.2.5. *Let the conditions of Theorem 1.2.4 be satisfied and let u' be a solution of (1.1.4) with l' instead of l. Then*

$$\|u - u'\| \leq \|l - l'\|.$$

Theorem 1.2.6. *Let H be a Hilbert space. Suppose that the given bilinear form $b(\cdot\,,\cdot)$, defined on $H \times H$, satisfies the following conditions:*

$$b(u,v) = b(v,u) \qquad \forall u,v \in H \quad \textit{(symmetry)}, \tag{1.2.2}$$

$$|b(u,v)| \leq C_1\|u\|\,\|v\| \quad \forall u,v \in H \quad \textit{(boundedness)}, \tag{1.2.3}$$

$$b(u,u) \geq C_0\|u\|^2, \qquad C_0 > 0, \quad \forall u \in H \quad \textit{(coercivity or V-ellipticity)}. \tag{1.2.4}$$

Then the problem (1.1.4) with $V = H$, $l \in H^*$ and

$$\Phi(v) \equiv b(v,v) - 2l(v) \tag{1.2.5}$$

is correct and equivalent to the problem of finding $u \in H$ such that

$$b(u,v) = l(v) \quad \forall v \in H. \tag{1.2.6}$$

Remark 1.2.7. It is easy to get the generalization of Theorem 1.2.6 if we replace H by $\varphi + H'$, where H' is some subset of H. This construction is useful for the inhomogeneous Dirichlet boundary value problems (see, e.g., [Axelsson, Barker]).

Also, some generalizations of the problems of minimization of $\Phi(v)$ on a non-empty, convex and closed set from V are known (see [Glowinski, Lions, Trémolières]).

1.3. On the concept of completion

In view of Section 1.2 we may say that the main difficulty in proving the validity of variational principles laid in making the extension of the incomplete space V to the Hilbert space H. The most convenient extension of spaces of type V was proposed by Sobolev in the 1930's. This approach (for details, see, e.g., [Mikhlin, 1970] and [Rektorys]) is very important for the modern theory of differential equations and numerical methods. The elements of such an extension are some classes of almost everywhere equivalent functions, and it is possible to identify them with usual functions which are called localizations of the classes. In the next section we present main features of such an extension; several important definitions and theorems will be given.

1.4. Sobolev spaces: definitions, main properties

In this section we introduce briefly several well-known definitions and results which we shall often use throughout the dissertation. The domain Ω is always supposed to be a bounded connected open set in $\mathbf{R}^d$. The conditions for the boundary will be given later.

Let $\mathbf{R}^d$ stand for the d-dimensional Euclidean space equipped with the norm

$$\|x\| = \left(\sum_{j=1}^{d} x_j^2 \right)^{1/2}, \quad x = (x_1, ..., x_d)^T \in \mathbf{R}^d. \tag{1.4.1}$$

The *Lebesgue space* of real functions defined over an open set $\Omega \subset \mathbf{R}^d$, which are integrable in the Lebesgue sense with the power $p \in [1, \infty)$, is denoted by $L^p(\Omega)$ and equipped with the norm

$$\|v\|_{0,p,\Omega} = \left(\int_\Omega |v|^p \, dx \right)^{1/p}, \quad v \in L^p(\Omega). \tag{1.4.2}$$

When $p = 2$, we write shortly $\| \cdot \|_{0,\Omega} \equiv \| \cdot \|_{0,2,\Omega}$ and we can define the scalar product as follows

$$(v, w)_{0,\Omega} \equiv \int_\Omega vw \, dx.$$

Recall the well-known *Hölder inequality*

$$\left| \int_\Omega vw \, dx \right| \leq \|v\|_{0,p,\Omega} \|w\|_{0,q,\Omega} \quad v \in L^p(\Omega), \; w \in L^q(\Omega), \tag{1.4.3}$$

which holds for any $p, q \in (1, \infty)$ satisfying the equality

$$\frac{1}{p} + \frac{1}{q} = 1.$$

The Lebesgue space of measurable essentially bounded functions over Ω is denoted by $L^\infty(\Omega)$ and equipped with the norm

$$\|v\|_{0,\infty,\Omega} = \operatorname*{ess\,sup}_{x \in \Omega} |v(x)|.$$

For any $p, q \in [1, \infty]$, $p \leq q$, the following algebraic imbedding

$$L^q(\Omega) \subset L^p(\Omega) \tag{1.4.4}$$

holds. Moreover, we have also the topological imbedding; namely, there exists a constant $C = C(\Omega) > 0$ such that

$$\|v\|_{0,p,\Omega} \leq C \|v\|_{0,q,\Omega} \quad \forall v \in L^q(\Omega), \tag{1.4.5}$$

see, e.g., [Vulih, Section 9.3].

By $\overline{\Omega}$ we denote the closure of Ω and by $\partial\Omega$ the boundary of Ω. Then clearly

$$\overline{\Omega} = \Omega \cup \partial\Omega, \quad \partial\Omega = \overline{\Omega} \cap (\mathbf{R}^d \setminus \Omega).$$

Recall that a *domain* Ω is an open and connected set in $\mathbf{R}^d$. Throughout the dissertation, the symbol d is solely reserved for the dimension of Ω.

If Ω is a bounded domain, then the space of continuous functions over $\overline{\Omega}$ is denoted by $C(\overline{\Omega})$ and equipped with the norm

$$\|v\|_{C(\overline{\Omega})} = \max_{x \in \overline{\Omega}} |v(x)|. \tag{1.4.6}$$

Obviously, we have

$$\|v\|_{C(\overline{\Omega})} = \|v\|_{0,\infty,\Omega} \quad \forall v \in C(\overline{\Omega}).$$

The symbol $C^k(\overline{\Omega})$, $k \in \{0, 1, ...\}$, stands for the space of functions whose classical derivatives up to order k belong to $C(\overline{\Omega})$. Moreover, we set

$$C^\infty(\overline{\Omega}) = \bigcap_{k=1}^{\infty} C^k(\overline{\Omega}),$$

i.e., $C^\infty(\overline{\Omega})$ is the space of infinitely differentiable functions over $\overline{\Omega}$. Finally, by $C_0^\infty(\Omega)$ we denote the space of infinitely differentiable functions with a compact support in Ω, i.e.,

$$C_0^\infty(\Omega) = \left\{ v \in C^\infty(\overline{\Omega}) \mid \operatorname{supp} v \subset \Omega \right\},$$

where

$$\operatorname{supp} v = \overline{\{x \in \Omega \mid v(x) \neq 0\}}. \tag{1.4.7}$$

This space plays a very important role in the theory of distributions.

We shall consider Sobolev spaces defined only on bounded domains with a Lipschitz continuous boundary which is a sufficiently wide class of domains for practical purposes. The next definition is due to [Mazja, p. 18].

Definition 1.4.1. *A bounded set $\Omega \subset \mathbf{R}^d$ is said to have a Lipschitz continuous boundary if for any $z \in \partial\Omega$ there exists a neighbourhood $U = U(z)$ such that the set $U \cap \Omega$ can be expressed, in some Cartesian coordinate system $(x_1, ..., x_d)$, by the inequality $x_d < F(x_1, ..., x_{d-1})$, where F is a Lipschitz continuous function.*

Denote by $\mathcal{L}$ the set of all bounded domains with a Lipschitz continuous boundary.

Actually, the assumption $\Omega \in \mathcal{L}$ is not too restrictive for applications on bounded domains. Moreover, if $\Omega \in \mathcal{L}$, then by [Nečas, 1967, p. 88] the outward normal exists almost everywhere (a.e.) on $\partial\Omega$. This is the reason why we will confine ourselves to domains with a Lipschitz continuous boundary, as in practice people usually deal with, e.g., the normal derivative, the normal component of flux, etc. A further reason is that there are several definitions of the Sobolev spaces and in the case $\Omega \notin \mathcal{L}$ the Sobolev spaces need not be uniquely defined.

So from now on let $\Omega \in \mathcal{L}$. For any $v \in C^\infty(\overline{\Omega})$ and the so-called *multi-index* $m = (m_1, ..., m_d)$, we define the *m-th classical derivative* as follows

$$D^m v \equiv \frac{\partial^{|m|} v}{\partial x_1^{m_1} \cdots \partial x_d^{m_d}},$$

where $m_1, ..., m_d$ are non-negative integers and $|m| \equiv m_1 + ... + m_d$. We define

$$D^{(0,\dots,0)} v \equiv v,$$

and

$$m! \equiv (m_1!) \cdot ... \cdot (m_d!).$$

A function $v \in L^2(\Omega)$ is said to have the m-th generalized derivative in $L^2(\Omega)$ if there exists a function $z \in L^2(\Omega)$ such that

$$\int_\Omega zw\, dx = (-1)^{|m|} \int_\Omega v D^m w\, dx \quad \forall w \in C_0^\infty(\Omega).$$

Then the function z is called the *m-th generalized derivative* of v. It is well-known that z is well-defined and the classical derivative is also generalized one. Hence, we will use the same symbol $D^m v$ for the generalized derivative.

Since for smooth functions $w(x)$ the derivative $D^m w$ does not depend on the order of differentiation, it follows from the uniqueness of the generalized derivative that the generalized derivative does not depend on the order of differentiation.

It is easy to check also that

$$D^m(c_1 u + c_2 v) = c_1 D^m u + c_2 D^m v,$$

where c_1, c_2 are constants, i.e., D^m is a linear operator.

In contrast to the corresponding classical derivative, the generalized derivative $D^m v$ is defined by the definition globally, at once in all of Ω. However, in every subregion $\Omega' \subset \Omega$ also the function $D^m v$ will be the generalized derivative of v.

Suppose that the function $v \in L^2(\Omega)$ has a generalized derivative $D^{m'} v = F$ and the function $F(x)$ has a generalized derivative $D^{m''} F = G$. Then there exists a generalized derivative $D^{m'+m''} v$, which is equal to G.

Besides, the generalized derivative $D^m v$ is defined at once for the order $|m|$ without assuming the existence of corresponding derivatives of lower orders.

The Sobolev space $W_p^k(\Omega)$ is defined as

$$W_p^k(\Omega) \equiv \{ v \in L^p(\Omega) \mid D^m v \in L^p(\Omega),\ |m| \le k \}.$$

This space is a Banach space with the norm

$$\|v\|_{k,p,\Omega} \equiv \|v\|_{W_p^k(\Omega)} \equiv \left(\sum_{|m|\le k} \int_\Omega |D^m v|^p \, dx \right)^{1/p} \tag{1.4.8}$$

and the seminorm ($k \ge 1$)

$$|v|_{k,p,\Omega} \equiv \left(\sum_{|m|=k} \int_\Omega |D^m v|^p \, dx \right)^{1/p} . \tag{1.4.9}$$

It is a closure of $C^\infty(\overline{\Omega})$ in the norm defined by (1.4.8).

Note that $W_p^0(\Omega) = L^p(\Omega)$ and $W_2^k(\Omega) \equiv H^k(\Omega)$, are Hilbert spaces with the following scalar product

$$(v, w)_{k,\Omega} \equiv \sum_{|m|\le k} \int_\Omega D^m v D^m w \, dx, \quad v, w \in H^k(\Omega). \tag{1.4.10}$$

Let us further introduce the induced norm in these spaces:

$$\|v\|_{k,\Omega} \equiv \left(\sum_{|m|\le k} \int_\Omega |D^m v|^2 \, dx \right)^{1/2}, \quad v \in H^k(\Omega) \tag{1.4.11}$$

and the seminorm ($k \ge 1$)

$$|v|_{k,\Omega} \equiv \left(\sum_{|m|=k} \int_\Omega |D^m v|^2 \, dx \right)^{1/2}, \quad v \in H^k(\Omega). \tag{1.4.12}$$

Note that the norm $\|\cdot\|_{k,\Omega}$ and the seminorm $|\cdot|_{k,\Omega}$ are invariants under substitution of some cartesian coordinate system by another one (see [Sobolev, 1974]).

Theorem 1.4.2. *(On equivalent norms). Let $l_1, ..., l_q$ be functionals from $(W_p^k(\Omega))^*$. Suppose that for any polynomial Q with degree at most $k-1$ the equalities $l_i(Q) = 0$, $i = 1, ..., q$ imply $Q = 0$. Then the norm defined by (1.4.8) is equivalent to the norm*

$$|v|'_{k,p} \equiv \left(\sum_{|m|=k} \int_\Omega |D^m v|^p \right)^{1/p} + \sum_{i=1}^{q} |l_i(v)|. \tag{1.4.13}$$

For the proof see [Sobolev, 1963].

Also for certain subspaces a norm can be of a more simple form. For example, the norm $\|\cdot\|_{m,\Omega}$ in $W_2^m(\Omega, \Gamma_0)$ is equivalent to $|\cdot|_{m,\Omega}$, where $W_2^m(\Omega;\Gamma_0)$ is a closure of the set $\{v \in C^\infty(\overline{\Omega}) \mid v = 0 \text{ on } \Gamma_0\}$ in the norm (1.4.8). Here $\Gamma_0 = \overline{\Gamma}_0 \subset \partial\Omega$ with $\text{meas}_{d-1}\,\Gamma_0 > 0$.

Further, $W_2^m(\Omega, \emptyset) = H^m(\Omega)$ and $W_2^m(\Omega, \partial\Omega) \equiv \overset{\circ}{W}{}_2^m(\Omega) \equiv \overset{\circ}{H}{}^m(\Omega)$.

For simplicity, the same symbols will be used also for vector functions; we shall write, e.g.,

$$\|v\|_{k,\Omega} = \left(\sum_{l=1}^{q} \|v_l\|_{k,\Omega}^2 \right)^{1/2} \quad \text{for } v = (v_1, ..., v_q)^T \in (H^k(\Omega))^q.$$

Moreover, the subscript $_\Omega$ will be often omitted, i.e.,

$$(\cdot\,,\cdot)_k = (\cdot\,,\cdot)_{k,\Omega}, \quad \|\cdot\|_k = \|\cdot\|_{k,\Omega}, \quad |\cdot|_k = |\cdot|_{k,\Omega}.$$

In particular, we have

$$\|v\|_1^2 = \|v\|_0^2 + \|\operatorname{grad} v\|_0^2.$$

Clearly,

$$\|v\|_{k-1} \leq \|v\|_k \quad \forall v \in H^k(\Omega),\ k = 1, 2, ...$$

and

$$L^2(\Omega) = H^0(\Omega) \supset H^1(\Omega) \supset H^2(\Omega) \supset \cdots .$$

Note that each classical derivative is also the generalized derivative, and, thus, we have

$$C^k(\overline{\Omega}) \subset H^k(\Omega), \quad k = 0, 1, \ldots .$$

Definition 1.4.3. *A set $\Gamma \subset \partial\Omega$ is said to be a (relatively) open set in $\partial\Omega$ if for any $x \in \Gamma$ there exists an open ball $B \subset \mathbf{R}^d$ containing x such that $B \cap \partial\Omega \subset \Gamma$.*

The *Lebesgue space* of square integrable functions over an open set $\Gamma \subset \partial\Omega$ is denoted by $L^2(\Gamma)$ and equipped with the standard norm

$$\|v\|_{0,\Gamma} = \left(\int_\Gamma |v|^2\, ds \right)^{1/2}, \quad v \in L^2(\Gamma).$$

Further, we recall some important properties of Sobolev spaces, proofs of which can be found, e.g., in [Adams], [Kufner, John, Fučík], [Mazja], [Nečas, 1967]. The proofs of the Eberlein–Schmulian theorem and the Brouwer fixed-point theorem can be found in [Yosida] and [Fučík, Kufner], respectively.

Theorem 1.4.4. *(Sobolev imbedding theorem). Let $\Omega \in \mathcal{L}$ and let k, p be integers such that $kp > d$. Then*

$$W_p^k(\Omega) \subset C(\overline{\Omega}) \tag{1.4.14}$$

and there exists a constant $C > 0$ such that

$$\|v\|_{k,p} \geq C\|v\|_{C(\overline{\Omega})} \quad \forall v \in W_p^k(\Omega). \tag{1.4.15}$$

Theorem 1.4.5. *(Trace theorem). Let $\Omega \in \mathcal{L}$. Then there exists exactly one linear continuous operator $\gamma\colon H^1(\Omega) \to L^p(\partial\Omega)$ such that*

$$\gamma(v) \equiv v\big|_{\partial\Omega} \quad \forall v \in C^\infty(\overline{\Omega}),$$

for any dimension $d \in \{1, 2, ...\}$ and $p = 2$ or $d = 2$ and $p \in [1, \infty)$ or $d = 3$ and $p \in [1, 4]$.

The function γv for $v \in H^1(\Omega)$ is called the *trace of* v and we denote it, for simplicity, by $v\big|_{\partial\Omega}$. Theorem 1.4.5, in fact, says that there exists a constant $C_p > 0$ such that

$$\|v\|_{0,p,\partial\Omega} \le C_p \|v\|_{1,\Omega} \quad \forall v \in H^1(\Omega). \tag{1.4.16}$$

The trace theorem enables us to define the space

$$H_0^1(\Omega) = \left\{ v \in H^1(\Omega) \mid v = 0 \text{ on } \partial\Omega \right\}.$$

Note that the spaces $H_0^1(\Omega)$ and $H^1(\Omega)$ can also be defined as the completion of $C_0^\infty(\Omega)$ and $C^\infty(\overline{\Omega})$, respectively, under the norm $\|\cdot\|_{1,\Omega}$, i.e.,

$$H_0^1(\Omega) = \overline{C_0^\infty(\Omega)}, \quad H^1(\Omega) = \overline{C^\infty(\overline{\Omega})}.$$

The following two theorems present stronger variants of the Sobolev Imbedding Theorem and the Trace Theorem.

Theorem 1.4.6. *(Rellich theorem). Let $\Omega \in \mathcal{L}$. Then the identity mapping from $H^1(\Omega)$ to $L^2(\Omega)$ is compact (i.e., any bounded sequence in $H^1(\Omega)$ contains a subsequence converging in $L^2(\Omega)$).*

Theorem 1.4.7. *(Kondrašov theorem). Let $\Omega \in \mathcal{L}$. Then the identity mapping from $H^1(\Omega)$ to $L^p(\Omega)$ is compact for $p \in [1, \infty)$ if $d = 2$ and $p \in [1, 6)$ if $d = 3$.*

Moreover, we have the following inclusion

$$H^1(\Omega) \hookrightarrow L^p(\Omega) \quad \text{for} \begin{cases} p \in [1, \infty) \text{ and } d = 2, \\ p \in [1, 6] \text{ and } d = 3, \end{cases} \tag{1.4.17}$$

and there exists a constant $C_p > 0$ dependent on p such that

$$\|v\|_{0,p,\Omega} \le C_p \|v\|_{1,\Omega} \quad \forall v \in H^1(\Omega).$$

Note that for $p = 6$ and $d = 3$ the above imbedding is not compact.

Theorem 1.4.8. *(Eberlein–Schmulian theorem). V is a reflexive Banach space if and only if for any bounded sequence from V there exists a weakly convergent (in V) subsequence.*

Note that any Hilbert space (even nonseparable) is reflexive.

Theorem 1.4.9. *(Green theorem). Let $\Omega \in \mathcal{L}$. Then for each $i \in \{1, ..., d\}$*

$$\int_\Omega w\partial_i v \, dx + \int_\Omega v\partial_i w \, dx = \int_{\partial\Omega} n_i v w \, ds \quad \forall v, w \in H^1(\Omega), \tag{1.4.18}$$

where n_i are the components of the outward unit normal to $\partial\Omega$ and $\partial_i v = \dfrac{\partial v}{\partial x_i}$.

Theorem 1.4.10. *Let $\Omega \in \mathcal{L}$ and let $\omega \neq \emptyset$ be an open set either in Ω or in $\partial\Omega_0$, where $\Omega_0 \subseteq \Omega$, $\Omega_0 \in \mathcal{L}$. Then there exists a constant $C > 0$ such that*

$$\|v\|_1 \leq C(|v|_1^2 + \|v\|_{0,\omega}^2)^{1/2} \quad \forall v \in H^1(\Omega). \tag{1.4.19}$$

This inequality is usually called *Friedrichs' inequality* when $\omega \subset \partial\Omega$. In particular,

$$\|v\|_1 \leq C|v|_1 \quad \forall v \in H_0^1(\Omega). \tag{1.4.20}$$

Inequalities (1.4.19) and (1.4.20) are usually used to prove the inequality (1.2.4), the so-called the *V-ellipticity condition*

$$a(v, v) \geq C\|v\|_1^2 \quad \forall v \in V,$$

when dealing with the solvability of a weak formulation of a linear elliptic problem of the second order over an appropriate subspace $V \subset H^1(\Omega)$. If the classical formulation of the elliptic problem contains first order terms, then the V-ellipticity need not be valid. In this case, the *Garding inequality*

$$a(v, v) \geq C_1\|v\|_1^2 + C_2\|v\|_0^2 \quad \forall v \in V \tag{1.4.21}$$

can be applied, where $C_1 > 0$, but the constant C_2 may be negative (for more details, see [Axelsson, Barker]).

Theorem 1.4.11. *(Brouwer fixed-point theorem). Let $\mathcal{B} \subset X$ be a non-empty closed ball in a normed finite dimensional space X. Let S be a continuous mapping which maps the ball $\mathcal{B}$ into itself, i.e., $S(\mathcal{B}) \subset \mathcal{B}$. Then there exists $x \in S$ such that $Sx = x$.*

1.5. Weak solutions of boundary value problems

The solutions of problems similar to that one considered in Theorem 1.2.6 are called *weak solutions* of problems similar to (1.1.1)–(1.1.2).

Note that the formulation (1.2.6) is very convenient for the description of the original problem and also for using numerical methods (cf. Chapter 2).

Sometimes it is convenient to rewrite the formulation similar to (1.2.6) as an operator equation. The most important result is the following two theorems.

Theorem 1.5.1. *Let H be a Hilbert (or Euclidean) space and let the bilinear form b satisfy (1.2.3). If $l(v) = (h, v)$, then the problem (1.2.6) is equivalent to the linear operator equation*

$$Lu = h, \tag{1.5.1}$$

where $\|L\| \le C_1$ and $(Lu, v) = b(u, v)$ for all $u, v \in H$. If, moreover, b satisfies (1.2.2), (1.2.4), then $L = L^*$ and $C_0 I \le L \le C_1 I$, where C_0 and C_1 are constants from (1.2.4) and (1.2.3).

Theorem 1.5.2. *(Lax–Milgram lemma). Let H be a Hilbert (or Euclidean) space, let b satisfy (1.2.3), (1.2.4), $l(v) = (h, v)$. Then the problem (1.2.6) is equivalent to (1.5.1), and there exists L^{-1} such that $\|L^{-1}\| \le C_0^{-1}$.*

Now, we present shortly an example of using Theorem 1.5.2:

$$\begin{aligned} &\Omega \in \mathcal{L}, \quad \Gamma_0 \subset \partial\Omega, \quad \mathrm{meas}_{d-1}\Gamma_0 > 0, \\ &\Gamma_1 \equiv \partial\Omega \setminus \Gamma_0, \quad H = W_2^1(\Omega, \Gamma_0), \\ &\|u\| \equiv |u|_{1,\Omega}, \\ &b(u, v) \equiv b^{(0)}(\Omega; u, v) + (\sigma, uv)_{0,\Gamma_1}, \end{aligned} \tag{1.5.2}$$

where

$$b^{(0)}(\Omega; u, v) \equiv \sum_{r,l=1}^{d} (a_{rl}, D_l u D_r v)_{0,\Omega} + \sum_{r=1}^{d} (b_r v D_r u - b'_r u D_r v, 1)_{0,\Omega} + (c, uv)_{0,\Omega} \tag{1.5.3}$$

and the coefficients a_{rl}, b_r, b'_r, c, σ are supposed to be piecewise smooth, i.e., all integrals in (1.5.2), (1.5.3) have sense when $u \in H$, $v \in H$.

Lemma 1.5.3. *The form $b(\cdot\,,\cdot)$ from (1.5.2) satisfies (1.2.3).*

Proof. It is enough to show that the modulus of each term from (1.5.2) and (1.5.3) is less or equal to $C\|u\|\,\|v\|$, where C is some constant.

For example,

$$\begin{aligned} |(b_r v, D_r u)_{0,\Omega}| &\le \sup |b_r(x)| (|v|, |D_r u|)_{0,\Omega} \\ &\le \sup |b_r(x)| \cdot |v|_{0,\Omega} |D_r u|_{1,\Omega} \le C\|u\|\,\|v\| \end{aligned}$$

(see (1.1.11)).

For the term with σ we have to use some variant of the Trace Theorem. □

Let

$$\begin{aligned} |b| &\equiv \sup_{\substack{x\in\Omega \\ r\in\{1,\ldots,d\}}} |b_r(x) - b'_r(x)|, \\ |\bar{c}| &\equiv \sup_{\substack{x\in\Omega \\ c(x)<0}} |c(x)|, \end{aligned}$$

ν_0 be some constant, I be a unit matrix $(d \times d)$.

Lemma 1.5.4. *Let for all $x \in \Omega$ the matrix*

$$A \equiv \{(a_{rl} + a_{lr})/2\}_{l,r=1}^{d} \geq \nu_0 I, \quad \nu_0 > 0,\ \sigma \geq 0,\ |b|Cd^{1/2} + |\bar{c}|C^2 \equiv \nu' < \nu_0, \tag{1.5.4}$$

where C is a constant from the Poincaré-Steklov inequality (1.1.11).
Then for $b(\cdot\,,\cdot)$ from (1.5.2) the property (1.2.4) is valid with $C_0 = \nu_0 - \nu'$.

Proof. As for the previous lemma we have

$$\begin{aligned} b(u,u) &\geq \nu_0 \sum_{r=1}^{d} |D_r u|_{0,\Omega}^2 - |b| \sum_{r=1}^{d} (|D_r u|, |u|)_{0,\Omega} - |\bar{c}|\, |u|_{0,\Omega}^2 \\ &\geq \nu_0 \|u\|^2 - d^{1/2}|b|\, |u|_{1,\Omega}|u|_{0,\Omega} - |\bar{c}|\, |u|_{0,\Omega}^2 \\ &\geq (\nu_0 - \nu')\|u\|^2. \end{aligned}$$

□

In view of Lemmas 1.5.3 and 1.5.4 and Theorem 1.5.2, the problem (1.2.6), (1.5.2) is correct. This problem corresponds to the mixed boundary value problem with the Dirichlet condition on Γ_0 and the boundary condition of the third type on Γ_1 for a common elliptic equation of the second order.

Remark 1.5.5. When $\Gamma_0 = \Gamma$, the condition on ν_0 can be replaced by the condition $D_r(b_r - b'_r) \leq 0$ for all r.

Remark 1.5.6. Conditions on the coefficient may be replaced by more common ones (see, e.g., [Ladyzenskaja, Ural'ceva]).

Remark 1.5.7. More examples on the point can be found, e.g., in [Axelsson, Barker]).

Chapter 2
Finite element method: main features

In this chapter we briefly describe the main idea of the Rayleigh–Ritz, the Bubnov–Galerkin and the finite element methods.

Several special questions of FE analysis, being related in a natural way with the next chapters, are outlined as well.

2.1. Rayleigh–Ritz method

Consider the problem from Theorem 1.2.6. The main idea of the Rayleigh–Ritz method is to replace (1.1.4) by the finite-dimensional problems

$$u^{\wedge} = \arg\min \Phi(v^{\wedge}), \quad v^{\wedge} \in V^{\wedge}, \tag{2.1.1}$$

where $V^{\wedge} \equiv V_N$ is some N-dimensional subspace in H with the basis $\{\psi_1, ..., \psi_N\}$ and expansion

$$v^{\wedge} = \alpha_1\psi_1 + ... + \alpha_N\psi_N, \quad v^{\wedge} \in V^{\wedge}. \tag{2.1.2}$$

The functions ψ_i did not depend on N and $V_N \subset V_{N+1}$ for the first variants of the method. For modern methods it is not true: the subspaces V_N are defined by some discretization parameter $h > 0$ instead of N. Such parameter presents a certain characteristic of the used mesh. Sometimes h can be even a vector parameter. Hence, we should regard $N \equiv N_h$ as elements of a certain increasing sequence of natural numbers. Moreover, we will usually use the denotation V_h instead of V_N to point out the dependence of finite element spaces on h (we will also use v_h instead of $v^{\wedge}$).

The problem (2.1.1) is closely related to the least square method (see, e.g., [Mikhlin, 1965]).

Theorem 2.1.1. *Let the conditions of Theorem 1.2.4 be satisfied. Then the Rayleigh–Riesz method (2.1.1) leads to the sequence of correct problems and its solutions u_N are the closest to u elements from V_N.*

Theorem 2.1.2. *Let the conditions of Theorem 1.2.6 be satisfied and $C = C_1^{1/2}/C_0^{1/2}$. Then (2.1.1) leads to correct problems, for solutions of which the following estimate holds:*

$$\|u_N - u\| \le C\,\mathrm{dist}\{u, V_N\}. \tag{2.1.3}$$

For the proof, see, e.g., [Axelsson, Barker, pp. 151–152].

Note that the method (2.1.1) can be reformulated in the following form (cf. (1.2.6)): find $u_N \in V_N$ such that

$$b(u_N, v_N) = l(v_N) \quad \forall v_N \in V_N. \tag{2.1.4}$$

When $L_N \equiv P_N L P_N$ and $f_N \equiv P_N f$, we can rewrite the method as the operator equation in V_N (cf. (1.5.1)):

$$L_N u_N = f_N. \tag{2.1.5}$$

The most important point is an algebraic form of (2.1.4), which uses the expansion (2.1.2) and the vectors

$$\mathbf{v} \equiv (\alpha_1, ..., \alpha_N)^T \in \mathbf{R}^N. \tag{2.1.6}$$

Then the vector $\mathbf{u}$ which corresponds to the expansion of u_N from (2.1.4) presents a solution of the following system of algebraic equations:

$$L\mathbf{u} = \mathbf{f}, \tag{2.1.7}$$

$$L \equiv L_h \equiv (b(\psi_j, \psi_i))_{i,j=1}^N, \tag{2.1.8}$$

where L is the Gramm matrix of the functions $\psi_1, ..., \psi_N$ and

$$\mathbf{f} \equiv (l(\psi_1), ..., l(\psi_N))^T. \tag{2.1.9}$$

2.2. Bubnov–Galerkin method

Formulation (2.1.4) for (2.1.1) enables us to introduce a more general method, which solves problems even when the symmetry condition (1.2.2) is violated.

Theorem 2.2.1. *(Céa's Lemma) Let the condition of Theorem 1.5.2 be satisfied. Consider problem (2.1.4). This problem is correct and for its solutions u_N the estimate (2.1.3) holds with $C = C_1/C_0$.*

For the proof, see, e.g., [Axelsson, Barker, Theorem 4.3].

We say that $\{V_N\}$ *approximates* H as $N \to \infty$ if

$$\operatorname{dist}\{u, V_N\} \underset{N\to\infty}{\longrightarrow} 0 \quad \forall u \in H. \tag{2.2.1}$$

An algebraic form for the Bubnov–Galerkin method is the same, see (2.1.7)–(2.1.9), the only difference is that the matrix L need not be symmetric.

2.3. Finite element method

The finite element method (FEM) can be considered as a special case of the Bubnov–Galerkin method, for which supports of almost all basic functions are just several cells of the mesh. Having domains with curved boundaries we have to take into account also an approximation of the boundary.

In the next sections we outline several topics from FE analysis which are related to the main questions of the dissertation.

2.4. On triangulation of domain

Definition 2.4.1. *Let Ω be some domain in $\mathbf{R}^d$. Then a finite number of subsets $\mathcal{T}_h = \{K_1, ..., K_N\}$ is called a triangulation of $\overline{\Omega}$ if the following assumptions hold:*

(1) $\overline{\Omega} = \bigcup_{K \in \mathcal{T}_h} K$,

(2) for each $K \in \mathcal{T}_h$, the set K is closed and its interior K^0 is non-empty,

(3) for each distinct $K_1, K_2 \in \mathcal{T}_h$, we have $K_1^0 \cap K_2^0 = \emptyset$,

(4) for each $K \in \mathcal{T}_h$, the boundary ∂K is Lipschitz continuous.

Sometimes triangulation will be also called decomposition, partition or division.

Mostly, we deal with elements which are convex polygons or polyhedra (then, of course, $\overline{\Omega}$ is a polygon or a polyhedron). In this case we add two more assumptions:

(5) any face of any $K_1 \in \mathcal{T}_h$ is either a subset of the boundary $\partial\Omega$, or a face of another element $K_2 \in \mathcal{T}_h$,

(6) the interior of any face of any $K \in \mathcal{T}_h$ is disjoint with $\overline{\Gamma}_0 \cap \overline{\Gamma}_1$, where Γ_0 and Γ_1 are defined as in (1.5.2).

From the assumption (5) we observe that the situation of Figure 2.4.1 is not allowed for all triangulations considered in the dissertation.

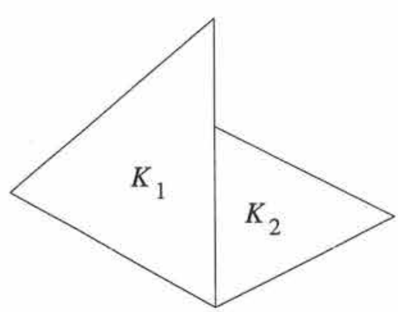

Figure 2.4.1.

The assumption (6) is called the *consistency condition* for mixed boundary conditions (see Figure 2.4.2). Note that triangulations which do not satisfy (5) are called *nonmatching.*

Further, we present two examples of triangulation of a cube.

Example 2.4.2. A cube $[0, 1]^d$ in $\mathbf{R}^d$ can be divided into $d!$ simplices $A_0A_1...A_d$

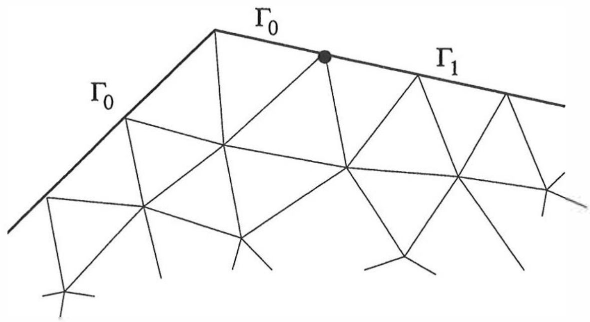

Figure 2.4.2.

with the following vertices

$$
\begin{aligned}
A_0 &\equiv (0, 0, ..., 0), \\
A_1 &\equiv A_0 + e_{j_1}, \\
A_2 &\equiv A_1 + e_{j_2}, \\
&\vdots \\
A_d &\equiv (1, 1, ..., 1),
\end{aligned}
\tag{2.4.1}
$$

where e_{j_k} is the ort in $\mathbf{R}^d$ and

$$
\begin{aligned}
&j_r \in \{1, 2, ..., d\}, \\
&j_2 \neq j_1, \\
&j_3 \neq j_1, j_2, ...
\end{aligned}
$$

Such a division is presented in Figure 2.4.3 for $d = 3$: the cube is divided into two prismas and each of them is decomposed into three tetrahedra.

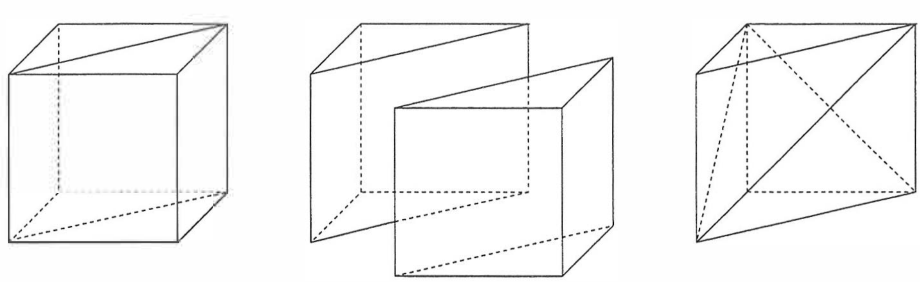

Figure 2.4.3.

Example 2.4.3. There also exists a different decomposition of the cube $[0, 1]^3$ in $\mathbf{R}^3$ into 5 tetrahedra (see Figure 2.4.4).

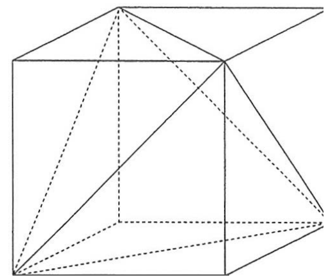

Figure 2.4.4.

A more general result on triangulation in a three-dimensional space is given by the following theorem.

Theorem 2.4.4. *For any polyhedron in* $\mathbf{R}^3$ *there exists a decomposition into tetrahedra.*

Proof. See [Křížek, Neittaanmäki, 1996, Theorem 4.21]. □

Remark 2.4.5. Computer generators of tetrahedral divisions are usually based on the Delaunay triangulation, which is uniquely determined by prescribing nodes only (cf., e.g., [George], [Weatherill, Hassan]). For a local refinement of tetrahedral meshes, see [Křížek, Strouboulis].

Remark 2.4.6. For a more general case, when Ω is not a polyhedron we use usually triangulations $\mathcal{T}_h(\Omega_h)$ of polyhedral domains Ω_h approximating in some way the given domain Ω.

Now we briefly describe the simplest way of construction of triangulations, which is suitable for polyhedral domains. The constructions of triangulations for domains with curved boundaries are presented in Part IV of the dissertation.

We omit the index h in what follows, keeping in mind that h can be chosen "equivalent" to $1/k$, i.e., there exist constants $\varkappa_1$, $\varkappa_2$, independent of h and k, such that

$$o < \varkappa_1 \leq hk \leq \varkappa_2.$$

Let Ω be a polyhedron. Having the initial triangulation, denoted by $\mathcal{T}^{(1)}(\Omega)$, we will construct new triangulations $\mathcal{T}^{(k)}(\Omega)$, $k \geq 2$, using a division of each simplex $K_i^{(1)}$ from $\mathcal{T}^{(1)}(\Omega)$ into k^d simplices $K_j^{(k)}$ forming $\mathcal{T}^{(k)}(\Omega)$. Constructions of such a type are described below for $d = 2$ and 3.

Example 2.4.7. Consider the triangle $K_i^{(1)} = A_0A_1A_2 \in \mathcal{T}^{(1)}(\Omega)$ (see Figure 2.4.5). In each side A_iA_j we put points $A_{ij,r}$, $r = 0, ..., k$, so that $A_{ij,0} = A_i$, $A_{ij,k} = A_j$ and the intermediate points divide $[A_i\ A_j]$ into k equivalent segments. Then the straight lines through all those points, parallel to the sides of $K_i^{(1)}$, define a division of $K_i^{(1)}$ into k^2 equivalent triangles forming $\mathcal{T}^{(k)}(\Omega)$ (in Figure 2.4.5 $k = 3$).

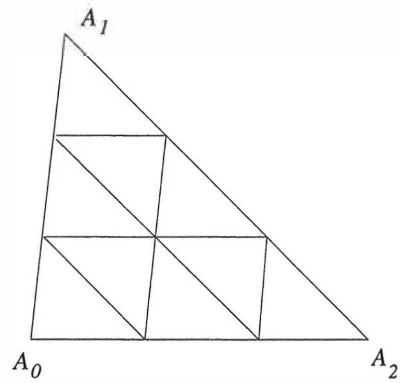

Figure 2.4.5.

Example 2.4.8. Figure 2.4.6 presents a similar decomposition of the tetrahedron $K = A_0A_1A_2A_3$ into 8 tetrahedra ($k = 2$).

Note that in three-dimensional space the situation is considerably more complicated than for $d = 2$, since it may not be possible to divide any tetrahedron into coinciding similar (congruent) tetrahedra (cf. [Goldberg]).

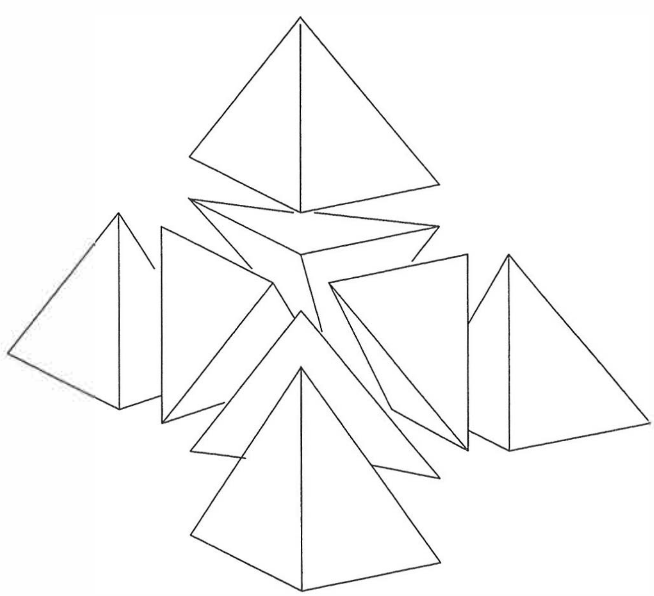

Figure 2.4.6.

2.5. On nonconforming finite element methods

When $\partial\Omega$ is piecewise curved, there are several ways of constructing finite element spaces. One way is to generate them by the so-called isoparametric (curved) elements (see, e.g., [Axelsson, Barker]). Note that a curved element K need not be convex (see Figure 2.5.1) and the space of shape functions P_K may be formed, e.g., by rational functions. Another way is to approximate Ω by a polygonal (polyhedral) domain $\Omega_h \subset \Omega$ and, then, to extend finite element functions from Ω_h to the whole Ω in an appropriate manner, see Part II for similar discussions.

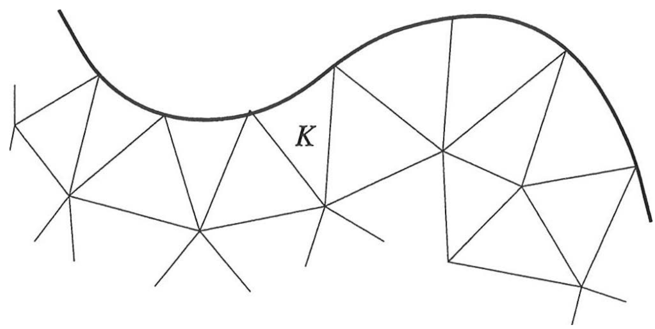

Figure 2.5.1.

When $V_h \subset H$ and when the bilinear and linear forms of the discrete problem are identical to the original ones, the finite element method is said to be conforming.

A nonconforming method arises when $V_h \not\subset H$ or when, e.g., some numerical integration is used.

Figures 2.5.2 and 2.5.3 show two different manners of an approximation of the boundary. The latter case leads to $V_h \not\subset H$ as well. The finite element method of this type is described in details, e.g., in [Feistauer, Ženíšek, 1987].

A nonconforming method is also obtained when the function giving the boundary conditions is approximated by some piecewise polynomial continuous function. More details on the point can be found in Part II of the dissertation, where all kinds of nonconformity are taken into account for three-dimensional nonlinear problem, and in the works mentioned there.

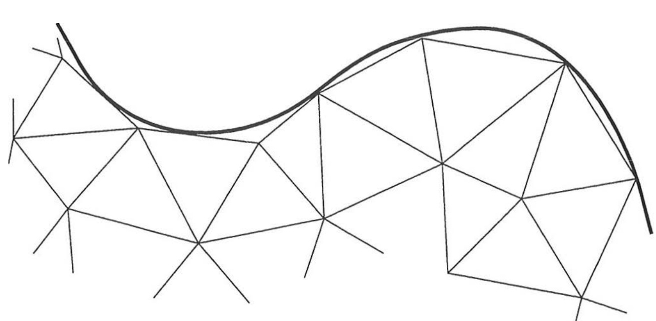

Figure 2.5.2.

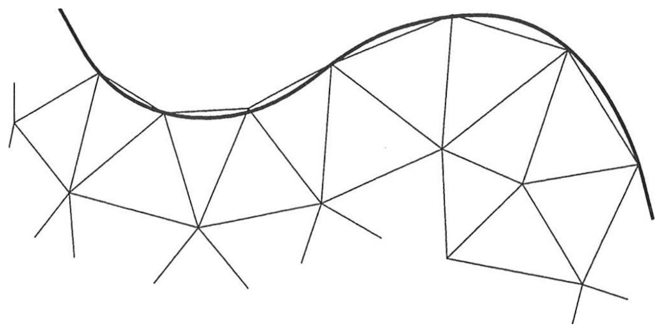

Figure 2.5.3.

2.6. Linear homotopy in plane

The standard element of triangulation in plane is either a usual triangle or a quasitriangle with one curved side (see Figure 2.6.1).

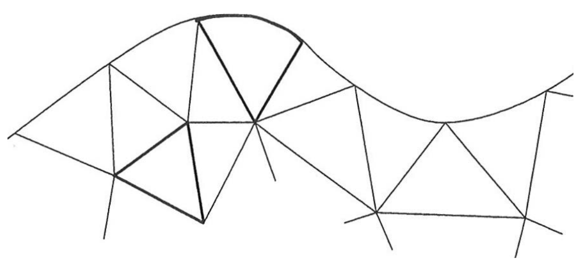

Figure 2.6.1.

A very convenient mapping of a quasitriangle $K' \equiv A_0A_1A_2$ (see Figure 2.6.2) into a straightline triangle $K \equiv \Delta A_0A_1A_2$ (with the same vertices) is a *linear homotopy* (see, e.g., [Korneev, 1977], [Mitchell, Wait], [Zlámal, 1973]). This mapping, in a local coordinate system (y_1, y_2) adapted to K', is a homothetic transformation with respect to y_2 (y_1 is fixed), such that straightline sides of K' are transformed into themselves.

Definition 2.6.1. *A quasitriangle K' with vertices A_0, A_1, A_2 is called a standard quasitriangle of the order m if (see Figure 2.6.2):*

(a) in the Decart coordinate system (y_1, y_2), where y_2 lies along the straight line (A_0A_2), $A_0 = (0,0)$, $A_1 = (\lambda, f(\lambda))$, $A_2 = (0, f(0))$ and the curve A_1A_2 is given by the equation

$$y_2 = f(y_1), \quad y_1 \in [0, \lambda], \ f \in C^{m+1}[0, \ \lambda];$$

(b) there exist two points A_2' and A_2'' on the ray $[A_0A_2)$ such that

$$\triangle A_0A_1A_2' \subset K' \subset \triangle A_0A_1A_2'';$$

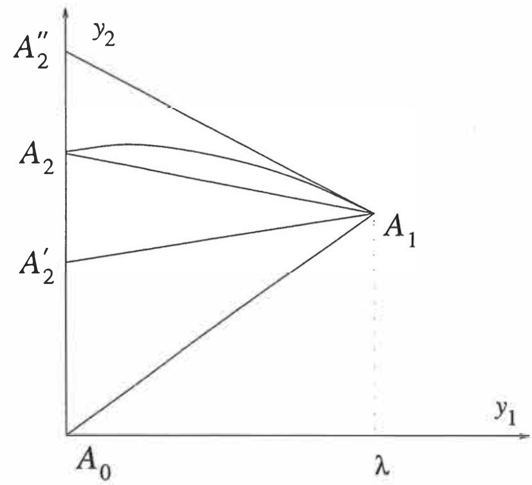

Figure 2.6.2.

(c) one of the following inequalities holds in the segment $[0, \lambda]$*:*

$$f''(y_1) \geq 0, \tag{2.6.1}$$

$$f''(y_1) \leq 0. \tag{2.6.2}$$

Let for K' (see Figure 2.6.2) the segments $[A_0A_1]$, $[A_2A_1]$, $[A_2'A_1]$ and $[A_2''A_1]$ lie on the straight lines described by the following equations:

$$y_2 = ky_1,$$
$$y_2 - f(\lambda) = k_1(y_1 - \lambda),$$
$$y_2 - f(\lambda) = k_1'(y_1 - \lambda),$$
$$y_2 - f(\lambda) = k_1''(y_1 - \lambda),$$

respectively. Then the linear homotopy $x = F(y)$, mapping K' into K, is defined by the following formulae:

$$x_1 = y_1 \equiv F_1(y_1, y_2) \tag{2.6.3}$$
$$x_2 = ky_1 + (k_1 - k)(y_1 - \lambda)[f(y_1) - ky_1]^{-1}(y_2 - ky_1) \equiv F_2(y_1, y_2).$$

The inverse mapping $y = Z(x)$ is defined by the formulae:

$$y_1 = x_1 \equiv Z_1(x_1, x_2) \tag{2.6.4}$$
$$y_2 = kx_1 + [f(x_1) - kx_1][(k_1 - k)(x_1 - \lambda)]^{-1}(x_2 - kx_1) \equiv Z_2(x_1, x_2).$$

2.7. Strongly regular families of triangulations

Let Ω be some polyhedron in $\mathbf{R}^d$. We denote by $\mathcal{T}_h$ a triangulation of $\overline{\Omega}$ formed by simplices which satisfy the requirements (1)–(6) from Section 2.4 and use the denotations $h_K = \operatorname{diam} K$, $K \in \mathcal{T}_h(\Omega)$, $h = \max\limits_{K \in \mathcal{T}_h(\Omega)} h_K$.

A set of triangulations $\mathcal{F}$ is called a *family of triangulations* if for every $\varepsilon > 0$ there exists $\mathcal{T}_h \in \mathcal{F}$ with $h < \varepsilon$.

Further, we introduce four equivalent definitions on a regularity of triangulations. The proof of equivalence is only technically complicated and can be done as in [Křížek, 1991]. The reason of introducing these definitions is that the first one is standard, see [Ciarlet], but the other ones are sometimes more convenient to work with and they all are, in fact, used in different parts of the dissertation.

Definition 2.7.1. *A family of triangulations $\mathcal{F}$ of a polyhedron $\overline{\Omega}$ into simplices is said to be strongly regular if there exists a constant $\varkappa > 0$, independent of h, such that for any triangulation $\mathcal{T}_h \in \mathcal{F}$ and for any simplex $K \in \mathcal{T}_h$ there exists a ball $\mathcal{B}_K$ of radius ϱ_K such that $\mathcal{B}_K \subset K$ and*

$$\varkappa h \le \varrho_K. \tag{2.7.1}$$

Definition 2.7.2. *A family of triangulations $\mathcal{F}$ of a polyhedron $\overline{\Omega}$ into simplices is said to be strongly regular if there exists a constant $\varkappa_1 > 0$, independent of h, such that for any triangulation $\mathcal{T}_h \in \mathcal{F}$ and any simplex $K \in \mathcal{T}_h$, we have*

$$\varkappa_1 h^d \le \operatorname{meas} K. \tag{2.7.2}$$

Definition 2.7.3. *A family of triangulations $\mathcal{F}$ of a polyhedron $\overline{\Omega}$ into simplices is said to be strongly regular if there exist constants $\varkappa_2 > 0$ and $\alpha_0 > 0$, independent of h, such that for any triangulation $\mathcal{T}_h \in \mathcal{F}$ and any simplex $K \in \mathcal{T}_h$, we have*

$$\varkappa_2 h \le l_K, \quad \alpha_0 \le \varphi_K \le \pi - \alpha_0, \tag{2.7.3}$$

where l_K is a length of any edge of K and φ_K is any angle between any edge and the hyperplane, being span of the other edges starting from the same vertex as the previous edge.

Definition 2.7.4. *A family of triangulations $\mathcal{F}$ of a polyhedron $\overline{\Omega}$ into simplices is said to be strongly regular if there exist constants $\varkappa_3 > 0$ and $\alpha_1 > 0$, independent of h, such that for any triangulation $\mathcal{T}_h \in \mathcal{F}$ and any simplex $K \in \mathcal{T}_h$, we have*

$$\varkappa_3 h \le l_K, \quad \alpha_1 \le \varphi_K, \tag{2.7.4}$$

where l_K is a length of any edge of K and φ_K is any angle between any two faces of K.

Remark 2.7.5. The strongly regularity property means, actually, two of the following facts: lengths of all edges of all simplices from $\mathcal{T}_h$ are “proportional” to h and there are no “too flat” simplices.

Such uniform meshes gives the FEM great strength, they allow to derive a priori error estimates. Nonuniform meshes are desirable when the solution of the problem is known to vary more rapidly in certain parts of the domain than in others and essential when the domain has an irregular geometry, see, e.g., [Křížek, Neittaanmäki, 1996, Section 4.4].

Remark 2.7.6. Note that a strongly regular family of triangulations of a polygon is easy to obtain due to the fact that any triangle in the triangulation is divided by midlines into four congruent triangles, which are similar to the original one. In three-dimensional space the situation is considerably more complicated, since it may not be possible to divide any tetrahedron into coinciding similar (congruent) tetrahedra.

Nevertheless, the following theorem holds.

Theorem 2.7.7. *For any polyhedron in* $\mathbf{R}^3$ *there exists a strongly regular family of decompositions into tetrahedra.*

Proof. See [Křížek, 1982]. □

Remark 2.7.8. The question of construction of strongly regular families of triangulations into triangles for planar domains with piecewise C^2-smooth boundaries is considered in Part IV.

PART II

Finite Element Analysis of Variational Crimes for Nonlinear Heat Conduction Equation in 3D

On contents of Part II

Part II of the dissertation consists of six chapters. First, in Chapter 3 we review shortly the main results obtained for nonlinear heat conduction equations, list the most important references on the point and give several common definitions and denotations used in the next chapters.

Further, the author presents his own contribution to the field, see [Korotov, Křížek]. The extended material of the above mentioned paper is given as Chapters 4–8.

The solved problem can be described as it follows. A finite element approximation of a nonlinear heat conduction equation in a three-dimensional bounded convex domain with a smooth boundary is examined. The domain is approximated by a polyhedron and a numerical integration is taken into account, i.e., so-called variational crimes are committed. We apply linear tetrahedral finite elements and prove the convergence of approximate solutions on polyhedral domains in the W_2^1-norm to the true solution without any regularity assumptions.

Chapter 3
Survey of main results

In this chapter we present a survey of main theoretical results, obtained for nonlinear heat conduction equations and introduce simultaneously definitions necessary for the next chapters.

More details and references on the topic are given in [Křížek, Neittaanmäki, 1996], [Křížek, Liu, 1997] and [Liu].

Several real-life examples of a calculation of a temperature distribution can be found, e.g., in [Křížek, Preiningerová, 1987], [Křížek, Preiningerová, 1991] and [Preiningerová, Křížek, Kahoun].

3.1. Classical formulation

A temperature distribution in large transformers is described by a quasilinear elliptic problem whose classical formulation reads:

Find $u \in C^1(\overline{\Omega})$ such that $u|_\Omega \in C^2(\Omega)$ and

$$-\operatorname{div}(\mathcal{A}(\cdot\,,u)\operatorname{grad} u) = f \quad \text{in } \Omega, \tag{3.1.1}$$

$$u = \overline{u} \quad \text{on } \Gamma_0, \tag{3.1.2}$$

$$\alpha u + n^T \mathcal{A}(\cdot\,,u)\operatorname{grad} u = g \quad \text{on } \Gamma_1, \tag{3.1.3}$$

where $\Omega \in \mathcal{L}$ (see Definition 1.4.1), $n = (n_1, ..., n_d)^T$ is the outward unit normal to $\partial\Omega$, $d \in \{1, 2, ...\}$, Γ_0 and Γ_1 are defined as in (1.5.2), u is the temperature, $f \in L^2(\partial\Omega)$ is the density of volume heat forces, $g \in L^2(\Gamma_1)$ is the density of surface heat sources, $\overline{u}$ is the temperature, maintained along the part Γ_0 of the boundary, the function $\alpha \geq 0$ is the heat transfer coefficient; the magnetic cores of transformers (consisting of iron sheets) are nonlinear orthotropic media, the heat conductivities of which are represented by a diagonal uniformly positive definite matrix $\mathcal{A} = \mathcal{A}(u)$.

Let the functions $\mathcal{A}$, α, f, $\overline{u}$ and g be sufficiently smooth for the time being (precise assumptions on these functions will be given later). The condition (3.1.3) is the so-called boundary condition of the third type.

First, we describe the main difficulties for treatment of such type of equations; they are brought by the nonlinearity of the problem and properties of the media.

Consider the following special case of (3.1.1)–(3.1.3) which describes a stationary heat conduction in a homogeneous and isotropic medium Ω:

$$\begin{aligned} -\operatorname{div}(\lambda(u)\operatorname{grad} u) &= f \quad && \text{in } \Omega, \\ u &= 0 && \text{on } \Gamma_0, \\ n^T\lambda(u)\operatorname{grad} u &= 0 && \text{on } \Gamma_1, \end{aligned}$$

where $\lambda\colon \mathbf{R}^1 \to \mathbf{R}^1$ is a measurable bounded function such that

$$\lambda(\xi) \geq C > 0 \quad \forall \xi \in \mathbf{R}^1. \tag{3.1.4}$$

Note that such a nonlinear problem can be converted by the well-known Kirchhoff transformation (cf. [Čermák, Zlámal], [Larsson, Thomée, Zhang])

$$\mathcal{K}(U) = \int_0^U \lambda(\xi)\,d\xi, \quad U \in \mathbf{R}^1,$$

to the linear problem

$$\begin{aligned} -\Delta z &= f \quad &&\text{in } \Omega,\\ z &= 0 \quad &&\text{on } \Gamma_0,\\ n^T \operatorname{grad} z &= 0 \quad &&\text{on } \Gamma_1, \end{aligned}$$

where $z(x) = \mathcal{K}(u(x))$. From (3.1.4) we observe that $\mathcal{K}$ is an increasing function, i.e., its inverse $\mathcal{K}^{-1}$ exists and we have $u(x) = \mathcal{K}^{-1}(z(x))$.

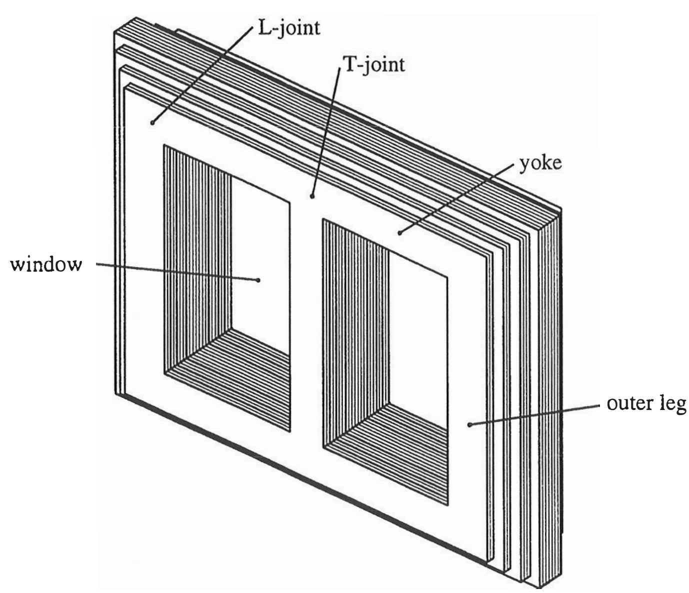

Figure 3.1.1.

The Kirchhoff transformation, however, cannot be applied in the case of anisotropic nonlinear media. For instance, in examining a temperature field in the magnetic circuit of a transformer from Figure 3.1.1, nonlinear temperature dependencies of the heat conductivities across and along the lamination differ. This is

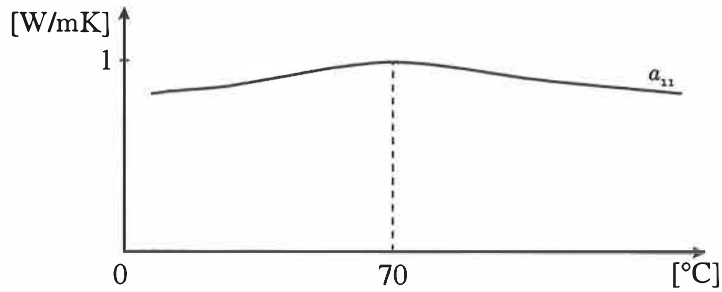

Figure 3.1.2.

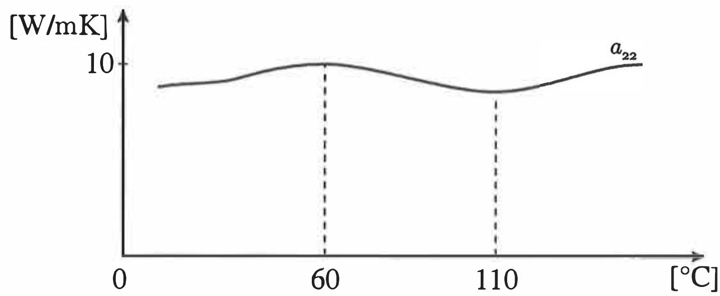

Figure 3.1.3.

the case of an orthotropic material. The associated 3×3 matrix $\mathcal{A}$ of heat conductivities is diagonal and such that $a_{11} \neq a_{22} = a_{33}$. The temperature dependencies of the diagonal entries are illustrated in Figures 3.1.2 and 3.1.3. We see that the type of nonlinearity is different in each direction (it, moreover, depends upon the type of iron used to construct the sheets).

3.2. Weak formulation

Since the Kirchhoff transformation cannot be used, the another approach to solve the problem is developed (see Section 3.3).

First, we state a weak formulation of the problem (3.1.1)–(3.1.3). We assume that $\mathcal{A} = \mathcal{A}(\cdot\,,\cdot)$ and $\alpha = \alpha(\cdot)$ are bounded measurable functions,

$$\operatorname*{ess\,sup}_{x,\xi,i,j} |a_{ij}(x,\xi)| \leq C, \quad \operatorname*{ess\,sup}_{s} |\alpha(s)| \leq C, \tag{3.2.1}$$

where $x \in \Omega$, $\xi \in \mathbf{R}^1$, $i,j \in \{1,...,d\}$ and $s \in \Gamma_1$. The components a_{ij} are assumed to be Lipschitz continuous with respect to the second variable, i.e., there exists $C_L > 0$ such that for all $\zeta, \xi \in \mathbf{R}^1$ and almost all $x \in \Omega$ we have

$$|a_{ij}(x,\zeta) - a_{ij}(x,\xi)| \leq C_L|\zeta - \xi|, \quad i,j = 1,...,d. \tag{3.2.2}$$

Moreover, let there exist $C_0 > 0$ such that for almost all $x \in \Omega$

$$C_0\eta^T\eta \leq \eta^T\mathcal{A}(x,\xi)\eta \quad \forall\xi \in \mathbf{R}^1 \quad \forall\eta \in \mathbf{R}^d \tag{3.2.3}$$

and let

$$0 \leq \alpha(s)$$

for almost all $s \in \Gamma_1$. Finally, let $f \in L^2(\Omega)$, $\overline{u} \in H^1(\Omega)$, $g \in L^2(\Gamma_1)$ and

$$V = \{v \in H^1(\Omega) \mid v = 0 \text{ on } \Gamma_0\}.$$

For simplicity, a possible dependence of $\mathcal{A}$ on x will usually not be explicitly indicated in what follows. Set

$$a(y; w, v) = (\mathcal{A}(y) \operatorname{grad} w, \operatorname{grad} v)_{0,\Omega} + \langle \alpha w, v \rangle_{0,\Gamma_1}, \quad y, w, v \in H^1(\Omega), \tag{3.2.4}$$

$$F(v) = (f, v)_{0,\Omega} + \langle g, v \rangle_{0,\Gamma_1}, \qquad v \in H^1(\Omega), \tag{3.2.5}$$

where $\langle \cdot\,, \cdot \rangle_{0,\Gamma_1}$ stands for the usual scalar product in $L^2(\Gamma_1)$. Since $\mathcal{A}$ and α are bounded, we observe that both the terms in (3.2.4) are finite, i.e., $a(\cdot\,;\cdot\,,\cdot)$ is well-defined.

Definition 3.2.1. *A function $u \in H^1(\Omega)$ is said to be a weak solution of the problem (3.1.1)–(3.1.3) if $u - \overline{u} \in V$ and*

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

Remark 3.2.2. From this place we suppose that the function $\overline{u}$, giving the Dirichlet boundary condition (3.1.2), is defined in the whole $\overline{\Omega}$ (cf. also (4.2.2) and Definition 4.2.1). In practice such a condition is usually given in the following form:

$$u = g(s) \quad \text{on } \Gamma_0,$$

where $g \in L_2(\Gamma_0)$. The reason of employing the function $\overline{u} \in H^1(\Omega)$ is the following. Let $\Gamma_0 = \partial\Omega$ for simplicity. Not every function $g \in L_2(\partial\Omega)$ need be the trace of some function $u(x)$ from the space $H^1(\Omega)$ (see, e.g., [Nečas, p. 22]), but we search the solution namely in this space. So, if the prescribed function $g(s)$ is not the trace of any function from $H^1(\Omega)$, then the corresponding problem cannot have a solution in the sense of Definition 3.2.1.

Also, there exists a relatively simple criteria, which holds for all considered in the dissertation cases, ensuring the existence of such a function $\overline{u}$, the trace of which coincides with the given function $g(s)$, see [Rektorys, Theorem 4.6.4]. Hence, from now on we accept the formulation given by Definition 3.2.1 when speaking about a fulfillment of the Dirichlet condition.

Remark 3.2.3. Note that to prove the existence of a weak solution $u \in V$ we cannot apply the main theorem for monotone operators, since our problem does not lead to a monotone operator, in general, see [Křížek, Neittaanmäki, 1996, Remark 9.2].

Moreover, the problem cannot be transformed to the minimization of a real functional, since the associated operator $\mathcal{A}$ is not potential in general, see again [Křížek, Neittaanmäki, 1996, Remark 9.3].

3.3. Existence of the weak and discrete solutions

Here we shall assume for simplicity that

$$d \in \{2,3\} \quad \text{and} \quad \Gamma_0 = \emptyset$$

which is an interesting case in practice. For the Dirichlet boundary conditions see Chapters 4–8, and for the mixed boundary conditions (3.1.2)–(3.1.3) (which can be possibly nonlinear) we refer to [Hlaváček, Křížek, Malý].

We shall look for the weak solution $u \in H^1(\Omega)$ such that (cf. Definition 3.2.1)

$$a(u; u, v) = F(v) \quad \forall v \in H^1(\Omega). \tag{3.3.1}$$

To guarantee the existence of such a function u, we moreover assume that there exist a constant $\alpha_0 > 0$ and a non-empty relatively open subset $\Gamma_2 \subset \Gamma_1 = \partial\Omega$ such that

$$\alpha(s) \geq \alpha_0$$

for almost all $s \in \Gamma_2$. Then there exists a constant $C_0 > 0$ such that

$$C_0 \|v\|_1^2 \leq a(y; v, v) \quad \forall y, v \in H^1(\Omega). \tag{3.3.2}$$

This inequality is a direct consequence of (3.2.3), (3.2.4) and the following Friedrichs' inequality (cf. (1.4.19)),

$$\|v\|_{1,\Omega}^2 \leq C(\|\operatorname{grad} v\|_{0,\Omega}^2 + \|v\|_{0,\Gamma_2}^2) \quad \forall v \in H^1(\Omega).$$

Remark 3.3.1. Using (3.2.4), (3.2.5), the boundedness of $\mathcal{A}$, α (see (3.2.1)) and the Trace Theorem, it is not difficult to verify that

$$|a(y; w, v)| \leq C\|w\|_1 \|v\|_1 \quad \forall y, w, v \in H^1(\Omega), \tag{3.3.3}$$

$$|F(v)| \leq C\|v\|_1 \quad \forall v \in H^1(\Omega). \tag{3.3.4}$$

Next theorem is useful.

Theorem 3.3.2. *Let $V_h \subset C(\overline{\Omega})$ be a non-empty finite-dimensional subspace. Then*

(i) there exists a Galerkin approximation $u_h \in V_h$ such that

$$a(u_h; u_h, v_h) = F(v_h) \quad \forall v_h \in V_h, \tag{3.3.5}$$

(ii) if the constant C_L from (3.2.2) is sufficiently small, there exists a unique Galerkin approximation u_h. Moreover, u_h can be calculated by means of the method of successive approximations (Kačanov's method) as follows:

Let $y^0 \in V_h$ be arbitrary. If $y^k \in V_h$ is known, $y^{k+1} \in V_h$ is defined by the relation

$$a(y^k; y^{k+1}, v_h) = F(v_h) \quad \forall v_h \in V_h \tag{3.3.6}$$

and

$$\|u_h - y^k\|_1 \to 0 \quad \text{as } k \to \infty.$$

Proof. The proof is based on the Lax–Milgram Lemma 1.5.2 and the Brouwer fixed point Theorem 1.4.11, and can be done as in [Křížek, Neittaanmäki, 1996, Theorem 9.5]. □

The problem of uniqueness of u_h is an open problem until now. Only if the coefficients are small or h is large (see [Hlaváček, Křížek, Malý]), we are able to prove the uniqueness.

Theorem 3.3.3 proves the existence of a solution u of the problem (3.3.1) as a weak limit of the Galerkin approximations u_h under the following assumption:

Let $\{V_h\}_{h\to 0}$ be a family of finite-dimensional subspaces of $H^1(\Omega) \cap C(\overline{\Omega})$ such that

$$\forall v \in C^\infty(\overline{\Omega})\ \exists \{v_h\}_{h\to 0} : v_h \in V_h,\ \|v - v_h\|_1 \to 0 \text{ as } h \to 0. \tag{3.3.7}$$

This condition is easy to satisfy for many families of finite element spaces.

Theorem 3.3.3. *Let (3.3.7) hold and let $\{u_h\}_{h\to 0}$ be a sequence of Galerkin approximations satisfying (3.3.5). Then there exist a subsequence $\{u_{\hat{h}}\} \subset \{u_h\}$ and an element $u \in H^1(\Omega)$ such that*

$$u_{\hat{h}} \rightharpoonup u \quad \text{(weakly) in } H^1(\Omega) \quad \text{as } \hat{h} \to 0, \tag{3.3.8}$$

and u is a solution of problem (3.3.1). Moreover, any weak cluster (accumulation) point of the sequence $\{u_h\}$ is a solution of (3.3.1).

Proof. The proof is based on the Eberlein–Schmulian Theorem 1.4.8 and is presented in [Křížek, Neittaanmäki, 1996, Theorem 9.6]. □

The corollary of the above theorem enables us to estimate the norm $\|u\|_1$ by the "data", i.e., the solution is stable in $\|\cdot\|_1$-norm.

Corollary 3.3.4. *There exists a constant $C > 0$ (independent of the data f and g) such that*

$$\|u\|_1 \le C(\|f\|_{0,\Omega} + \|g\|_{0,\partial\Omega}).$$

Remark 3.3.5. The existence of the weak solution for the nonlinear equation (3.1.1) with the pure Neumann boundary condition is examined in [Hlaváček].

3.4. Uniqueness and nonuniqueness

First of all note that if an elliptic equation is not in the divergence form there exist examples of nonunique solutions (see, e.g., [Gilbarg, Trudinger, p. 209], [Meyers, p. 178]). We can also get non-unique solutions of our problem (3.1.1)–(3.1.3) if the Lipschitz condition (3.2.2) is violated (see [Hlaváček, Křížek, Malý]).

Note that a proof of the uniqueness of the classical solution of the problem (3.1.1)–(3.1.3) is given by [Douglas, Dupont, Serrin] for the Dirichlet boundary conditions and by [Hlaváček, Křížek, 1993a] for mixed conditions. The uniqueness of the weak solution for the mixed nonlinear boundary conditions (the condition (3.2.2) is imposed) is proved in [Hlaváček, Křížek, Malý].

Throughout this section assume for simplicity again that $\Gamma_0 = \emptyset$, $d \in \{1, 2, \ldots\}$ and that there exists a constant $\alpha_0 > 0$ and a non-empty relatively open subset $\Gamma_2 \subset \Gamma_1 = \partial\Omega$ such that

$$\alpha \geq \alpha_0 \text{ on } \Gamma_2. \tag{3.4.1}$$

Theorem 3.4.1. *Let (3.2.1)–(3.2.3) and (3.4.1) hold and let u_1, $u_2 \in H^1(\Omega)$ be two weak solutions of the problem (3.3.1). Then $u_1 = u_2$ a.e. in Ω.*

For the proof see [Hlaváček, Křížek, Malý], where also mixed boundary conditions are considered. The proof is based on a special choice of a test function $v \in V$.

Another uniqueness theorem for nonlinear elliptic problem with the Dirichlet boundary condition is given in [Jensen].

3.5. Convergence of finite element approximations

From (3.3.8) and the compactness of the imbedding operator $H^1(\Omega) \to L^2(\Omega)$ we can easily prove the convergence of the Galerkin approximations in the $\|\cdot\|_0$-norm. To prove even the (strong) convergence in the $\|\cdot\|_1$-norm, we shall, in addition, require that

$$V_h \subset W_4^1(\Omega), \quad \|v_h\|_{1,4} \leq C(v) \quad \forall h, \tag{3.5.1}$$

where v_h satisfies (3.3.7) and $C(v)$ is a constant independent of h.

Remark 3.5.1. Using the standard interpolation theory one can verify (3.3.7) and (3.5.1) for many families of finite element spaces (see, e.g., [Ciarlet, p. 123]). The functions v_h can be defined, e.g., as the V_h-interpolant of v. Then we find that

$$\|v_h\|_{1,4} \leq \|v - v_h\|_{1,4} + \|v\|_{1,4} \leq C(v).$$

Theorem 3.5.2. *Let the assumptions of Theorem 3.3.3 be fulfilled and let (3.5.1) hold. Then the convergence (3.3.8) is strong, i.e.,*

$$\|u - u_{\hat{h}}\|_1 \to 0 \quad \text{as } \hat{h} \to 0. \tag{3.5.2}$$

Moreover, if there exists precisely one solution of the problem (3.3.1) then (3.5.2) holds for the whole sequence $\{u_h\}$.

For the proof see [Křížek, Neittaanmäki, 1996, Theorem 9.10].

Remark 3.5.3. We will list briefly another important results. The comparison principle is proved in [Křížek, Liu, 1996]. The discrete maximum principle is considered in [Křížek, Lin] for a special type of a triangulation. The rate of convergence is given in [Liu, Křížek, Neittaanmäki].

The other important questions as an effect of a numerical integration and an approximation of the curved boundary are discussed in details in Chapters 4–8.

Chapter 4
Setting the problem

From this chapter we begin to consider the main questions of Part II. Chapters 4–8 are based on the work [Korotov, Křížek], which is under review in "Numerische Mathematik".

4.1. Introduction

A finite element analysis of variational crimes, which arise in solving a nonlinear heat conduction problem in planar domains with a piecewise curved boundary, is presented in [Feistauer, Křížek, Sobotíková]. The main goal of Chapters 4–8 is to generalize this analysis to three-dimensional domains. In particular, we approximate a smooth boundary by a polyhedral one, we use appropriate numerical quadrature formulae (see [Engels]) to evaluate all integrals and, finally, we interpolate boundary conditions. Doing these approximations, we commit so-called *variational crimes*, in virtue of which the used finite element method becomes nonconforming, cf. Section 2.5.

A detailed analysis of variational crimes for linear boundary value problems is given, e.g., in [Ciarlet], [Ciarlet, Raviart], [Strang]. Its extension to a class of nonlinear plane elliptic problems of monotone type was first done in [Feistauer, Ženíšek, 1987].

This analysis was later generalized into several directions: to pseudomonotone operators [Feistauer, Křížek, Sobotíková], [Feistauer, Ženíšek, 1988], to nonlinear elliptic problems with discontinuous coefficients [Feistauer, Sobotíková, 1990], [Ženíšek, 1990a], to nonlinear boundary conditions [Feistauer, Felcman, Rokyta, Vlášek], [Feistauer, Najzar], to polyhedral domains [Feistauer, Křížek, Sobotíková], [Křížek, Lin], etc.

Some of these results are also surveyed in monographs [Feistauer, 1993], [Křížek, Neittaanmäki, 1996], [Ženíšek, 1990b]. However, the above mentioned references do not contain any finite element analysis of a nonlinear elliptic problem in three-dimensional domains with curved boundaries.

For simplicity, let $\Omega \subset \mathbf{R}^3$ be a bounded domain with a smooth boundary. There are several approaches how to treat the curved boundary $\partial\Omega$. The first one is to employ isoparametric elements. However, they do not have a simple form in $\mathbf{R}^3$ (see [Bernardi] and [Ciarlet]). The ansatz shape functions are nonpolynomial; in favorite cases rational, but, in general, they are quite complicated. Note that for a small discretization parameter h it is not possible to decompose $\overline{\Omega}$ into tetrahedral elements having at most one face curved (cf. [Křížek, Neittaanmäki, 1990, p. 76]). Thus, isoparametric elements with at least two curved faces are used (see [Lenoir]).

Another approach is to approximate Ω by a polyhedron Ω_h and then decompose $\overline{\Omega}_h$ into straight (non-curved) elements. Since we usually have no a priori information about the regularity of the true solution of a nonlinear problem, lower order finite elements on Ω_h have to be applied. This manner is often used in practical calculations, but for a theoretical finite element analysis the entire domain Ω should be taken into account.

In [Knobloch], the shape of tetrahedral elements near $\partial\Omega_h$ are slightly changed in such a way that Ω is completely covered by them.

Here we present a different approach. We assume for simplicity that Ω is convex and $\Omega_h \subset \Omega$. The set $\overline{\Omega} \setminus \Omega_h$ is decomposed into two kinds of special elements – hat and slice elements, see Chapter 6. They are not applied for computer implementation, but only to prove the convergence.

Problems which we meet, when $\partial\Omega$ is not smooth, are outlined in Remarks 6.1.4 and 6.1.5. The mixed boundary conditions are not considered, since the set of points, where one boundary condition changes to another, can have a very complicated structure.

Throughout Chapters 4–8 the standard Sobolev space notation from Chapter 1 is used. The symbol C (possibly with indices) stands for a generic constant, which will always be independent of the discretization parameter h.

4.2. Classical and variational formulations. Assumptions on the data

Consider the following quasilinear elliptic problem with the Dirichlet boundary condition whose classical formulation reads: Find $u \in C(\overline{\Omega})$ such that $u|_\Omega \in C^2(\Omega)$ and

$$-\operatorname{div}(\mathcal{A}(\cdot\,,u)\operatorname{grad} u) = f \quad \text{in } \Omega, \tag{4.2.1}$$

$$u = \overline{u} \quad \text{on } \partial\Omega, \tag{4.2.2}$$

where Ω is a bounded convex domain in $\mathbf{R}^3$ with a C^2-smooth boundary $\partial\Omega$. Since quadrature formulae will be employed later, we need stronger smoothness assumptions on data (cf. 3.2.1)–(3.2.3)), in particular, let $\mathcal{A} = (a_{ij})_{i,j=1}^3$, $a_{ij} = a_{ij}(x,\xi)$, $x \in \overline{\Omega}$, $\xi \in \mathbf{R}^1$ and let a_{ij}, $\partial a_{ij}/\partial x_k$, $\partial a_{ij}/\partial\xi$ be continuous and bounded in $\overline{\Omega} \times \mathbf{R}^1$ for all $i,j,k = 1,2,3$. Thus, there exists a positive constant C such that

$$|a_{ij}(x,\xi)| \le C, \quad \left|\frac{\partial a_{ij}}{\partial x_k}(x,\xi)\right| \le C, \quad \left|\frac{\partial a_{ij}}{\partial \xi}(x,\xi)\right| \le C \quad \forall x \in \Omega \quad \forall \xi \in \mathbf{R}^1. \tag{4.2.3}$$

The boundedness of the derivatives $\partial a_{ij}/\partial\xi$ obviously implies the Lipschitz-continuity of a_{ij} with respect to ξ, i.e.,

$$|a_{ij}(x,\zeta) - a_{ij}(x,\xi)| \le C_L|\zeta - \xi|, \quad i,j = 1,2,3. \tag{4.2.4}$$

Moreover, let there exist $C_0 > 0$ such that for almost all $x \in \Omega$

$$C_0\eta^T\eta \le \eta^T\mathcal{A}(x,\xi)\eta \quad \forall \xi \in \mathbf{R}^1 \quad \forall \eta \in \mathbf{R}^3. \tag{4.2.5}$$

Finally, let $f \in W^1_\infty(\Omega)$, $\overline{u} \in W^1_p(\Omega)$ with $p > 3$ fixed and

$$V = \overset{\circ}{W}{}^1_2(\Omega) \equiv W^1_2(\Omega, \partial\Omega). \tag{4.2.6}$$

For simplicity, a possible dependence of $\mathcal{A}$ on x will usually not be explicitly indicated in what follows. Set

$$a(y; w, v) = (\mathcal{A}(y) \operatorname{grad} w, \operatorname{grad} v)_{0,\Omega}, \quad y, w, v \in W^1_2(\Omega), \tag{4.2.7}$$

$$F(v) = (f, v)_{0,\Omega}, \qquad v \in W^1_2(\Omega). \tag{4.2.8}$$

Since the elements of $\mathcal{A}$ are bounded (cf. (4.2.3)), we observe that the term in (4.2.7) is finite, i.e., $a(\cdot\,;\cdot\,,\cdot)$ is well-defined.

Definition 4.2.1. *A function $u \in W^1_2(\Omega)$ is said to be a weak solution of the problem (4.2.1)–(4.2.2) if $u - \overline{u} \in V$ and*

$$a(u; u, v) = F(v) \quad \forall v \in V. \tag{4.2.9}$$

According to [Hlaváček, Křížek, Malý] there exists precisely one weak solution $u \in W^1_2(\Omega)$.

From the properties (4.2.3) of the matrix $\mathcal{A}$ it follows that

$$|a(y; w, v)| \le C\|w\|_{1,\Omega}\|v\|_{1,\Omega} \quad \forall y, w, v \in W^1_2(\Omega). \tag{4.2.10}$$

Moreover, from the Lipschitz continuity (4.2.4) we have

$$\begin{aligned} &|a(y; w, v) - a(z; w, v)| \le |((\mathcal{A}(\cdot\,, y) - \mathcal{A}(\cdot\,, z)) \operatorname{grad} w, \operatorname{grad} v)_{0,\Omega}| \\ &\le \|\mathcal{A}(\cdot\,, y) - \mathcal{A}(\cdot\,, z)\|_{0,\Omega}\|\operatorname{grad} w\|_{0,\Omega}\|\operatorname{grad} v\|_{0,\infty,\Omega} \\ &\le d^2 C_L \|y - z\|_{0,\Omega}\|w\|_{1,\Omega}\|v\|_{1,\infty,\Omega} \quad \forall y, z, w \in W^1_2(\Omega) \quad \forall v \in W^1_\infty(\Omega). \end{aligned} \tag{4.2.11}$$

Further we introduce a polyhedral approximation Ω_h of Ω. For a given discretization parameter $h \in (0, h_0)$ let $\mathcal{T}_h$ consist of closed tetrahedra K such that, cf. Section 2.4:

1) $h_K = \operatorname{diam} K \le Ch$ for all $K \in \mathcal{T}_h$,
2) $\overline{\Omega}_h = \bigcup_{K \in \mathcal{T}_h} K \subset \overline{\Omega}$,
3) Ω_h is convex,
4) all vertices of $\overline{\Omega}_h$ belong to $\partial\Omega$,
5) any face of any tetrahedron $K \in \mathcal{T}_h$ is a face of another tetrahedron from $\mathcal{T}_h$ or a part of the boundary $\partial\Omega_h$.

Assumption 3) is used in Chapters 6 and 7. Note, that it does not follow from the convexity of Ω. To see this we can easily construct a nonconvex polyhedron, which consists only of two tetrahedra and is inscribed to a ball.

The set $\mathcal{T}_h$ will be called a *partition* of $\overline{\Omega}_h$ into tetrahedra. Consider families $\{\Omega_h\}$, $h \in (0, h_0)$, of polyhedral approximations of Ω and $\{\mathcal{T}_h\}$, $h \in (0, h_0)$, of partitions of $\overline{\Omega}_h$ into tetrahedra (with $h_0 > 0$ sufficiently small). Let the family $\{\mathcal{T}_h\}$, $h \in (0, h_0)$, be strongly regular, i.e., there exists $C > 0$ such that (cf. (2.7.2))

$$Ch^3 \leq \operatorname{meas} K$$

for all $K \in \mathcal{T}_h$ and all $h \in (0, h_0)$. It is clear, that at most three vertices of any $K \in \mathcal{T}_h$ belong to $\partial\Omega$ provided h is sufficiently small.

For any $h \in (0, h_0)$ we set

$$\begin{aligned} X_h &= \{v_h \in C(\overline{\Omega}_h) \mid v_h|_K \in P_1(K) \quad \forall K \in \mathcal{T}_h\}, \\ V_h &= \{v_h \in X_h \mid v_h|_{\partial\Omega_h} = 0\}, \\ \tilde{a}_h(y; w, v) &= (\mathcal{A}(y) \operatorname{grad} w, \operatorname{grad} v)_{0,\Omega_h}, \quad y, w, v \in W_2^1(\Omega_h), \\ \tilde{F}_h(v) &= (f, v)_{0,\Omega_h}, \quad v \in W_2^1(\Omega_h), \end{aligned} \tag{4.2.12}$$

where $P_1(K)$ is the space of linear polynomials over the tetrahedron K.

In view of the Sobolev Imbedding Theorem we have $\overline{u} \in C(\overline{\Omega})$ and, thus, it makes sense to define the Lagrange interpolant $\pi_h\overline{u} \in X_h$. Recall that $\pi_h\overline{u}(P) = \overline{u}(P)$ for every vertex P of $K \in \mathcal{T}_h$. We set

$$\overline{u}_h = \pi_h\overline{u}. \tag{4.2.13}$$

Lemma 4.2.2. *The Lagrange interpolant $\overline{u}_h$ has the following properties:*

$$\lim_{h\to 0} \|\overline{u} - \overline{u}_h\|_{1,\Omega_h} = 0, \tag{4.2.14}$$

$$\|\overline{u}_h\|_{1,\Omega_h} \leq C \quad \forall h \in (0, h_0). \tag{4.2.15}$$

Proof. Recall that $\overline{\Omega}_h \subset \overline{\Omega}$ for all $h \in (0, h_0)$ with sufficiently small $h_0 > 0$ due to the condition 2).

From [Ciarlet, Theorem 3.1.6] it follows that there exists a constant $C > 0$ independent of h such that

$$\|v - \pi_h v\|_{1,p,\Omega_h} \leq C\|v\|_{1,p,\Omega_h} \quad \forall v \in W_p^1(\Omega) \quad \forall h \in (0, h_0). \tag{4.2.16}$$

From (4.2.16) and the inclusion $\overline{\Omega}_h \subset \overline{\Omega}$ we immediately get

$$\begin{aligned} &\|\pi_h v\|_{1,p,\Omega_h} \leq \|\pi_h v - v\|_{1,p,\Omega_h} + \|v\|_{1,p,\Omega_h} \leq \\ &\leq C'\|v\|_{1,p,\Omega_h} \leq C'\|v\|_{1,p,\Omega} \quad \forall v \in W_p^1(\Omega) \quad \forall h \in (0, h_0), \end{aligned} \tag{4.2.17}$$

where $C' = C + 1$.

Again from [Ciarlet, Theorem 3.1.6] we have

$$\|v - \pi_h v\|_{1,p,\Omega_h} \leq C'' h \|v\|_{2,p,\Omega_h} \quad \forall v \in W_p^2(\Omega), \tag{4.2.18}$$

where C'' is a constant independent of h and v.

Let ε be an arbitrary small positive number. According to [Nečas, 1967, Chapter 2], the space $C^\infty(\overline{\Omega})$ is dense in $W_p^1(\Omega)$ and one can choose $v \in C^\infty(\overline{\Omega})$ such that

$$\|\overline{u} - v\|_{1,p,\Omega_h} \leq \|\overline{u} - v\|_{1,p,\Omega} \leq \frac{\varepsilon}{3C'}, \tag{4.2.19}$$

where C' is a constant from inequality (4.2.17). Further, from (4.2.17) and (4.2.19) we have

$$\|\pi_h(\overline{u} - v)\|_{1,p,\Omega_h} \leq C' \|\overline{u} - v\|_{1,p,\Omega} \leq \frac{\varepsilon}{3}. \tag{4.2.20}$$

According to (4.2.18) and the imbedding $C^\infty(\overline{\Omega}) \subset W_p^2(\Omega)$, there exists a constant $h_\varepsilon \in (0, h_0)$ such that

$$\|v - \pi_h v\|_{1,p,\Omega_h} \leq \frac{\varepsilon}{3} \quad \forall h \in (0, h_\varepsilon). \tag{4.2.21}$$

Further, (4.2.19), (4.2.20) and (4.2.21) imply that

$$\begin{aligned}\|\overline{u} - \overline{u}_h\|_{1,p,\Omega_h} &= \|\overline{u} - \pi_h \overline{u}\|_{1,p,\Omega_h} \\ &\leq \|\overline{u} - v\|_{1,p,\Omega_h} + \|v - \pi_h v\|_{1,p,\Omega_h} + \|\pi_h v - \pi_h \overline{u}\|_{1,p,\Omega_h} \\ &\leq \frac{\varepsilon}{3C'} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} < \varepsilon.\end{aligned}$$

The last inequality is valid due to the fact that $C' = C + 1 > 1$. Hence, we have

$$\lim_{h \to 0} \|\overline{u} - \overline{u}_h\|_{1,p,\Omega_h} = 0. \tag{4.2.22}$$

Using the Hölder inequality 1.4.3 and the inclusion $\Omega_h \subset \Omega$, we get the following inequalities (under our assumptions that $p > 3$ and fixed)

$$\begin{aligned}\|\psi\|_{0,\Omega_h} &= \left(\int_{\Omega_h} |\psi|^2 \, dx \right)^{1/2} \\ &\leq \left(\int_{\Omega_h} 1^{\frac{p}{p-2}} \, dx \right)^{\frac{p-2}{2p}} \cdot \left(\int_{\Omega_h} |\psi|^p \, dx \right)^{1/p} \\ &\leq (\operatorname{meas} \Omega_h)^{\frac{p-2}{2p}} \|\psi\|_{0,p,\Omega_h} \\ &\leq (\operatorname{meas} \Omega)^{\frac{p-2}{2p}} \|\psi\|_{0,p,\Omega_h} \\ &\leq C''' \|\psi\|_{0,p,\Omega_h} \qquad \forall \psi \in L_p(\Omega_h),\end{aligned} \tag{4.2.23}$$

where the constant C''' does not depend on h, ψ and p. Applying (4.2.23) to $\overline{u}_h - \overline{u}$ and $\partial_i(\overline{u}_h - \overline{u})$, $i = 1, 2, 3$, and using (4.2.22), we get (4.2.14).

Further,

$$\|\overline{u}_h\|_{1,\Omega_h} \leq \|\overline{u}_h - \overline{u}\|_{1,\Omega_h} + \|\overline{u}\|_{1,\Omega_h} \leq \|\overline{u}_h - \overline{u}\|_{1,\Omega_h} + \|\overline{u}\|_{1,\Omega} \leq \overline{C},$$

where $\overline{C}$ is some constant independent of h. Note, that for derivation of (4.2.15) we used (4.2.14) with sufficiently small h_0. □

Chapter 5
Approximate solution

5.1. Auxiliary results

We will employ the following numerical integration formula over an element $K \in \mathcal{T}_h$

$$\int_K g(x)\,dx \approx \operatorname{meas} K \sum_{k=1}^{M_K} c_{K,k}\, g(x_{K,k}), \tag{5.1.1}$$

where the weights $c_{K,k} \in \mathbf{R}^1$ are such that

$$c_{K,k} > 0 \quad \text{and} \quad \sum_{k=1}^{M_K} c_{K,k} = 1, \tag{5.1.2}$$

and the nodes $x_{K,k} \in K$ for $k = 1, \ldots, M_K$.

For any $v_h \in X_h$ and any $K \in \mathcal{T}_h$ we set

$$v_K = v_h|_K.$$

Definition 5.1.1. *A function $u_h \in X_h$ is said to be an approximate solution of the problem (4.2.1)–(4.2.2) if*

$$u_h - \overline{u}_h \in V_h, \tag{5.1.3}$$

$$a_h(u_h; u_h, v_h) = F_h(v_h) \quad \forall v_h \in V_h, \tag{5.1.4}$$

where

$$a_h(y_h; w_h, v_h) = \sum_{K \in \mathcal{T}_h} \operatorname{meas} K \sum_{k=1}^{M_K} c_{K,k} (\mathcal{A}(x_{K,k}, y_K(x_{K,k})) \operatorname{grad} w_K)^T \operatorname{grad} v_K \tag{5.1.5}$$

for $y_h, w_h, v_h \in X_h$ and

$$F_h(v_h) = \sum_{K \in \mathcal{T}_h} \operatorname{meas} K \sum_{k=1}^{M_K} c_{K,k} f(x_{K,k}) v_K(x_{K,k}), \quad v_h \in X_h. \tag{5.1.6}$$

Lemma 5.1.2. *The seminorm* $|\cdot|_{1,\Omega_h}$ *(in* $W_2^1(\Omega_h)$*) is a norm on* V_h *uniformly equivalent to the norm* $\|\cdot\|_{1,\Omega_h}$*, i.e., there exists a constant* C $(0 < C \leq 1)$ *independent of* h *and* v *such that*

$$C\|v\|_{1,\Omega_h} \leq |v|_{1,\Omega_h} \leq \|v\|_{1,\Omega_h} \quad \forall v \in V_h \quad \forall h \in (0, h_0). \tag{5.1.7}$$

Proof. Consider the space $V = \overset{\circ}{W}{}_2^1(\Omega)$. It is well-known that $|\cdot|_{1,\Omega}$ and $\|\cdot\|_{1,\Omega}$ are equivalent on V (see, e.g., [Mikhailov, 1978, p. 149]), i.e., there exists a constant $C > 0$ independent of v such that

$$C\|v\|_{1,\Omega} \leq |v|_{1,\Omega} \leq \|v\|_{1,\Omega} \quad \forall v \in V. \tag{5.1.8}$$

For sufficiently small h_0 we have $\Omega_h \subset \Omega$ for all $h \in (0, h_0)$. Any function $v \in V_h$ may be extended by zero on $\Omega \setminus \Omega_h$ and, if we define a function v' as

$$v' = \begin{cases} v & \text{for } x \in \Omega_h, \\ 0 & \text{for } x \in \Omega \setminus \overline{\Omega}_h, \end{cases}$$

then $v' \in V$.

We will prove that the constant C from (5.1.8) can be used for (5.1.7) as well. Indeed, for any $v \in V_h$ the corresponding function $v' \in V$ and, moreover,

$$\begin{aligned} |v|_{1,\Omega_h} &= |v'|_{1,\Omega}, \\ \|v\|_{1,\Omega_h} &= \|v'\|_{1,\Omega}. \end{aligned} \tag{5.1.9}$$

Using (5.1.8) for the function v', we have

$$C\|v'\|_{1,\Omega} \leq |v'|_{1,\Omega} \leq \|v'\|_{1,\Omega},$$

which implies, in view of (5.1.9), that

$$C\|v\|_{1,\Omega_h} \leq |v|_{1,\Omega_h} \leq \|v\|_{1,\Omega_h}$$

with C independent of v and h. □

Remark 5.1.3. The following inverse inequality holds, see [Ciarlet, Theorem 3.2.6],

$$|v_h|_{1,q,\Omega_h} \leq \frac{C}{(h^d)^{\max(0,1/2-1/q)}}|v_h|_{1,\Omega_h} \quad \forall h \in (0, h_0) \quad \forall v_h \in X_h \tag{5.1.10}$$

for any strongly regular family of partitions and $q \in [1, \infty]$, where the constant C depends on q only.

5.2. Properties of the forms a_h, F_h

Lemma 5.2.1. *There exists a positive constant C_1 independent of $h \in (0, h_0)$ such that*

$$a_h(y_h; v_h, v_h) \geq C_1 \|v_h\|^2_{1,\Omega_h} \quad \forall y_h \in X_h \quad \forall v_h \in V_h. \tag{5.2.1}$$

Proof. Recall that $v_K = v_h|_K$, where $v_h \in X_h$. Hence, for fixed $v_h \in X_h$, v_K is a linear function and $\operatorname{grad} v_K$ is a constant vector for any $K \in \mathcal{T}_h$. From (5.1.5), (4.2.5), (5.1.2) and (5.1.7) we have (C_0 is a constant from (4.2.5))

$$\begin{aligned}
a_h(y_h; v_h, v_h) &= \sum_{K \in \mathcal{T}_h} \operatorname{meas} K \sum_{k=1}^{M_K} c_{K,k} (\mathcal{A}(x_{K,k}, y_K(x_{K,k})) \operatorname{grad} v_K)^T \operatorname{grad} v_K \\
&\geq C_0 \sum_{K \in \mathcal{T}_h} \operatorname{meas} K \sum_{k=1}^{M_K} c_{K,k} \| \operatorname{grad} v_K \|^2 \\
&= C_0 \sum_{K \in \mathcal{T}_h} \operatorname{meas} K \| \operatorname{grad} v_K \|^2 \\
&= C_0 |v_h|^2_{1,\Omega_h} \geq C_1 \|v_h\|^2_{1,\Omega_h},
\end{aligned}$$

where C_1 is independent of h. □

Lemma 5.2.2. *There exists a positive constant C_2 independent of $h \in (0, h_0)$ such that*

$$|a_h(y_h; w_h, v_h)| \leq C_2 \|w_h\|_{1,\Omega_h} \|v_h\|_{1,\Omega_h} \quad \forall y_h, w_h, v_h \in X_h. \tag{5.2.2}$$

Proof. First, we prove in details, an elementary auxiliary result. Let x, y be vectors from $\mathbf{R}^3$, $A = (a_{ij})^3_{i,j=1}$ be a matrix with $|a_{ij}| \leq C$, $i, j \in \{1, 2, 3\}$, where C is some positive constant. Then we have the following useful estimate:

$$\begin{aligned}
|(Ax)^T y| &= \left| \sum_{i=1}^{3} \sum_{j=1}^{3} x_i a_{ij} y_j \right| \\
&\leq \sum_{i=1}^{3} \sum_{j=1}^{3} |a_{ij}|\, |x_i|\, |y_j| \leq C \sum_{i=1}^{3} \sum_{j=1}^{3} |x_i|\, |y_j| \\
&= C \Big(\sum_{i=1}^{3} |x_i| \Big) \Big(\sum_{j=1}^{3} |y_j| \Big) \\
&\leq 3C \Big(\sum_{i=1}^{3} |x_i|^2 \Big)^{1/2} \Big(\sum_{j=1}^{3} |y_j|^2 \Big)^{1/2}.
\end{aligned}$$

The last inequality is due to the well-known "norm inequality".

Now, from the above auxiliary result, (5.1.5), (5.1.2), (5.1.7) and the discrete Hölder inequality we have

$$
\begin{aligned}
|a_h(y_h; w_h, v_h)| &= \Big| \sum_{K\in\mathcal{T}_h} \operatorname{meas} K \sum_{k=1}^{M_K} c_{K,k} (\mathcal{A}(x_{K,k}, y_K(x_{K,k})) \operatorname{grad} w_K)^T \operatorname{grad} v_K \Big| \\
&\le \sum_{K\in\mathcal{T}_h} \operatorname{meas} K \sum_{k=1}^{M_K} c_{K,k} \big| (\mathcal{A}(x_{K,k}, y_K(x_{K,k}) \operatorname{grad} w_K)^T \operatorname{grad} v_K \big| \\
&\le C_2 \sum_{K\in\mathcal{T}_h} \operatorname{meas} K \sum_{k=1}^{M_K} c_{K,k} \|\operatorname{grad} w_K\| \|\operatorname{grad} v_K\| \\
&= C_2 \sum_{K\in\mathcal{T}_h} \big[\sqrt{\operatorname{meas} K} \|\operatorname{grad} w_K\|\big] \big[\sqrt{\operatorname{meas} K} \|\operatorname{grad} v_K\|\big] \\
&= C_2 |w_h|_{1,\Omega_h} |v_h|_{1,\Omega_h} \le C_2 \|w_h\|_{1,\Omega_h} \|v_h\|_{1,\Omega_h},
\end{aligned}
$$

where C_2 is some positive constant independent of h.

□

Lemma 5.2.3. *Let $f \in W^1_\infty(\Omega)$ and let (5.1.2) hold. Then there exists some positive constant C_3 such that*

$$|F_h(v_h)| \le C_3 \|v_h\|_{1,\Omega_h} \quad \forall v_h \in X_h. \tag{5.2.3}$$

Proof. First we prove that

$$|\widetilde{F}_h(v_h) - F_h(v_h)| \le \overline{C} h \|f\|_{1,\infty,\Omega_h} \|v_h\|_{1,\Omega_h} \quad \forall v_h \in X_h, \tag{5.2.4}$$

where $\overline{C}$ is some positive constant independent of h. Indeed, in view of [Ciarlet, Theorem 4.1.5] we have the following estimate

$$
\begin{aligned}
|\widetilde{F}_h(v_h) - F_h(v_h)| &\le \sum_{K\in\mathcal{T}_h} \Big| \int_K f v_K \, dx - \operatorname{meas} K \sum_{k=1}^{M_K} c_{K,k} f(x_{K,k}) v_K(x_{K,k}) \Big| \\
&\le \sum_{K\in\mathcal{T}_h} C h_K (\operatorname{meas} K)^{1/2-1/q} \|f\|_{1,q,K} \|v_K\|_{1,2,K}
\end{aligned}
$$

for $q \in (3, \infty]$, where $h_K = \operatorname{diam} K$.

Taking $q = \infty$ (note that in the above inequality the constant C may depend on q, but if q is fixed then C is fixed) and using the obvious inequality

$$\|f\|_{1,\infty,K} \le \|f\|_{1,\infty,\Omega_h}, \tag{5.2.5}$$

we have by the Cauchy–Schwarz inequality that

$$
\begin{aligned}
|\widetilde{F}_h(v_h) - F_h(v_h)| &\le C\|f\|_{1,\infty,\Omega_h} \sum_{K\in\mathcal{T}_h} h_K (\operatorname{meas} K)^{1/2} \|v_K\|_{1,2,K} \\
&\le C\|f\|_{1,\infty,\Omega_h} \Big(\sum_{K\in\mathcal{T}_h} h_K^2 \operatorname{meas} K \Big)^{1/2} \Big(\sum_{K\in\mathcal{T}_h} \|v_K\|_{1,2,K}^2 \Big)^{1/2} \\
&\le \overline{C} h \|f\|_{1,\infty,\Omega_h} \|v_h\|_{1,\Omega_h}.
\end{aligned}
$$

Moreover, $\|f\|_{1,\infty,\Omega_h} \le \|f\|_{1,\infty,\Omega}$ and, thus, we have

$$
|\widetilde{F}_h(v_h) - F_h(v_h)| \le \overline{C} h \|f\|_{1,\infty,\Omega} \|v_h\|_{1,\Omega_h}. \tag{5.2.6}
$$

Since

$$
\widetilde{F}_h(v) = \int_{\Omega_h} f v \, dx, \quad v \in H^1(\Omega_h),
$$

then

$$
|\widetilde{F}_h(v_h)| \le C' \|v_h\|_{1,\Omega_h} \quad \forall v_h \in X_h, \tag{5.2.7}
$$

with a positive constant C' independent of v_h and h. Hence, from (5.2.6) and (5.2.7) we prove (5.2.3) in the following manner

$$
\begin{aligned}
|F_h(v_h)| &\le |F_h(v_h) - \widetilde{F}_h(v_h)| + |\widetilde{F}_h(v_h)| \\
&\le \overline{C} h \|f\|_{1,\infty,\Omega} \|v_h\|_{1,\Omega_h} + C' \|v_h\|_{1,\Omega_h} \\
&\le C_3 \|v_h\|_{1,\Omega_h} \qquad \forall v_h \in X_h.
\end{aligned}
$$

□

5.3. Existence of the approximate solution

Theorem 5.3.1. *Let the assumptions of Lemma 5.2.3 and assumptions (4.2.3) on* a_{ij} *be satisfied. Then, for any* $h \in (0, h_0)$ *there exists an approximate solution* $u_h \in X_h$ *from Definition 5.1.1. Moreover, there exists a constant* $C > 0$ *independent of* h *such that*

$$
\|u_h\|_{1,\Omega_h} \le C \quad \forall h \in (0, h_0). \tag{5.3.1}
$$

Proof. To prove the theorem we use the scheme of proof as described in [Křížek, Neittaanmäki, 1996, Theorem 9.16] with some modifications.

In view of (5.1.3), we can write the solution u_h in the following form

$$
u_h = \overline{u}_h + z_h, \quad z_h \in V_h. \tag{5.3.2}
$$

Consider a mapping S defined as follows: if $z \in V_h$ then $Sz \in V_h$ is the unique solution of the problem

$$a_h(\overline{u}_h + z; \overline{u}_h + Sz, v) = F_h(v) \quad \forall v \in V_h. \tag{5.3.3}$$

We show now that S is well-defined. In view of linearity of $a_h(\cdot\,;\cdot\,,\cdot)$ with respect to the second argument identity (5.3.3) can be rewritten in the following form

$$a_h(\overline{u}_h + z; Sz, v) = G_h(v) := F_h(v) - a_h(\overline{u}_h + z; \overline{u}_h, v) \quad \forall v \in V_h. \tag{5.3.4}$$

For $z \in V_h$ fixed, G_h is obviously a linear functional on V_h.

Let us show that it is also continuous. By (5.2.2), (5.2.3) and (4.2.15) we have

$$\begin{aligned} |G_h(v)| &= |F_h(v) - a_h(\overline{u}_h + z; \overline{u}_h, v)| \\ &\le |F_h(v)| + |a_h(\overline{u}_h + z; \overline{u}_h, v)| \\ &\le C_2\|v\|_{1,\Omega_h} + C_3\|\overline{u}_h\|_{1,\Omega_h}\|v\|_{1,\Omega_h} \\ &\le C_4\|v\|_{1,\Omega_h}, \end{aligned}$$

i.e.,

$$|G_h(v)| \le C_4\|v\|_{1,\Omega_h} \quad \forall v \in V_h, \tag{5.3.5}$$

where C_4 does not depend on h.

The Lax–Milgram Lemma 1.5.2, (5.2.1) and (5.2.2) imply that the mapping S is uniquely determined.

Setting $v = Sz$ in (5.3.4) and using (5.2.1) and (5.3.5), we may write that

$$C_1\|Sz\|^2_{1,\Omega_h} \le a_h(\overline{u}_h + z; Sz, Sz) = G_h(Sz) \le C_4\|Sz\|_{1,\Omega_h}, \tag{5.3.6}$$

where the constants C_1 and C_4 are from (5.2.1) and (5.3.5), respectively. Consequently, the mapping S is bounded,

$$\|Sz\|_{1,\Omega_h} \le C_4/C_1 \quad \forall z \in V_h. \tag{5.3.7}$$

Note, that the constant C_4/C_1 does not depend on h.

We show now that the mapping S is Lipschitz continuous in V_h, i.e., there exists a constant $C_h > 0$ such that

$$\|Sy - Sz\|_{1,\Omega_h} \le C_h\|y - z\|_{1,\Omega_h} \quad \forall y, z \in V_h. \tag{5.3.8}$$

Let y and z be arbitrary elements of V_h. Consider the matrix $\overline{A} = (\overline{a}_{ij})^3_{i,j=1} = \mathcal{A}(x_{K,k}, (u_h + z)(x_{K,k})) - \mathcal{A}(x_{K,k}, (u_h + y)(x_{K,k}))$. Its elements can be evaluated in view of (4.2.4) as follows

$$|\overline{a}_{ij}| \le C_L|z(x_{K,k}) - y(x_{K,k})| \le C_L\|z - y\|_{0,\infty,K} \le C_L\|z - y\|_{0,\infty,\Omega_h}, \quad i, j \in \{1, 2, 3\}.$$

Let $v = Sy - Sz$. Since $v \in V_h$, we have by the above inequality, (5.2.1), (5.3.3), (5.1.2), the triangle inequality (1.1.8), (5.3.7), the discrete Hölder inequality and (4.2.15) the following estimates:

$$
\begin{aligned}
C_1\|v\|^2_{1,\Omega_h} &\le a_h(\overline{u}_h + y; v, v)\\
&= a_h(\overline{u}_h + y; \overline{u}_h + Sy, v) - a_h(\overline{u}_h + y; \overline{u}_h + Sz, v)\\
&= F_h(v) - a_h(\overline{u}_h + y; \overline{u}_h + Sz, v)\\
&= a_h(\overline{u}_h + z; \overline{u}_h + Sz, v) - a_h(\overline{u}_h + y; \overline{u}_h + Sz, v)\\
&= \sum_{K\in\mathcal{T}_h} \operatorname{meas} K \sum_{k=1}^{M_K} c_{K,k}\big[(\mathcal{A}(x_{K,k}, (u_h + z)(x_{K,k}))\\
&\qquad - \mathcal{A}(x_{K,k}, (u_h + y)(x_{K,k})))\operatorname{grad}(\overline{u}_h + Sz)_K\big]\operatorname{grad} v_K\\
&\le C_L \sum_{K\in\mathcal{T}_h} \operatorname{meas} K \sum_{k=1}^{M_K} c_{K,k}\|z - y\|_{0,\infty,\Omega_h}\|\operatorname{grad}(\overline{u}_h + Sz)_K\|\|\operatorname{grad} v_K\|\\
&\le C_L\|z - y\|_{0,\infty,\Omega_h} \sum_{K\in\mathcal{T}_h} \operatorname{meas} K\|\operatorname{grad}(\overline{u}_h + Sz)_K\|\|\operatorname{grad} v_K\|\\
&\le C_L\|z - y\|_{0,\infty,\Omega_h}\big(|\overline{u}_h|_{1,\Omega_h} + |Sz|_{1,\Omega_h}\big)|v|_{1,\Omega_h}\\
&\le C_L\|z - y\|_{0,\infty,\Omega_h}(\overline{C} + C_4/C_1)\|v\|_{1,\Omega_h}\\
&\le C_5\|z - y\|_{0,\infty,\Omega_h}\|v\|_{1,\Omega_h}.
\end{aligned}
$$

Hence, from the equivalence of all norms in the finite-dimensional space V_h, we find that

$$\|Sy - Sz\|_{1,\Omega_h} = \|v\|_{1,\Omega_h} \le C_5\|z - y\|_{0,\infty,\Omega_h} \le C_h\|z - y\|_{1,\Omega_h}.$$

Therefore, (5.3.8) is proved.

Using (5.3.7), (5.3.8) and the Brouwer fixed-point theorem 1.4.11, we get the existence of a fixed point $z = Sz$. Setting $u_h = \overline{u}_h + z$, we obtain the approximate solution.

The estimation (5.3.1) is a consequence of the boundedness of S (see (5.3.7)) and (4.2.15). □

5.4. The error of numerical integration

Remark 5.4.1. Proving Lemma 5.2.3 we established (cf. (5.2.4))

$$|\widetilde{F}_h(v_h) - F_h(v_h)| \le Ch\|f\|_{1,\infty,\Omega}\|v_h\|_{1,\Omega_h} \quad \forall v_h \in X_h. \tag{5.2.6}$$

Lemma 5.4.2. *Let (5.1.2) hold and let the assumptions (4.2.3) on a_{ij} be satisfied. Then there exists some positive constant C independent of h such that*

$$|\tilde{a}_h(y_h; w_h, v_h) - a_h(y_h; w_h, v_h)| \leq$$
$$\leq Ch|w_h|_{1,\Omega_h}|v_h|_{1,\Omega_h} + Ch^{1-\frac{3}{p}}|y_h|_{1,\Omega_h}|w_h|_{1,\alpha,\Omega_h}|v_h|_{1,\beta,\Omega_h} \quad \forall y_h, w_h, v_h \in X_h, \tag{5.4.1}$$

where either

$$\alpha = p,\ \beta = 2 \quad or \quad \alpha = 2,\ \beta = p \quad (p > 3),$$

and $\tilde{a}_h(\cdot\,;\cdot\,,\cdot)$ and $a_h(\cdot\,;\cdot\,,\cdot)$ are defined by (4.2.12) and (5.1.5), respectively.

Proof. The proof is based on the inverse inequality 5.1.10, and can be easily done as in [Křížek, Neittaanmäki, 1996, Lemma 9.17] with Ω replaced by Ω_h. □

Chapter 6
Slice and hat elements

Let $\omega_h = \Omega \setminus \overline{\Omega}_h$. Hence,

$$\overline{\Omega} = \overline{\Omega}_h \cup \overline{\omega}_h.$$

We will decompose the set $\overline{\omega}_h$ into special elements of two kinds. Let $K_1, ..., K_{\bar{I}}$ be those tetrahedra from $\mathcal{T}_h$ which have three vertices on $\partial\Omega$. Let

$$F_i = K_i \cap \partial\Omega_h, \ i = 1, ..., \bar{I},$$

i.e.,

$$\partial\Omega_h = \bigcup_{i=1}^{\bar{I}} F_i.$$

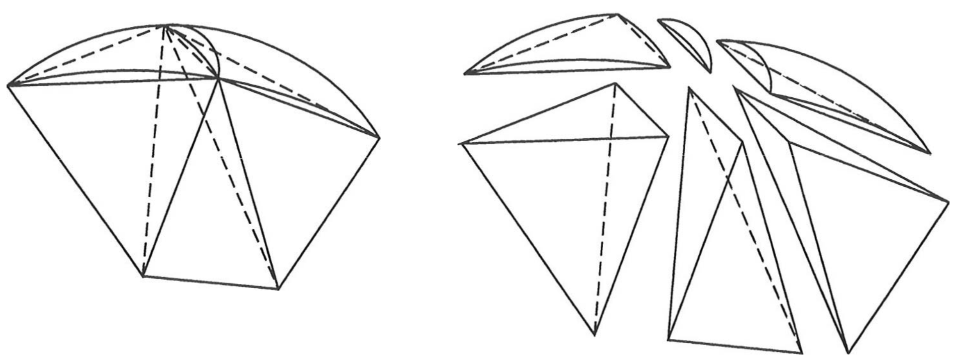

Figure 6.1.1.

Let $F_1, ..., F_I$, $I \leq \bar{I}$, be only those faces such that $F_i \not\subset \partial\Omega$, $i = 1, ..., I$. If Ω is strictly convex then $I = \bar{I}$. However, if $\partial\Omega$ contains a part of a plane then we can get $I < \bar{I}$.

Let $K_i^{\perp}$ be a tetrahedron symmetric to K_i with respect to the plane containing the face F_i, i.e., $F_i = K_i \cap K_i^{\perp}$. (Note that $K_i^{\perp} \notin \mathcal{T}_h$.) Define

$$H_i = K_i^{\perp} \cap \overline{\Omega} \quad \text{for} \quad i = 1, ..., I.$$

Such a set will be called a *hat element* because it looks like a hat on the element K_i (see Figures 6.1.1 and 6.1.2). Since Ω_h is convex, the interiors of all hat elements are mutually disjoint.

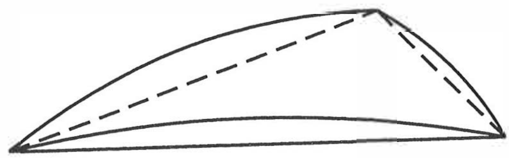

Figure 6.1.2.

Let $S_1, ..., S_J$ be the closures of components of the set $\omega_h \setminus \bigcup_{i=1}^{I} H_i$. Any such set S_j, $j = 1, ..., J$, will be called a *slice element* because it looks like an orange slice (see Figures 6.1.1 and 6.1.3).

Thus, we can decompose $\overline{\omega}_h$ into hat and slice elements, i.e.,

$$\overline{\omega}_h = (\bigcup_{i=1}^{I} H_i) \cup (\bigcup_{j=1}^{J} S_j).$$

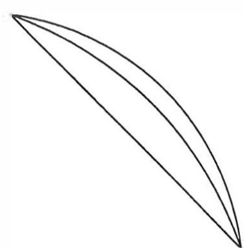

Figure 6.1.3.

Remark 6.1.1. We introduce a simple example, which shows that the slice element need not occur between two neighbouring $K_i^{\perp}$ and $K_j^{\perp}$ even if Ω is a strictly convex domain. Let Ω be the unit ball and let c be a small circle on its surface $\partial\Omega$. Figure 6.1.4 shows a position of four elements K_1, K_2, K_3, K_4, each of which has three vertices on c and the fourth vertex is common for all these elements. For simplicity we marked only faces $F_1, ..., F_4$ in the figure. It is easy to see that the hat element H_1 is surrounded by three hat elements H_2, H_3, H_4 in a such way, that there are no slice elements between H_1 and H_i, $i = 2, 3, 4$.

Remark 6.1.2. The reason, why we define two kinds of elements to decompose $\overline{\omega}_h$, is that it helps us to extend continuous piecewise linear functions from Ω_h to the whole domain Ω so that the extended functions remain continuous and piecewise linear. If we would consider only one kind of elements (say, "hat" elements) then we could not get the continuity of extended functions between adjacent "hat" elements, since any linear function in $\mathbf{R}^3$ is uniquely determined by values only at four points which do not belong to one plane. Three of these values are given

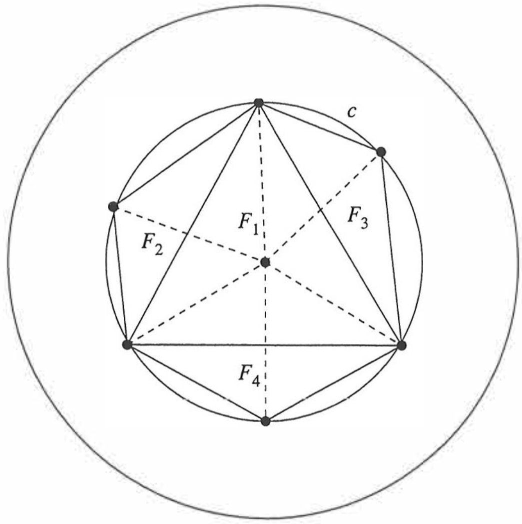

Figure 6.1.4.

at vertices of F_i, $i \in \{1, ..., I\}$. The fourth value in each "hat" element cannot guarantee the required continuity over the whole domain Ω.

Lemma 6.1.3. *There exists a positive constant C independent of h such that*

$$\text{meas}\, \omega_h \leq Ch^2, \quad h \in (0, h_0). \tag{6.1.1}$$

Proof. We will present only a short sketch of the proof. There exists a finite number of overlapping parts of $\partial\Omega$ such that each part is a graph of a C^2-function in some coordinate system. In the same system, a part of $\partial\Omega_h$ represents a graph of a continuous and piecewise linear function which interpolates the part of $\partial\Omega$. Using interpolation properties of linear elements in the C-norm, we come to

$$\max_{x \in \partial\Omega} \text{dist}(x, \partial\Omega_h) \leq Ch^2, \tag{6.1.2}$$

which proves the lemma. □

Remark 6.1.4. An arbitrarily short part of an edge of a convex domain need not lie in one plane, see, e.g., the intersection of two cylinders in Figure 6.1.5. Thus, it is not possible to decompose a neighbourhood of such an edge into elements (hat elements, slice elements, tetrahedral elements, ...) which have at most one face curved. It is obvious that any edge of one face curved elements belong to some plane. This is the reason why we do not consider $\overline{\Omega}$ with edges.

Remark 6.1.5. There are problems also with isolated vertices. Consider, for example, a domain, which looks like a drop (see Figure 6.1.6). Its boundary is, of

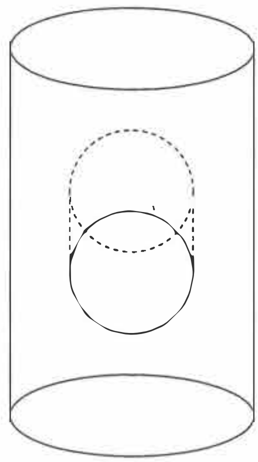

Figure 6.1.5.

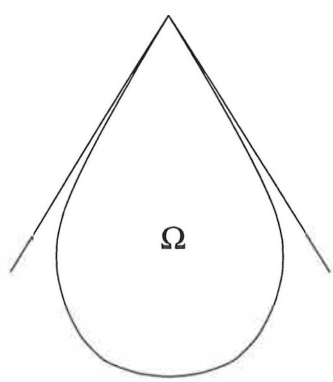

Figure 6.1.6.

course, non-differentiable in one point (vertex) only. It is easy to show that the second derivatives are unbounded near this point. Hence, we would not be able to prove Lemma 6.1.3 for a drop domain by the presented technique.

To see that the second derivatives are unbounded, we present the following simple example.

Example 6.1.6. Consider a "tangentional cone" to a given drop-domain (see Figure 6.1.7).

Choose the Cartesian coordinates (x_1, x_2, x_3) so that

$$x_1^2 + x_2^2 - \gamma^2 x_3^2 = 0 \quad \text{for } x_3 \le 0$$

is the equation of the tangentional cone (see Figure 6.1.7), where $\gamma > 0$ is a given constant.

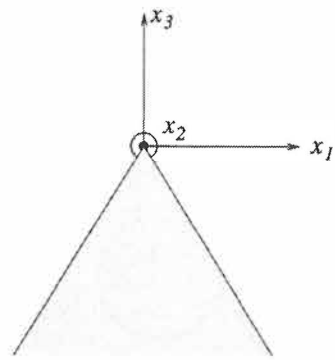

Figure 6.1.7.

Let $x_3 = -\varepsilon$, where ε is some small number, then

$$x_1^2 + x_2^2 = \gamma^2\varepsilon^2.$$

If $x_2 \geq 0$ then

$$x_2 = f(x_1) \equiv (\gamma^2\varepsilon^2 - x_1^2)^{1/2}.$$

Obviously,

$$f'(x_1) = -x_1(\gamma^2\varepsilon^2 - x_1^2)^{-1/2}$$

and

$$f''(x_1) = -(\gamma^2\varepsilon^2 - x_1^2)^{-1/2} - x_1^2(\gamma^2\varepsilon^2 - x_1^2)^{-3/2}.$$

Then

$$f''(0) = -1/\gamma\varepsilon \to \infty \quad \text{as} \quad \varepsilon \to 0.$$

Chapter 7
Extension of the finite element approximations to the whole domain

Using the terminology of [Feistauer, Ženíšek, 1987], we first define an analogue of "associated function" under our assumption on the given three-dimensional convex domain Ω: if $v_h \in V_h$ then we extend it by zero in $\Omega \setminus \overline{\Omega}_h$ and the resulting function (defined in Ω) is denoted as $\hat{v}_h$ and called the *associated function* with v_h. A situation with an analogue of "natural extension" (see also [Feistauer, Ženíšek, 1987] for terminology) is much more complicated in the three-dimensional case. As we shall see later we have to extend functions from X_h to hat and slice elements being parts of $\overline{\omega}_h$.

Note that we must also provide a continuity of extended functions. This is done "automatically" in two-dimensional case, but in the three-dimensional case we meet several obstacles to guarantee the continuity over $\overline{\Omega}$.

7.1. Natural extension

The *natural extension* v_h^* of $v_h \in X_h$ will be a piecewise linear function which is continuous on $\overline{\Omega}$ such that $v_h^*|_{\Omega_h} = v_h$. If $v_h \in X_h$ then v_h^* on each hat element H_i is defined by $v_h|_{K_i}$ using the symmetry with regard to the face F_i, $i = 1, ..., I$, i.e., for any $x^\perp \in H_i$ we set $v_h^*(x^\perp) = v_h(x)$, where x is the mirror image of $x^\perp$ with respect to F_i.

Before we define v_h^* on slice elements, we mention the following property of partitions of $\overline{\Omega}_h$, which is provided by the property of the strong regularity, see Definition 2.7.4. There exists a constant $\theta_0 > 0$ such that

$$0 < \theta_0 \leq \theta, \tag{7.1.1}$$

where θ is the angle between any two faces of any element K_i, $i = 1, ..., \bar{I}$.

Consider two "neighbouring" elements K_{i_1} and K_{i_2}, $i_1, i_2 \leq I$. For simplicity we write 1 and 2 instead of i_1 and i_2 in what follows. There may occur two different situations.

Case (i). Let K_1 and K_2 have a common face which we denote by F_{12}. Consider the profile being the result of the following procedure. Let $l_{12} = F_1 \cap F_2$ (it is a straight line segment with the end points on $\partial\Omega$). We cut the elements K_1 and K_2 (and H_1, H_2) by the plane orthogonal to l_{12} and passing through the middle point of l_{12}. A sketch of such a profile is presented in Figure 7.1.1.

Let φ be the angle formed by those faces of $K_1^\perp$, $K_2^\perp$ which are images of F_{12} (denoted by F_{12}^1, F_{12}^2) after symmetrical reflections with respect to the faces F_1 and F_2. They have l_{12} as a common edge and α_i, $i = 1, 2$, are the angles between F_i and F_{12}, $i = 1, 2$.

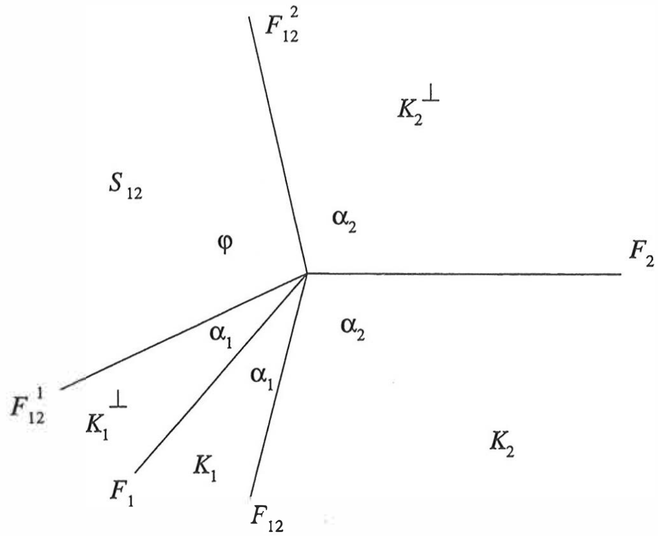

Figure 7.1.1.

We will show how to extend a function $v_h \in X_h$ on slice elements. We suppose that the angle φ satisfies the condition

$$0 < \varphi \leq \pi - 2\theta_0. \tag{7.1.2}$$

When h is small enough then (7.1.2) is, obviously, valid. Denote by S_{12} the slice element between H_1 and H_2, if it exists.

Vertices of the triangles F_{12}^1 and F_{12}^2 form a tetrahedron, which lies between $K_1^{\perp}$ and $K_2^{\perp}$. At those vertices, which are the end points of l_{12}, the values of the extended function v_h^* are equal to the values of v_h at these points. At other two vertices, the values of v_h^* are the same and equal to the value of v_h at that vertex of the triangle F_{12}, which does not belong to l_{12}. Obviously, values at those four points define uniquely v_h^* on the slice element S_{12}.

Case (ii). Let K_1 and K_2 have only one common edge l_{12}. Suppose that there are M tetrahedra $K^1, ..., K^M \in \mathcal{T}_h$, $M \geq 1$, between K_1 and K_2, which have l_{12} as a common edge. We make the profile in a similar manner as in Figure 7.1.1, see Figure 7.1.2. Here, by $\beta_1, ..., \beta_M$ we denote the angles formed by the faces of $K^1, ..., K^M$ having l_{12} as the common edge. By α_1 (α_2) we denote the angle between two faces of K_1 (K_2), which have l_{12} as a common edge.

Let φ be the angle described in the same manner as for case (i). We cut φ into M angles $\varphi_1, ..., \varphi_M$ such that

$$\varphi = \sum_{j=1}^{M} \varphi_j \qquad \text{and} \qquad \frac{\varphi_1}{\beta_1} = ... = \frac{\varphi_M}{\beta_M}.$$

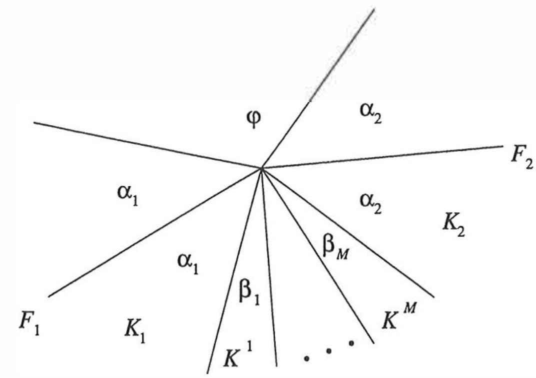

Figure 7.1.2.

First, let $M = 1$, i.e., we have only one tetrahedron K^1 between K_1 and K_2. Again suppose that $\varphi = \varphi_1 \leq \pi - 2\theta_0$ and define v_h^* in the following manner. Values of v_h^* at those vertices, which are the end points of l_{12} are equal to values of v_h at these points.

Further, we denote by F_1^1 and F_2^1 the faces common for K^1, K_1 and K^1, K_2, respectively. Let $F_1^{1,\perp}$ and $F_2^{1,\perp}$ be the faces-images of F_1^1 and F_2^1, respectively, after "reflection procedure".

Then, the values of v_h^* at those vertices of $F_1^{1,\perp}$ and $F_2^{1,\perp}$ which do not belong to l_{12} are equal to the values of v_h at the corresponding vertices of $F_1^{1,\perp}$ and $F_2^{1,\perp}$ (which do not belong to l_{12} as well).

These four given values define, obviously, v_h^* on the slice element S_{12}.

The case $M \geq 2$ is only technically more complicated.

7.2. Basic estimates

In this section we establish basic estimates for v_h^*. First, prove a simple auxiliary result.

Lemma 7.2.1. *Let $ABCD$ be a tetrahedron. Then there exists such a vector $\vec{l} \in \mathbf{R}^3$, $\|l\| = 1$, that the projections of all points of the tetrahedron along the vector $\vec{l}$ into the plane containing the triangle ABC belong to this triangle.*

Proof. Take some point $O \in \triangle ABC$ and consider the vector $\vec{DO}$. Then we may choose $\vec{l}$ as $\dfrac{\vec{DO}}{\|\vec{DO}\|}$. □

Lemma 7.2.2. *There exists a constant C such that*

$$\sum_{i=1}^{I} \|v_h^*\|_{1,H_i}^2 \leq Ch\|v_h\|_{1,\Omega_h}^2 \quad \forall v_h \in X_h. \tag{7.2.1}$$

Proof. Consider an element K_i, $i = 1, ..., I$, and the corresponding hat element $H_i = K_i^\perp \cap \overline{\Omega}$. Let $\vec{l} = (l_1, l_2, l_3)$ be a vector such that all points of $K_i^\perp$ (and of H_i) can be projected along it into F_i.

Choose the local Cartesian coordinate system $(y_i^1,\ y_i^2,\ y_i^3)$ such that $y_i^1,\ y_i^2$ are in F_i and y_i^3 is along the outward normal to F_i. Recall that the extended function v_h^* is linear in H_i. If $y = (y_1, y_2, y_3) \in H_i$ and $y^0 = (y_1^0, y_2^0, 0)$ is a projection of y into F_i along the vector $\vec{l}$ then

$$v_h^*(y) = v_h^*(y^0) + \int_{y^0}^{y} \frac{\partial v_h^*}{\partial \vec{l}}\, dt. \tag{7.2.2}$$

Since $\left.\dfrac{\partial v_h^*}{\partial \vec{l}}\right|_{K_i^\perp}$ is constant, we have

$$|v_h^*(y)|^2 = \left| v_h^*(y^0) + \int_{y^0}^{y} \frac{\partial v_h^*}{\partial \vec{l}}\, dt \right|^2 \leq 2\left\{ |v_h^*(y^0)|^2 + \|y - y^0\|^2 \left| \left.\frac{\partial v_h^*}{\partial \vec{l}}\right|_{K_i^\perp} \right|^2 \right\}. \tag{7.2.3}$$

As $\|y - y^0\| \leq Ch$ for $y^0, y \in K_i^\perp$, we get

$$|v_h^*(y)|^2 \leq 2|v_h^*(y^0)|^2 + 2C^2h^2 \left| \left.\frac{\partial v_h^*}{\partial \vec{l}}\right|_{K_i^\perp} \right|^2. \tag{7.2.4}$$

Therefore, from (7.2.4) we come to

$$\begin{aligned} \|v_h^*\|_{0,H_i}^2 &= \int_{H_i} |v_h^*(y)|^2\, dy \leq 2 \int_{H_i} |v_h^*(y^0)|^2\, dy + 2C^2h^2 \int_{H_i} \left| \frac{\partial v_h^*}{\partial \vec{l}} \right|^2 dy \\ &\leq C_2 h^2 \|v_h^*\|_{0,F_i}^2 + C_3 h^2 \|\operatorname{grad} v_h^*\|_{0,H_i}^2 \end{aligned}$$

and thus we get

$$\|v_h^*\|_{0,H_i}^2 \leq C_4 h^2 \left(\|v_h\|_{0,F_i}^2 + \|v_h\|_{1,K_i}^2 \right). \tag{7.2.5}$$

In a similar manner we derive

$$\|\operatorname{grad} v_h^*\|_{0,H_i}^2 \leq C_5 h \|\operatorname{grad} v_h\|_{0,K_i}^2 \leq Ch\|v_h\|_{1,K_i}^2. \tag{7.2.6}$$

To proceed our proof we use a discrete analogue of the trace theorem (7.2.7), which is given in Theorem 7.2.3. Having (7.2.5), (7.2.6), taking sum over all H_i and using (7.2.7), we derive easily (7.2.1). □

Theorem 7.2.3. (*The trace theorem in the spaces $W_2^1(\Omega_h)$*). *There exists a constant C such that*

$$\|v\|_{0,\partial\Omega_h} \leq C\|v\|_{1,\Omega_h} \quad \forall v \in W_2^1(\Omega_h) \quad \forall h \in (0, h_0). \tag{7.2.7}$$

Proof. See the proof of Theorem 3.3.5 from [Feistauer, 1987], which is based on the proof of the standard trace theorem (see, e.g., [Nečas, 1967]). □

Lemma 7.2.4. *There exists a constant $C > 0$ independent of h such that*

$$\sum_{j=1}^{J} \|v_h^*\|_{1,S_j}^2 \leq Ch^2 \|v_h\|_{1,\Omega_h}^2 \quad \forall v_h \in X_h. \tag{7.2.8}$$

Proof. For a given slice element S_j let H_{j_1} and H_{j_2} be its adjacent hat elements, associated to K_{j_1} and K_{j_2}. Let us write 1 and 2 instead of j_1 and j_2 in what follows.

Consider again two possible different cases:

(i) K_1 and K_2 have a common face F_{12},

(ii) K_1 and K_2 have only a common edge l_{12}.

Case (i). Consider the following local Cartesian coordinate system (x_1, x_2, x_3). The axis x_1 lies along l_{12}, the axis x_3 is perpendicular to x_1 and lies in the plane bisecting the angle φ and the axis x_2 is such that the system (x_1, x_2, x_3) is Cartesian.

Let F_{12}^1 and F_{12}^2 be defined as in Section 7.1 and let T denote the tetrahedron formed by vertices of F_{12}^1 and F_{12}^2 ($S_j \subset T$).

From interpolation properties in two-dimensional case, condition (7.1.2) and from C^2-smoothness of $\partial\Omega$ we observe the existence of such a positive constant C that

$$S_j \subset T \cap L, \tag{7.2.9}$$

where $L = \{(x_1, x_2, x_3) | \ 0 \leq x_3 \leq Ch^2\}$ is a layer of thickness $O(h^2)$.

Obviously, for case (i) we have

$$\partial_2 v_h^*|_T = 0. \tag{7.2.10}$$

Then (7.1.2), (7.2.9) and (7.2.10) imply

$$\|v_h^*\|_{0,S_j}^2 \leq Ch^2 \|v_h^*\|_{0,F_{12}^1 \cap L}^2. \tag{7.2.11}$$

There exists a unit vector $\vec{l}$ such that all points of $K_1^\perp$ can be projected along $\vec{l}$ into F_1. Let the point $x = (x_1, x_2, x_3)$ belong to $F_{12}^1 \cap L$ and $x^0 = (x_1^0, x_2^0, x_3^0)$ be the projection of x into F_1 along $\vec{l}$. Then we have

$$v_h^*(x) = v_h^*(x^0) + \int_{x^0}^{x} \frac{\partial v_h^*}{\partial \vec{l}}\, dt.$$

Further, (cf. (7.2.3)),

$$|v_h^*(x)|^2 \leq 2\left\{ |v_h^*(x^0)|^2 + \|x - x^0\|^2 \left| \frac{\partial v_h^*}{\partial \vec{l}}\bigg|_{K_1^\perp} \right|^2 \right\}.$$

Since there is a linear affine one-to-one mapping between x and x^0, we get by the substitution theorem and the above inequality that

$$\|v_h^*\|_{0,F_{12}^1\cap L}^2 \le C(\|v_h^*\|_{0,F_1}^2 + h^2\|\operatorname{grad} v_h^*\|_{0,H_1}^2). \tag{7.2.12}$$

From the symmetry with respect to F_1, (7.2.11) and (7.2.12) we get

$$\|v_h^*\|_{0,S_j}^2 \le Ch^2(\|v_h\|_{0,F_1}^2 + h^2\|v_h\|_{1,K_1}^2). \tag{7.2.13}$$

Now we deal with $\operatorname{grad} v_h^*$ over S_j. For its first component we have

$$\partial_1 v_h^*|_{S_j} = \partial_1 v_h|_{K_1},$$

because v_h^* is linear continuous on S_j, and l_{12} lies in the axis x_1. The second component is zero in view of (7.2.10). So let us examine the third component.

Consider a unit vector $\vec{\nu}$ parallel to the segment $F_{12}^1 \cap \{x_1 = 0\}$. Then

$$\partial_3 v_h^*|_{S_j} = \frac{\partial v_h^*}{\partial \vec{\nu}} / \cos(\varphi/2).$$

Condition (7.1.2) implies that $\cos(\varphi/2) \ge \cos(\pi/2 - \theta_0)$, i.e., $1/\cos(\varphi/2) \le 1/\sin\theta_0$, which leads to

$$|\partial_3 v_h^*|_{S_j}|^2 \le C\left|\frac{\partial v_h^*}{\partial \vec{\nu}}\right|^2 \le C_1\|\operatorname{grad} v_h|_{K_1}\|^2.$$

Hence, there exists a constant C such that

$$\|\operatorname{grad} v_h^*\|_{0,S_j}^2 \le Ch^2\|\operatorname{grad} v_h\|_{0,K_1}^2 \le Ch^2\|v_h\|_{1,K_1}^2. \tag{7.2.14}$$

Further, (7.2.14) and (7.2.13) imply

$$\|v_h^*\|_{1,S_j} \le Ch^2(\|v_h\|_{1,K_1}^2 + \|v_h\|_{0,F_1}^2). \tag{7.2.15}$$

Case (ii). Let $M = 1$ for simplicity, let us form the same coordinate system (x_1, x_2, x_3) as for case (i). We will also use denotations introduced for case (i).

Property (7.2.10) is violated, but (7.2.9) is valid. Then analogously to (7.2.11), using the Taylor formula, we get

$$\|v_h^*\|_{0,S_j}^2 \le Ch^2(\|v_h^*\|_{0,F_{12}^1\cap L}^2 + h^2\|\partial_2 v_h^*\|_{0,T\cap L}^2). \tag{7.2.16}$$

The term $\|v_h^*\|_{0,F_{12}^1\cap L}^2$ can be estimated as in case (i). And the only problem is to estimate the value $\partial_2 v_h^*|_T$. From geometrical considerations we have

$$\partial_2 v_h^*|_T = \frac{\sin(\beta_1/2)}{\sin(\varphi/2)} \frac{\partial v_h}{\partial \vec{\mu}}\bigg|_{K^1},$$

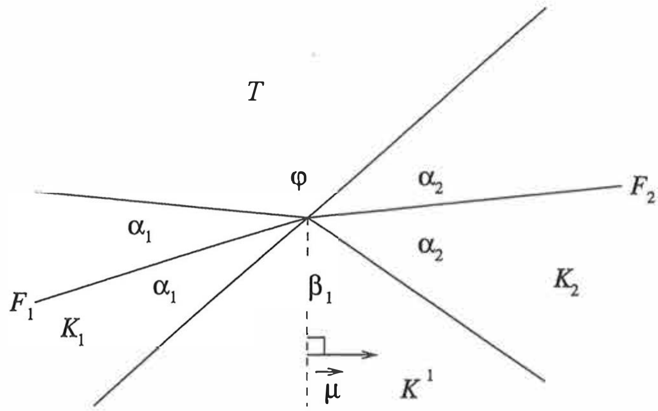

Figure 7.2.1.

where $K^1 \in \mathcal{T}_h$ is the tetrahedron between K_1 and K_2 and $\vec{\mu}$ is a unit vector in $\mathbf{R}^3$, which is perpendicular to the plane bisecting the angle β_1 (see Figure 7.2.1).

Obviously,

$$\theta_0 \le \beta_1 \le \varphi < \pi - 2\theta_0,$$

and thus

$$\frac{\sin(\beta_1/2)}{\sin(\varphi/2)} \le \frac{\cos\theta_0}{\sin(\theta_0/2)},$$

i.e.,

$$|\partial_2 v_h^*|_T|^2 \le C\|\operatorname{grad} v_h|_{K^1}\|^2. \tag{7.2.17}$$

Doing an analogous analysis to that one of case (i), we easily get the estimation similar to (7.2.15). The case $M > 2$ can be treated in a similar way. Having (7.2.15), (7.2.17), taking sum over all S_j and using (7.2.7), we get (7.2.8). □

Theorem 7.2.5. *There exists a constant C such that*

$$\|v_h^*\|_{1,\omega_h} \le Ch^{1/2}\|v_h\|_{1,\Omega_h} \quad \forall v_h \in X_h. \tag{7.2.18}$$

Proof. Obviously, (7.2.1) and (7.2.8) imply (7.2.18). □

Theorem 7.2.6. *There exists a constant C such that*

$$\|v_h^* - \hat{v}_h\|_{1,\Omega} \le Ch^{1/2}\|v_h\|_{1,\Omega_h} \quad \forall v_h \in V_h, \tag{7.2.19}$$

$$|v_h^*|_{1,q,\omega_h} \le Ch^{1/q}|v_h|_{1,q,\Omega_h} \quad \forall v_h \in X_h \quad \forall h \in (0, h_0) \quad \forall q \in [1,\infty]. \tag{7.2.20}$$

Proof. Formula (7.2.19) is a direct consequence of Theorem 7.2.5 and the definition of $\hat{v}_h$.

Using (7.2.6) and (7.2.14), we see that (7.2.20) holds, since v_h^* is piecewise linear. □

Lemma 7.2.7. *The following relations hold*

$$\|w - \pi_h w\|_{1,\Omega_h} \to 0 \quad as\ h \to 0, \tag{7.2.21}$$

$$\|\pi_h w\|_{1,q,\Omega_h} + |(\pi_h w)^*|_{1,q,\Omega} \leq C(w,q) \quad \forall w \in W_q^2(\Omega) \quad \forall h \in (0,h_0) \quad \forall q \in (\tfrac{3}{2}, \infty]. \tag{7.2.22}$$

Proof. By the Sobolev imbedding theorem $w \in W_q^2(\Omega)$ is continuous. Hence, from the standard interpolation theorem (see [Ciarlet, Theorem 3.1.6]) we have

$$\|w - \pi_h w\|_{1,\Omega_h} \leq Ch\|w\|_{2,\Omega_h} \leq Ch\|w\|_{2,\Omega} \to 0 \quad \text{as } h \to 0,$$

i.e., (7.2.21) is valid.

Further, by (7.2.20),

$$|(\pi_h w)^*|_{1,q,\omega_h} \leq C_1 h^{1/q} |\pi_h w|_{1,q,\Omega_h}.$$

From here and [Ciarlet, Theorem 3.1.6], we see that

$$\|\pi_h w\|_{1,q,\Omega_h} + |(\pi_h w)^*|_{1,q,\Omega} \leq C_2 \|\pi_h w\|_{1,q,\Omega_h}$$
$$C_2(\|w - \pi_h w\|_{1,q,\Omega_h} + \|w\|_{1,q,\Omega_h}) \leq C_3(h\|w\|_{1,q,\Omega_h} + \|w\|_{1,q,\Omega}) \leq C_4\|w\|_{2,q,\Omega}.$$

□

Lemma 7.2.8. *The following relations hold*

$$\|v_h^*\|_{1,\Omega} + \|\hat{v}_h\|_{1,\Omega} \leq C\|v_h\|_{1,\Omega_h} \quad \forall v_h \in V_h \quad \forall h \in (0,h_0), \tag{7.2.23}$$

$$\|w\|_{1,\omega_h} \to 0 \quad as \quad h \to 0 \quad \forall w \in W_2^1(\Omega). \tag{7.2.24}$$

Proof. Estimation (7.2.23) is a consequence of (7.2.20) and the definition of $\hat{v}_h$. Estimation (7.2.24) follows from (6.1.1) and the absolute continuity of the Lebesgue integral. □

Lemma 7.2.9. *The following relations hold*

$$\|\overline{u}_h - \overline{u}\|_{1,\Omega_h} + \|\overline{u}_h^* - \overline{u}\|_{1,\Omega} \to 0 \quad as\ h \to 0, \tag{7.2.25}$$

$$\|\overline{u}_h\|_{1,\Omega_h} + \|\overline{u}_h^*\|_{1,\Omega} \leq C, \tag{7.2.26}$$

$$\|\overline{u}_h\|_{1,q,\Omega_h} + |\overline{u}_h^*|_{1,q,\Omega} \leq C(q) \quad \forall h \in (0,h_0) \quad \forall q \in [1,p]. \tag{7.2.27}$$

Proof. The first term in (7.2.25) tends to 0 as $h \to 0$ in view of (4.2.14). Further, $\|\overline{u}_h^* - \overline{u}\|_{1,\Omega}^2 = \|\overline{u}_h - \overline{u}\|_{1,\Omega_h}^2 + \|\overline{u}_h^* - \overline{u}\|_{1,\omega_h}^2$. The first term in the right-hand side of the above inequality tends to zero (see (4.2.14)). For the last term we have

$$\|\overline{u}_h^* - \overline{u}\|_{1,\omega_h} \leq \|\overline{u}_h^*\|_{1,\omega_h} + \|\overline{u}\|_{1,\omega_h}.$$

From (7.2.18) and the absolute continuity of the Lebesgue integral (cf. (7.2.24)) both terms in the right-hand side of the above inequality tends to zero. Thus, (7.2.25) is proved. Also, from (7.2.18) and (4.2.15) we get easily (7.2.26). The proof of (7.2.27) is similar to the proof of Lemma 3.1.3 from [Feistauer, Sobotíková, 1990]. □

Chapter 8
Main convergence results

In this final chapter of Part II we present the main convergence results: first, in Theorem 8.1.2 we prove the weak convergence, and, then, in Theorem 8.2.1, we prove the strong convergence of the approximate solutions to the exact (weak) one.

Note that these theorems generalize Theorems 9.24 and 9.25 from [Křížek, Neittaanmäki, 1996] to the three-dimensional space.

8.1. Weak convergence

For each $h \in (0, h_0)$ we define a function $u'_h \in W_2^1(\Omega)$ corresponding to the approximate solution u_h as follows: according to Definition 5.1.1 we express u_h in the following form:

$$u_h = \overline{u}_h + z_h$$

with $z_h \in V_h$. Then

$$u'_h = \overline{u}_h^* + \hat{z}_h, \tag{8.1.1}$$

where $\overline{u}_h^*$ is the natural extension of $\overline{u}_h$ and $\hat{z}_h$ is the associated function.

Lemma 8.1.1. *There exists a constant C such that*

$$\|z_h\|_{1,\Omega_h} + \|u'_h\|_{1,\Omega} + \|u_h^*\|_{1,\Omega} + \|\hat{z}_h\|_{1,\Omega} + \|z_h^*\|_{1,\Omega} \leq C. \tag{8.1.2}$$

Proof. The proof is trivial and similar to [Feistauer, Sobotíková, 1990, Lemma 3.1.6]. □

Let Ω_0 be an open subset in Ω. We put

$$\tilde{a}_{\Omega_0}(y; w, v) = (\mathcal{A}(\cdot, y) \operatorname{grad} w, \operatorname{grad} v)_{0,\Omega_0}, \quad y, w, v \in W_2^1(\Omega_0). \tag{8.1.3}$$

From the boundedness of $\mathcal{A}$ in $\Omega \times \mathbf{R}^1$ we, obviously have

$$|\tilde{a}_{\Omega_0}(y; w, v)| \leq C\|w\|_{1,\Omega_0}\|v\|_{1,\Omega_0} \quad \forall y, w, v \in W_2^1(\Omega_0). \tag{8.1.4}$$

Theorem 8.1.2. *If $u_h = \overline{u}_h + z_h$, $z_h \in V_h$, is an approximate solution from Definition 5.1.1 and u'_h is the function corresponding to u_h by (8.1.1), then*

$$u'_h \rightharpoonup u \quad \textit{(weakly) in } W_2^1(\Omega) \quad \textit{as } h \to 0,$$

where $u \in W_2^1$ is the weak solution of the problem (4.2.1)–(4.2.2).

Proof. From (8.1.2) we have $\|u'_h\|_{1,\Omega} \leq C$ for all $h \in (0, h_0)$. Since the functions u'_h belong to $W_2^1(\Omega)$, then, as a consequence of the Eberlein–Schmulyan Theorem

1.4.10, there exists an element $u \in W_2^1(\Omega)$ and a subsequence $\{u'_{\hat{h}}\} \subset \{u'_h\}$ such that $u'_{\hat{h}} \rightharpoonup u$ as $\hat{h} \to 0$. Hence, we must prove, in fact, that any weak cluster point of the family $\{u'_h\}$, $h \in (0, h_0)$, is the weak solution of the problem (4.2.1)–(4.2.2).

Let u be one of such cluster points. We denote the corresponding subsequence $\{u'_{\hat{h}}\}$ convergent to u as $\{u'_h\}$ for simplicity, i.e.,

$$u'_h \rightharpoonup u \quad \text{(weakly) in } W_2^1(\Omega) \text{ as } h \to 0. \tag{8.1.5}$$

We will show that u is the solution of (4.2.1)–(4.2.2).

Note that $V = \overset{\circ}{W}{}_2^1(\Omega)$ is weakly closed and the sequence $\hat{z}_h = u'_h - \overline{u}_h^*$ belongs to V. It means that the weak limit of $\{\hat{z}_h\}$ also belongs to V. From (8.1.5) $u'_h \rightharpoonup u$ as $h \to 0$ and from (7.2.25) we have $\overline{u}_h^* \to \overline{u}$ as $h \to 0$, it implies $\overline{u}_h^* \rightharpoonup \overline{u}$ as $h \to 0$. Hence, $u - \overline{u}$ is a weak limit of $\hat{z}_h$ as $h \to 0$ and, therefore, belongs to V.

To prove the condition $a(u; u, v) = F(v)$ for all $v \in V$, we first consider an arbitrary $v \in V \cap C^\infty(\overline{\Omega})$. Let $v_h = \pi_h v \in V_h$. Then, by (7.2.18),

$$\begin{aligned} \|v - v_h^*\|_{1,\Omega}^2 &= \|v - v_h^*\|_{1,\Omega_h}^2 + \|v - v_h^*\|_{1,\omega_h}^2 \\ &\le C(h\|v\|_{2,\Omega}^2 + \|v\|_{1,\omega_h}^2 + \|v_h^*\|_{1,\omega_h}^2) \\ &\le C(h\|v\|_{2,\Omega}^2 + \|v\|_{1,\infty,\Omega}^2 \text{meas } \omega_h + h\|v_h\|_{2,\Omega_h}^2). \end{aligned}$$

From here, (6.1.1), (7.2.21) and (7.2.22) we get

$$\begin{aligned} \|v - v_h^*\|_{1,\Omega} &\to 0 && \text{as } h \to 0, \\ \|v_h\|_{1,q,\Omega_h} &\le C(v, q), && q \in (\tfrac{3}{2}, \infty]. \end{aligned} \tag{8.1.6}$$

From the fact that $u_h^* = \overline{u}_h^* + z_h^*$, the definition of u'_h and (7.2.19) we obtain

$$\|u_h^* - u'_h\|_{1,\Omega} = \|z_h^* - \hat{z}_h\|_{1,\Omega} \le Ch^{1/2}\|z_h\|_{1,\Omega_h}. \tag{8.1.7}$$

Estimates (8.1.7) and (8.1.2) imply that

$$\|u_h^* - u'_h\|_{1,\Omega} \to 0 \quad \text{as } h \to 0. \tag{8.1.8}$$

Further, from (8.1.8) and (8.1.5) we easily find that

$$u_h^* \rightharpoonup u \quad \text{in } W_2^1(\Omega) \text{ as } h \to 0. \tag{8.1.9}$$

Using (8.1.9) and the compact imbedding $W_2^1(\Omega) \hookrightarrow L^2(\Omega)$, see the Kondrašov Theorem 1.4.7, we get the strong convergence

$$u_h^* \to u \quad \text{in } L^2(\Omega). \tag{8.1.10}$$

Obviously,

$$\begin{aligned} a(u;u,v) - F(v) &= [a(u;u-u_h^*,v)] + [a(u;u_h^*,v) - a(u_h^*;u_h^*,v)] \\ &\quad + [a(u_h^*;u_h^*,v-v_h^*)] + [\tilde{a}_{\omega_h}(u_h^*;u_h^*,v_h^*)] \\ &\quad + [\tilde{a}_h(u_h;u_h,v_h) - a_h(u_h;u_h,v_h)] + [F_h(v_h) - \tilde{F}_h(v_h)] \\ &\quad + [F(v_h^*) - F(v)] + \left[- \int_{\omega_h} f v_h^* \, dx \right] \\ &\equiv I_1 + \dots + I_8. \end{aligned}$$

The terms $I_1, \dots, I_8$ can be estimated for $h \to 0$ in the following manner:

1. The mapping $y \in W_2^1(\Omega) \mapsto a(u,y,v) \in \mathbf{R}^1$ is a continuous linear functional (see (4.2.10)). Thus, u_h^* converges to u weakly in $W_2^1(\Omega)$ and $I_1 \to 0$ as $h \to 0$.
2. By (4.2.11), (8.1.2) and (8.1.10),
$$|I_2| \le C\|u - u_h^*\|_{0,\Omega}\|u_h^*\|_{1,\Omega}\|v\|_{1,\infty,\Omega} \le C\|u - u_h^*\|_{0,\Omega}\|v\|_{1,\infty,\Omega} \to 0.$$
3. In view of (4.2.10), (8.1.2) and (8.1.6),
$$|I_3| \le C\|u_h^*\|_{1,\Omega}\|v - v_h^*\|_{1,\Omega} \le C\|v - v_h^*\|_{1,\Omega} \to 0.$$
4. It follows from (8.1.4), (7.2.18) and (8.1.6) that
$$|I_4| \le C\|u_h^*\|_{1,\omega_h}\|v_h^*\|_{1,\omega_h} \le Ch^2\|u_h\|_{1,\Omega_h}\|v_h\|_{1,\Omega_h} \le C_1(v,2)h^2 \to 0.$$
5. By (5.4.1), (5.3.1) and (8.1.6),
$$|I_5| \le Ch\|u_h\|_{1,\Omega_h}\|v_h\|_{1,\Omega_h} + Ch^{1-3/p}\|u_h\|_{1,\Omega_h}^2\|v_h\|_{1,p,\Omega_h} \to 0.$$
6. We have $|I_6| \to 0$ in virtue of (5.2.6) and (8.1.6).
7. The Cauchy–Schwarz–Buniakovskii inequality and (8.1.6) imply
$$|I_7| = |F(v) - F(v_h^*)| = |(f, v - v_h^*)_0| \le \|f\|_{0,\Omega}\|v - v_h^*\|_{0,\Omega} \le C\|v - v_h^*\|_{1,\Omega} \to 0.$$
8. By the Cauchy–Schwarz–Buniakovskii inequality, (7.2.18) and (8.1.6),
$$|I_8| \le \int_{\omega_h} |f v_h^*| \, dx \le \|f\|_{0,\Omega}\|v_h^*\|_{0,\omega_h} \le Ch\|v_h\|_{1,\Omega_h} \to 0.$$

As a result we see that u satisfies the identity

$$a(u;u,v) = F(v) \quad \forall v \in V \cap C^\infty(\overline{\Omega})$$

and, hence, by the density of $V \cap C^\infty(\overline{\Omega})$ in V and the continuity of $a(u;u,\cdot)$ and $F(\cdot)$ we get $a(u;v,v) = F(v)$ for all $v \in V$. By [Hlaváček, Křížek, Malý], u is unique and, thus, the whole sequence u_h' converges weakly to u, i.e., the theorem is proved. □

8.2. Strong convergence

Now we will establish the strong convergence.

Theorem 8.2.1. *We have $u_h' \to u$ (strongly) in $W_2^1(\Omega)$ as $h \to 0$ and*

$$\lim_{h\to 0} \|u - u_h\|_{1,\Omega_h} = 0.$$

Proof. We write again $u = \overline{u} + z$, $z \in V$ and $u_h = \overline{u}_h + z_h$, $z_h \in V_h$. Let us consider a sequence $\{z^m\}_{m=1}^{\infty} \subset V \cap C^\infty(\Omega)$ such that

$$\|z - z^m\|_{1,\Omega} \to 0 \quad \text{as } m \to \infty. \tag{8.2.1}$$

Let us set

$$z_h^m = \pi_h z^m.$$

Thus, $z_h^m \in V_h$ and from (7.2.21) and (7.2.22) we have

$$\begin{aligned} &\|z^m - z_h^m\|_{1,\Omega_h} \to 0 && \text{as } h \to 0, \\ &\|z_h^m\|_{1,q,\Omega_h} + |(z_h^m)^*|_{1,q,\Omega} \le C(z^m, q) && \forall h \in (0, h_0)\ \forall q \in (\tfrac{3}{2}, \infty]. \end{aligned} \tag{8.2.2}$$

Let us define

$$v_h^m = z_h - z_h^m \ (\in V_h).$$

By (5.2.1), (5.1.4) and the definition of a weak solution we obtain

$$\begin{aligned} C_1 \|v_h^m\|_{1,\Omega_h}^2 &\le a_h(u_h; v_h^m, v_h^m) \\ &= a_h(u_h; \overline{u}_h + z_h, v_h^m) - a_h(u_h; \overline{u}_h + z_h^m, v_h^m) \\ &= F_h(v_h^m) - a_h(u_h; \overline{u}_h + z_h^m, v_h^m) \\ &= [F_h(v_h^m) - \tilde{F}_h(v_h^m)] + \Big[-\int_{\omega_h} f (v_h^m)^*\, dx\Big] \\ &\quad + [F((v_h^m)^*) - F(\hat{v}_h^m)] + [a(u; u, \hat{v}_h^m) - a_h(u_h; \overline{u}_h + z_h^m, v_h^m)] \\ &\equiv J_1 + \ldots + J_4. \end{aligned} \tag{8.2.3}$$

Further, we can write

$$\begin{aligned} J_4 &\equiv a(u; u, \hat{v}_h^m) - a_h(u_h; \overline{u}_h + z_h^m, v_h^m) \\ &= [a(u; \overline{u} + z - \overline{u}_h^* - (z_h^m)^*, \hat{v}_h^m)] \\ &\quad + [a(u; \overline{u}_h^* + (z_h^m)^*, \hat{v}_h^m) - a(u_h'; \overline{u}_h^* + (z_h^m)^*, \hat{v}_h^m)] \\ &\quad + [a(u_h'; \overline{u}_h^* + (z_h^m)^*, \hat{v}_h^m) - a(u_h'; \overline{u}_h^* + (z_h^m)^*, (v_h^m)^*)] \\ &\quad + [a(u_h'; \overline{u}_h^* + (z_h^m)^*, (v_h^m)^*) - a(u_h^*; \overline{u}_h^* + (z_h^m)^*, (v_h^m)^*)] \\ &\quad + [\tilde{a}_{\omega_h}(u_h^*; \overline{u}_h^* + (z_h^m)^*, (v_h^m)^*)] \\ &\quad + [\tilde{a}_h(u_h; \overline{u}_h + z_h^m, v_h^m) - a_h(u_h; \overline{u}_h + z_h^m, v_h^m)] \\ &\equiv D_1 + \ldots + D_6. \end{aligned} \tag{8.2.4}$$

We will estimate the individual terms J_1, J_2, J_3 and $D_1, ..., D_6$. By (5.2.6),

$$|J_1| \leq Ch\|f\|_{1,\infty,\Omega}\|v_h^m\|_{1,\Omega_h} \leq C'h\|v_h^m\|_{1,\Omega_h}. \tag{8.2.5}$$

The Cauchy–Schwarz inequality and the estimate (7.2.18) imply that

$$|J_2| \leq Ch^{1/2}\|v_h^m\|_{1,\Omega_h}. \tag{8.2.6}$$

Further, using the continuity of the functional F and (7.2.19), we get

$$|J_3| \leq C\|(v_h^m)^* - \hat{v}_h^m\|_{1,\Omega} \leq Ch^{1/2}\|v_h^m\|_{1,\Omega_h}. \tag{8.2.7}$$

In virtue of (4.2.10), the following inequalities

$$\|z - (z_h^m)^*\|_{1,\Omega} \leq \|z - z^m\|_{1,\Omega} + \|z^m - (z_h^m)^*\|_{1,\Omega},$$
$$\|z^m - (z_h^m)^*\|_{1,\Omega} \leq C(\|z^m - z_h^m\|_{1,\Omega_h} + \|z^m\|_{1,\omega_h} + \|(z_h^m)^*\|_{1,\omega_h}),$$

(7.2.18), (8.2.2) with $q = 2$ and (7.2.23), we have

$$\begin{aligned} |D_1| &= |a(u; \overline{u} - \overline{u}_h^* + z - z^m + z^m - z_h^m + z_h^m - (z_h^m)^*, \hat{v}_h^m)| \\ &\leq C(\|\overline{u} - \overline{u}_h^*\|_{1,\Omega} + \|z - z^m\|_{1,\Omega} + \|z^m - z_h^m\|_{1,\Omega_h} \\ &\quad + \|z^m\|_{1,\omega_h} + C(z^m, 2)h^{1/2})\|v_h^m\|_{1,\Omega_h}. \end{aligned} \tag{8.2.8}$$

Using (7.2.23), (7.2.26) and (8.2.2), we can show that

$$\begin{aligned} |D_2| &\leq C\|u - u_h'\|_{0,q,\Omega}|\overline{u}_h^* + (z_h^m)^*|_{1,p,\Omega}\|\hat{v}_h^m\|_{1,\Omega} \\ &\leq C_1\|u - u_h'\|_{0,q,\Omega}C(z^m, p)\|v_h^m\|_{1,\Omega_h} \end{aligned} \tag{8.2.9}$$

with $1/p + 1/q = 1/2$.

Estimates (4.2.10), (7.2.26), (8.2.2) and (7.2.19) imply that

$$|D_3| \leq C(z^m, 2)Ch^{1/2}\|v_h^m\|_{1,\Omega_h}. \tag{8.2.10}$$

The term D_4 can be estimated in a similar way as D_2:

$$|D_4| \leq C\|u_h' - u_h^*\|_{0,q,\Omega}|\overline{u}_h^* + (z_h^m)^*|_{1,p,\Omega}\|(v_h^m)^*\|_{1,\Omega}.$$

Now, using the compact imbedding $W_2^1(\Omega) \hookrightarrow L^q(\Omega)$ for $q \in [1, 6)$, see Kondrašov Theorem 1.4.7, (7.2.19), (8.1.2), (7.2.26), (8.2.2) and (7.2.23), we find that

$$|D_4| \leq C(z^m, p)Ch^{1/2}\|v_h^m\|_{1,\Omega_h}. \tag{8.2.11}$$

In order to estimate D_5, we use (8.1.4), (7.2.18), (7.2.26) and (8.2.2) with $q = 2$. Then

$$|D_5| \leq Ch^{1/2}\|v_h^m\|_{1,\Omega_h}. \tag{8.2.12}$$

Finally, by (5.4.1) with $\alpha = p$, $\beta = 2$, (5.3.1), (7.2.26) and (8.2.2)

$$\begin{aligned} |D_6| &\le C(h\|\overline{u}_h + z_h^m\|_{1,\Omega_h} + h^{1-3/p}\|u_h\|_{1,\Omega_h}\|\overline{u}_h + z_h^m\|_{1,p,\Omega_h})\|v_h^m\|_{1,\Omega_h} \\ &\le C_1(h + C(z^m,p)h^{1-3/p})\|v_h^m\|_{1,\Omega_h}. \end{aligned} \tag{8.2.13}$$

Now taking into account (8.2.3)–(8.2.13), we obtain the estimate

$$\begin{aligned} C_1\|v_h^m\|_{1,\Omega_h} \le C(&h^{1/2} + \|\overline{u} - \overline{u}_h^*\|_{1,\Omega} + \|z - z^m\|_{1,\Omega} + C(z^m,2)h^{1/2} \\ &+ \|z^m - z_h^m\|_{1,\Omega_h} + \|z^m\|_{1,\omega_h} \\ &+ C(z^m,p)\|u - u_h'\|_{0,q,\Omega} + C(z^m,p)h^{1/2} + C(z^m,p)h^{1-2/p}). \end{aligned}$$

Let us consider m fixed and pass to the limit for $h \to 0$. Using (8.2.2), (7.2.24) with $w = z^m$, (7.2.25), (8.1.5) and the compact imbedding $W_2^1(\Omega) \hookrightarrow L^q(\Omega)$, see the Kondrašov Theorem 1.4.7, which imply that $u_h' \to u$ in $L^q(\Omega)$, $q \in [1,6)$, as $h \to 0$, we conclude that

$$\limsup_{h\to 0} \|v_h^m\|_{1,\Omega_h} \le C\|z - z^m\|_1, \quad m = 1,2,...,$$

where C is independent of m.

Further,

$$\begin{aligned} \|u - u_h\|_{1,\Omega_h} &= \|(\overline{u} + z) - (\overline{u}_h + z_h)\|_{1,\Omega_h} \\ &\le \|\overline{u} - \overline{u}_h\|_{1,\Omega_h} + \|z - z^m\|_{1,\Omega_h} + \|z^m - z_h^m\|_{1,\Omega_h} + \|z_h^m - z_h\|_{1,\Omega_h}. \end{aligned}$$

Taking into account that $z_h - z_h^m = v_h^m$, and passing to the limit for $h \to 0$, we obtain by (7.2.25) and (8.2.2) that

$$\limsup_{h\to 0} \|u - u_h\|_{1,\Omega_h} \le C\|z - z^m\|_{1,\Omega}, \quad m = 1,2,\dots .$$

Passing to the limit for $m \to \infty$ and using (8.2.1), we find that

$$\lim_{h\to 0} \|u - u_h\|_{1,\Omega_h} = 0.$$

The strong convergence $u_h' \to u$ in $W_2^1(\Omega)$ is a consequence of (4.2.15), (7.2.19), (7.2.18), (7.2.24) and (8.1.1). □

PART III

Equilibrium Finite Elements in Three-Dimensional Space

On contents of Part III

Part III consists of one chapter – Chapter 9.

The main result can be formulated as it follows: the space of divergence-free functions with vanishing normal flux on the boundary is approximated by subspaces of finite elements that have the same property. The easiest way of generating basis functions in these subspaces is considered.

This chapter is based on the published work [Korotov, 1997a].

Chapter 9
Equilibrium finite elements

9.1. Introduction

The main result is a construction of finite element subspaces of the spaces of divergence-free functions. Such a problem is frequently met when we treat numerically some phenomena in continuum mechanics, electromagnetism, heat and fluid flow problems, etc.

Namely, we shall describe an internal finite element approximation of the following space which appears in variational formulations of a considerable number of problems, see, e.g., [Girault, Raviart], [Hlaváček, Křížek, 1984], [Křížek, Neittaanmäki, 1990], [Nedelec], [Temam]:

$$H_0(\operatorname{div}^0;\Omega) = \{\vec{q} \in [L^2(\Omega)]^d \mid (\vec{q}, \nabla z)_0 = 0 \ \forall z \in H^1(\Omega)\}, \quad d = 2, 3. \tag{9.1.1}$$

We will deal only with the three-dimensional case: $\Omega \subset \mathbf{R}^3$ is a bounded domain with a Lipschitz continuous boundary, $(\cdot\,,\cdot)_0$ is the inner product in $[L_2(\Omega)]^l$, $l = 1, 2, 3$, $H^k(\Omega)$ is the standard Sobolev space with the norm $\|\cdot\|_k$ and $\vec{l} \cdot \vec{w}$ is the standard inner product of vectors $\vec{l}$ and $\vec{w}$ in $\mathbf{R}^3$. In this chapter we keep the symbol "$\rightarrow$" for all vectors.

We will generalize the results which were obtained in [Křížek, Neittaanmäki, 1986] for wider class of domains.

9.2. Auxiliary results

First we recall some known important facts.

Introduce a space of vector-functions the divergence of which exists in the sense of distributions (see, for example, [Girault, Raviart])

$$H(\operatorname{div};\Omega) = \{\vec{q} \in [L^2(\Omega)]^3 \mid \exists \varphi \in L^2(\Omega) : (\vec{q}, \nabla z)_0 + (\varphi, z)_0 = 0 \ \forall z \in H_0^1(\Omega)\} \tag{9.2.1}$$

and its subspace of divergence-free (so-called solenoidal) functions

$$H(\operatorname{div}^0;\Omega) = \{\vec{q} \in [L^2(\Omega)]^3 \mid (\vec{q}, \nabla z)_0 = 0 \ \forall z \in H_0^1(\Omega)\}. \tag{9.2.2}$$

Note that for both spaces the test-functions z vanish on the boundary $\partial\Omega$, so there are no conditions upon the normal flux $\vec{n} \cdot \vec{q}$ on $\partial\Omega$, where $\vec{n}$ is the outward normal to Ω.

Let $\vec{w} = (w_1, w_2, w_3) \in [H^1(\Omega)]^3$ and $z \in C_0^\infty(\Omega)$ be arbitrary functions. Then $(\operatorname{curl} \vec{w}, \nabla z)_0 = (\vec{w}, \operatorname{curl} \nabla z)_0 = 0$ due to the Green formula (1.4.18), where

$$\operatorname{curl} \vec{w} = (\partial_2 w_3 - \partial_3 w_2, \partial_3 w_1 - \partial_1 w_3, \partial_1 w_2 - \partial_2 w_1). \tag{9.2.3}$$

Hence, the density $C_0^\infty(\Omega)$ in $H_0^1(\Omega)$ implies

$$\operatorname{curl}\vec{w} \in H(\operatorname{div}^0;\Omega) \quad \forall \vec{w} \in [H^1(\Omega)]^3. \tag{9.2.4}$$

Recall (see [Girault, Raviart, p. 16]) that the functional $\vec{q} \to \vec{n}\cdot\vec{q}\,|_{\partial\Omega}$ defined on $[C^\infty(\overline{\Omega})]^3$ can be extended by continuity to a linear continuous mapping from the space $H(\operatorname{div};\Omega)$ into $H^{-1/2}(\partial\Omega)$, where the latter is the dual space to the space of traces $H^{1/2}(\partial\Omega)$ of functions from $H^1(\Omega)$. In this case, the Green formula takes the form

$$(\vec{q},\nabla z)_0 + (\operatorname{div}\vec{q}, z)_0 = \langle \vec{n}\cdot\vec{q}, z\rangle_{\partial\Omega} \quad \forall \vec{q} \in H(\operatorname{div};\Omega) \forall z \in H^1(\Omega), \tag{9.2.5}$$

where $\langle\cdot\,,\cdot\rangle_{\partial\Omega}$ denotes the duality pairing between $H^{-1/2}(\partial\Omega)$ and $H^{1/2}(\partial\Omega)$.

Now we will formulate and prove Theorem 9.2.1.

Theorem 9.2.1. *Let $\vec{l} = (l_1, l_2, l_3)$ be a constant vector in $\mathbf{R}^3$ and $\Omega \subset \mathbf{R}^3$ a bounded domain with a Lipschitz continuous boundary. Then*

$$H_0(\operatorname{div}^0;\Omega) = \operatorname{curl} W, \tag{9.2.6}$$

where

$$W = \{\vec{w} = (w_1, w_2, w_3) \in [H^1(\Omega)]^3 \mid \vec{n}\cdot\operatorname{curl}\vec{w} = 0 \text{ on } \partial\Omega,\ \vec{l}\cdot\vec{w} = 0 \text{ in } \Omega\}. \tag{9.2.7}$$

Proof. Let $\vec{w} \in W$ be given. Then

$$(\operatorname{curl}\vec{w}, \nabla z)_0 = (-\operatorname{div}\operatorname{curl}\vec{w}, z)_0 + \langle \vec{n}\cdot\operatorname{curl}\vec{w}, z\rangle_{\partial\Omega} = 0 \quad \forall z \in H^1(\Omega)$$

(see formulae (9.2.4), (9.2.5)). Hence, it follows from $\operatorname{curl}\vec{w} \in H_0(\operatorname{div}^0;\Omega)$ that

$$H_0(\operatorname{div}^0;\Omega) \supset \operatorname{curl} W. \tag{9.2.8}$$

Conversely, let $\vec{q} = (q_1, q_2, q_3) \in H_0(\operatorname{div}^0;\Omega)$, i.e.,

$$\begin{aligned} \operatorname{div}\vec{q} &= 0 \quad \text{in } \Omega, \\ \langle \vec{q}\cdot\vec{n}, 1\rangle_{\partial\Omega} &= 0. \end{aligned}$$

We can extend $\vec{q}$ (according to [Girault, Raviart, pp. 27–28]) to the whole space so that the extended function $\tilde{\vec{q}} \in [L^2(\mathbf{R}^3)]^3$ would be still divergence-free and have a compact support. Let $\hat{q}_j$ be the Fourier transform of q_j, $j = 1, 2, 3$,

$$\hat{q}_j(\xi) = \int_{\mathbf{R}^3} e^{-2i\pi x\cdot\xi}\tilde{q}_j(x)\,dx, \quad \xi \in \mathbf{R}^3. \tag{9.2.9}$$

Here i is the imaginary unit, i.e., $i^2 = -1$. In what follows we will write $\mathbf{R}^3_\xi$ for the three-dimensional space with coordinates (ξ_1, ξ_2, ξ_3). The condition $\operatorname{div} \vec{\bar{q}} = 0$ implies that

$$\sum_{i=1}^{3} \xi_i \hat{q}_i = 0. \tag{9.2.10}$$

We seek a function $\vec{\bar{\varphi}}$ in $[L^2(\mathbf{R}^3_\xi)]^3$ such that $\operatorname{curl} \vec{\varphi} = \vec{\bar{q}}$, i.e.,

$$\begin{cases} \hat{q}_1 = 2i\pi(\xi_2\hat{\varphi}_3 - \xi_3\hat{\varphi}_2), \\ \hat{q}_2 = 2i\pi(\xi_3\hat{\varphi}_1 - \xi_1\hat{\varphi}_3), \\ \hat{q}_3 = 2i\pi(\xi_1\hat{\varphi}_2 - \xi_2\hat{\varphi}_1). \end{cases} \tag{9.2.11}$$

Obviously, the third equation of (9.2.11) is a consequence of the first two and equation (9.2.10), hence, in fact, we have only two equations to define three unknown functions $\hat{\varphi}_1$, $\hat{\varphi}_2$, $\hat{\varphi}_3$.

Further, we add the following third condition which is suitable for our purposes:

$$\sum_{i=1}^{3} l_i \varphi_i = 0 \tag{9.2.12}$$

which, after the Fourier transform, takes the form

$$\sum_{i=1}^{3} l_i \hat{\varphi}_i = 0, \tag{9.2.13}$$

due to the fact that $\vec{l}$ is a constant vector. Equation (9.2.13) is the third relation connecting the functions $\hat{\varphi}_1$, $\hat{\varphi}_2$, $\hat{\varphi}_3$.

Hence, taking the first two equations from system (9.2.11) and equation (9.2.13) we obtain the system

$$\begin{cases} \xi_2\hat{\varphi}_3 - \xi_3\hat{\varphi}_2 = \dfrac{\hat{q}_1}{2\pi i}, \\ \xi_3\hat{\varphi}_1 - \xi_1\hat{\varphi}_3 = \dfrac{\hat{q}_2}{2\pi i}, \\ l_1\hat{\varphi}_1 + l_2\hat{\varphi}_2 + l_3\hat{\varphi}_3 = 0. \end{cases} \tag{9.2.14}$$

In the matrix form it may be rewritten as follows:

$$\begin{bmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ l_1 & l_2 & l_3 \end{bmatrix} \begin{bmatrix} \hat{\varphi}_1 \\ \hat{\varphi}_2 \\ \hat{\varphi}_3 \end{bmatrix} = \begin{bmatrix} \hat{q}_1/2\pi i \\ \hat{q}_2/2\pi i \\ 0 \end{bmatrix}. \tag{9.2.15}$$

The solution is

$$\begin{bmatrix} \hat{\varphi}_1 \\ \hat{\varphi}_2 \\ \hat{\varphi}_3 \end{bmatrix} = \frac{1}{2\pi i \xi_3 \vec{l} \cdot \vec{\xi}} \begin{bmatrix} \xi_1 l_2 \hat{q}_1 + \xi_3 l_3 \hat{q}_2 + \xi_2 l_2 \hat{q}_2 \\ -\xi_3 l_3 \hat{q}_1 - \xi_1 l_1 \hat{q}_1 - \xi_2 l_1 \hat{q}_2 \\ \xi_3 l_2 \hat{q}_1 - \xi_3 l_1 \hat{q}_2 \end{bmatrix}. \tag{9.2.16}$$

The function defined by (9.2.16) represents the unique solution of system (9.2.14), because the determinant of the matrix in (9.2.15) is not equal to zero.

Now, we have the following facts:

1) $\hat{q}_j$ are holomorphic in $\mathbf{R}^3_\xi$, since the supports of $\tilde{q}_j$ are compact (see [Girault, Raviart, p. 27]).
2) The Fourier transform is a linear continuous operator from $L_2(\mathbf{R}^3)$ to $L_2(\mathbf{R}^3_\xi)$, hence $\hat{q}_j \in L_2(\mathbf{R}^3_\xi), j = 1,2,3$.

We recall the following theorem:

Theorem 9.2.2. *Let k and d be any integers. Then*

$$u(x) \in H^k(\mathbf{R}^d) \iff \xi^\alpha \hat{u}(\xi) \in L_2(\mathbf{R}^d_\xi) \quad \forall \alpha \text{ such that } |\alpha| \leq k$$

(see, for example, [Vladimirov]), where the sign "$\hat{\ }$" means the Fourier transform and α is a multi-index.

According to Theorem 9.2.2, in order to get $\varphi_j \in H^1(\Omega)$, $j = 1,2,3$, we shall prove the following theorem.

Theorem 9.2.3. *The statements*

(a) $\xi_j \hat{\varphi}_i(\xi) \in L_2(\mathbf{R}^3_\xi)$, $i, j = 1,2,3$,

(b) $\hat{\varphi}_i(\xi) \in L_2(\mathbf{R}^3_\xi)$, $i = 1,2,3$

are valid, where $\hat{\varphi}_i(\xi)$ and $\mathbf{R}^3_\xi$ are described above.

Proof. Condition (a) can be proved immediately from formula (9.2.10). We also have

$$|\hat{\varphi}_i| \leq \frac{C(|\hat{q}_j| + |\hat{q}_k|)}{\|\xi\|},$$

where $C > 0$ is a constant. Hence, we must check only the boundedness of $\hat{\varphi}_i$ in the neighbourhood of zero.

Condition (9.2.10) implies

$$\hat{q}_i(0) = 0. \tag{9.2.17}$$

From 1) we have

$$\hat{q}_i(\xi) = \sum_{j=1}^{3} \xi_j \frac{\partial \hat{q}_i}{\partial \xi_j}(0) + \mathcal{O}(\|\xi\|^2)$$

in a neighbourhood of 0. Here $\|\xi\|$ means the usual Euclidean norm of the vector $\vec{\xi} = (\xi_1, \xi_2, \xi_3)$. Hence, $\vec{\hat{\varphi}}$ is bounded as $\xi \to 0$. □

By restricting the inverse transform $\vec{\varphi}$ of $\vec{\hat{\varphi}}$ to Ω, we get a function $\vec{\varphi} \in [H^1(\Omega)]^3$ such that

$$\operatorname{curl} \vec{\varphi} = \vec{q}$$

and, moreover, the important identity $\vec{l} \cdot \vec{\varphi} = 0$ is valid. Note that in [Girault, Raviart] and [Křížek, Neittaanmäki, 1986] the vector $\vec{l}$ is, in fact, equal to $(0, 0, 1)$. □

9.3. Equilibrium finite elements

Let W_h be an arbitrary finite element space of W whose functions are continuous and piecewise polynomial on some partition of $\overline{\Omega}$. We define the space of equilibrium finite elements as

$$Q_h = \operatorname{curl} W_h. \tag{$*$}$$

Due to Theorem 9.2.1, Q_h is a subspace of $H_0(\operatorname{div}^0; \Omega)$. Recall (see [Křížek, Neittaanmäki, 1986], Corollary of Theorem 1) that, if $\{W_h\}$ is a system of finite element subspaces of W such that the union $\bigcup_h W_h$ is dense in W with respect to the $\|\cdot\|_1$ norm, then $\bigcup_h Q_h$ is dense in $H_0(\operatorname{div}^0; \Omega)$ in the $\|\cdot\|_0$ norm.

Definition 9.3.1. *A domain $\Omega \subset \mathbf{R}^3$ is said to belong to the class $\mathcal{L}^*$ if it can be transformed by a rotation in $\mathbf{R}^3$ to the domain Ω' from the class $\mathcal{L}$ (see [Křížek, Neittaanmäki, 1986]), i.e.,*

(i) Ω' is a bounded domain with a Lipschitz boundary,

(ii) there exists a simply connected domain $\omega \subset \mathbf{R}^2$ and a positive function $F\colon \omega \to \mathbf{R}^1$ (in general discontinuous) such that

$$\Omega' = \{(x_1, x_2, x_3) \in \mathbf{R}^3 \mid (x_1, x_2) \in \omega,\ 0 < x_3 < F(x_1, x_2)\}.$$

Remark 9.3.2. Denote by $\partial\Omega_0$ the base of the domain Ω, i.e., ω is the image of $\partial\Omega_0$ under the above rotation. Then there exists a constant vector $\vec{l} \in \mathbf{R}^3$ which is perpendicular to the base of such a domain (see Figure 9.3.1).

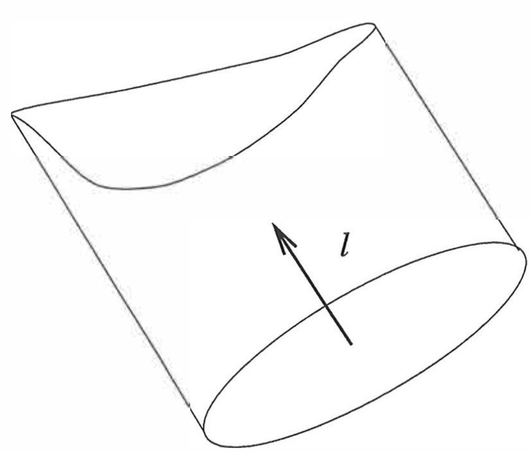

Figure 9.3.1.

Further we shall require the following property of finite element subspaces ($\Omega \in \mathcal{L}^*$ with the vector $\vec{l}$) to be valid:

$$\vec{w} \in W_h \implies \vec{\overset{\circ}{w}} \in W_h, \tag{9.3.1}$$

where

$$\vec{\overset{\circ}{w}}(x_1, x_2, x_3) = \vec{w}(x_1^0, x_2^0, x_3^0) \tag{9.3.2}$$

and the points

$$x^0 = (x_1^0, x_2^0, x_3^0) \in \partial\Omega_0, \quad x = (x_1, x_2, x_3) \in \Omega$$

are connected by the following relation:

$$x_i - x_i^0 = \alpha \cdot l_i, \quad i = 1, 2, 3 \quad (\alpha \text{ is a constant}), \tag{9.3.3}$$

i.e., the point x^0 is the projection of the point x onto the base of the domain along the vector $\vec{l}$.

For simplicity we choose the vector $\vec{l}$ to be of the unit length, i.e.,

$$\|\vec{l}\| = (l_1^2 + l_2^2 + l_3^2)^{1/2} = 1. \tag{9.3.4}$$

Note that the operator curl: $W_h \to Q_h = \operatorname{curl} W_h$ is not bijective in general, so we need to define $V_h \subset W_h$ such that curl: $V_h \to Q_h$ is bijective.

The next theorem generalizes Theorem 2 from [Křížek, Neittaanmäki, 1986].

Theorem 9.3.3. *Let $\Omega \in \mathcal{L}^*$ and let the vector $\vec{l}$ correspond to this domain (see Remark 9.3.2). Let $W_h \subset W$ satisfy (9.3.2) and $Q_h = \operatorname{curl} W_h$. Then for the space $V_h \subset W_h$ such that*

$$V_h = \{\vec{v} \in W_h \mid \vec{v} = 0 \text{ on } \partial\Omega_0\}$$

the mapping

$$\operatorname{curl}\colon V_h \to Q_h \subset H_0(\operatorname{div}^0; \Omega)$$

is bijective.

Proof. Injectivity. If $\operatorname{curl} \vec{v} = 0$ for some $\vec{v} \in W_h$ then there exists $s \in H^1(\Omega)$ (note that Ω is simply connected) such that

$$\vec{v} = \operatorname{grad} s.$$

Moreover, $s \in H^2(\Omega) \subset C(\overline{\Omega})$. Hence, s is continuous in $\overline{\Omega}$ and, of course, s is a piecewise polynomial function. Due to these facts the following formula makes sense:

$$s(x_1, x_2, x_3) = s(x_1^0, x_2^0, x_3^0) + \int_{x_0}^{x} \frac{\partial s}{\partial \vec{l}} \, d\xi,$$

where the point $(x_1^0, x_2^0, x_3^0) \in \partial\Omega_0$ is the projection of the point (x_1, x_2, x_3) to the base of Ω along the vector $\vec{l}$. It is obvious that

$$\frac{\partial s}{\partial \vec{l}} = \vec{l} \cdot \nabla s = \vec{l} \cdot \vec{v} = 0,$$

which implies that $s(x) = s(x^0)$.

Since $\vec{v} = 0$ on $\partial\Omega_0$, we get that s is constant on $\partial\Omega_0$ and, then, in the whole domain Ω. This means that $\vec{v} \equiv 0$ in $\overline{\Omega}$.

Surjectivity. Let $\vec{q} \in Q_h$ be an arbitrary vector function. According to Theorem 9.2.1, there exists a continuous piecewise polynomial function $\vec{w} = (w_1, w_2, w_3)$ such that $\vec{w} \in W_h$, $\vec{l} \cdot \vec{w} = 0$ and

$$\vec{q} = \operatorname{curl} \vec{w}.$$

Let $\vec{v} = \vec{w} - \vec{\overset{\circ}{w}}$, where $\vec{\overset{\circ}{w}} = (\overset{\circ}{w}_1, \overset{\circ}{w}_2, \overset{\circ}{w}_3)$ is defined by (9.3.2) and (9.3.3). Then $\vec{v} = 0$ on $\partial\Omega_0$ and $\vec{v} \in V_h \subset W_h$.

Now we check whether the relation

$$\vec{q} = \operatorname{curl} \vec{v}$$

holds.

In fact, we must show that

$$\operatorname{curl} \vec{\overset{\circ}{w}} = 0 \quad \text{in } \Omega.$$

Let us introduce the following convenient notation:

$$\begin{cases} \dfrac{\partial \overset{\circ}{w}_3}{\partial x_2} - \dfrac{\partial \overset{\circ}{w}_2}{\partial x_3} = \Delta_1, \\ \dfrac{\partial \overset{\circ}{w}_1}{\partial x_3} - \dfrac{\partial \overset{\circ}{w}_3}{\partial x_1} = \Delta_2, \\ \dfrac{\partial \overset{\circ}{w}_2}{\partial x_1} - \dfrac{\partial \overset{\circ}{w}_1}{\partial x_2} = \Delta_3. \end{cases} \tag{9.3.5}$$

Since $\vec{\overset{\circ}{w}}$ is, in fact, the trace of $\vec{w}$ on $\partial\Omega_0$ along the vector $\vec{l}$ we have the following obvious conditions:

$$\sum_{i=1}^{3} l_i \frac{\partial \overset{\circ}{w}_j}{\partial x_i} = 0, \; j = 1, 2, 3 \quad \text{in } \Omega. \tag{9.3.6}$$

As $\vec{l} \cdot \vec{\overset{\circ}{w}} = 0$, we also have

$$\sum_{i=1}^{3} l_i \overset{\circ}{w}_i = 0 \quad \text{in } \Omega. \tag{9.3.7}$$

And, of course, the following condition will be taken into account:

$$\vec{n} \cdot \operatorname{curl} \vec{\overset{\circ}{w}}\big|_{\partial\Omega_0} = 0. \tag{9.3.8}$$

For simplicity, we suppose that $l_3 \neq 0$. Then (9.3.7) yields

$$\overset{\circ}{w}_3 = -\frac{l_2}{l_3}\overset{\circ}{w}_2 - \frac{l_1}{l_3}\overset{\circ}{w}_1.$$

Hence,

$$\begin{aligned}
\Delta_1 &= \frac{\partial \overset{\circ}{w}_3}{\partial x_2} - \frac{\partial \overset{\circ}{w}_2}{\partial x_3} \\
&= \frac{\partial}{\partial x_2}\Big(-\frac{l_2}{l_3}\overset{\circ}{w}_2 - \frac{l_1}{l_3}\overset{\circ}{w}_1\Big) - \frac{\partial}{\partial x_3}\overset{\circ}{w}_2 \\
&= -\frac{1}{l_3}\Big(l_2\frac{\partial \overset{\circ}{w}_2}{\partial x_2} + l_1\frac{\partial \overset{\circ}{w}_1}{\partial x_2} + l_3\frac{\partial \overset{\circ}{w}_2}{\partial x_3}\Big) \\
&= -\frac{1}{l_3}\Big(l_1\frac{\partial \overset{\circ}{w}_2}{\partial x_1} + l_2\frac{\partial \overset{\circ}{w}_2}{\partial x_2} + l_3\frac{\partial \overset{\circ}{w}_2}{\partial x_3} + l_1\frac{\partial \overset{\circ}{w}_1}{\partial x_2} - l_1\frac{\partial \overset{\circ}{w}_2}{\partial x_1}\Big) \\
&= \frac{l_1}{l_3}\Delta_3.
\end{aligned}$$

and consequently,

$$l_3\Delta_1 = l_1\Delta_3.$$

It is easy to check that if l_3 is zero then the above equality also holds. Similar argument leads to the equalities $l_2\Delta_1 = l_1\Delta_2$ and $l_3\Delta_2 = l_2\Delta_3$.

These equalities constitute the system

$$\begin{cases} l_3\Delta_1 = l_1\Delta_3, \\ l_2\Delta_1 = l_1\Delta_2, \\ l_3\Delta_2 = l_2\Delta_3. \end{cases} \tag{9.3.9}$$

Obviously, only two equalities from system (9.3.9) are independent. Condition (9.3.8) implies

$$l_1\Delta_1 + l_2\Delta_2 + l_3\Delta_3 = 0$$

(since $\|\vec{l}\| = 1$ and $\vec{l} = -\vec{n}$ on $\partial\Omega_0$, if $\Omega \in \mathcal{L}^*$).

Taking the system

$$\begin{cases} l_3\Delta_1 - l_1\Delta_3 = 0, \\ l_2\Delta_1 - l_1\Delta_2 = 0, \\ l_1\Delta_1 + l_2\Delta_2 + l_3\Delta_3 = 0, \end{cases}$$

with zero right-hand side, we see that, if

$$\det \begin{bmatrix} l_3 & 0 & -l_1 \\ l_2 & -l_1 & 0 \\ l_1 & l_2 & l_3 \end{bmatrix} = -l_1 \cdot \|\vec{l}\| = -l_1 \neq 0,$$

then the only solution is $\Delta_1 = \Delta_2 = \Delta_3 = 0$. Obviously, if $l_1 = 0$ then we take other two equations from (9.3.9). □

Remark 9.3.4. Note that we have to form finite elements according to the position of the base of such domain in the space, so conditions (9.3.2) and (9.3.3) are quite natural and can be easily satisfied when employing prismatic or rectangular C^0-elements.

Also, the restriction $\vec{l} \cdot \vec{w} = 0$ is not very difficult, because $\vec{l}$ is a constant vector. The basis in V_h can be obtained from the finite element basis of finite element subspaces of $H^1(\Omega)$.

Under the assumptions of Theorem 9.3.3 we can introduce basis functions of the space $Q_h = \operatorname{curl} V_h$. Namely, let $\{\mathbf{v}^i\}_{i=1}^m$ be basis in V_h, then

$$\mathbf{q}^i = \operatorname{curl} \mathbf{v}^i, \quad i = 1, \ldots, m \tag{9.3.10}$$

are basis functions in $Q_h \subset H_0(\operatorname{div}^0; \Omega)$ since the linear mapping

$$\operatorname{curl}\colon V_h \to Q_h$$

is bijective under our conditions on the domain Ω. Moreover, we see that $\operatorname{supp} \mathbf{q}^i \subseteq \operatorname{supp} \mathbf{v}^i$, $i = 1, \ldots, m$.

Note that $\mathbf{q}^i$ are not, in general, continuous in the whole $\overline{\Omega}$. However, similarly to [Křížek, 1982] we may prove, by the Green formula (9.2.5) that the normal component $n_f \cdot \mathbf{q}^i$ is continuous at each common face f of any two adjacent elements, where n_f is a normal to f.

Example 9.3.5. Let the vector $l = (0, 0, 1)$ for simplicity. We introduce a typical shape of divergence-free basis functions derived from the usual trilinear elements, the ansatz-polynomials, which are of the form

$$c_0 + c_1 x_1 + c_2 x_2 + c_3 x_3 + c_4 x_1 x_2 + c_5 x_1 x_3 + c_6 x_2 x_3 + c_7 x_1 x_2 x_3$$

on every rectangular element, see [Ciarlet]. Assume for simplicity that a uniform mesh (with the mesh size h) is given and let, e.g., $\mathbf{y} = (0, h, h)$ be a nodal point in $\overline{\Omega}$. If $\mathbf{y} \notin \partial\Omega$ then we can have two standard basis functions $\mathbf{v}^i, \mathbf{v}^{i+1} \in V_h$ for some $i \in \{1, \ldots, m\}$ such that

$$\operatorname{supp} \mathbf{v}^i = \operatorname{supp} \mathbf{v}^{i+1} = [-h, h] \times [0, 2h] \times [0, 2h].$$

This support consists of eight elements. One of them is, for instance, $K = [0, h] \times [0, h] \times [0, h]$, and we may immediately obtain

$$\begin{aligned} \mathbf{v}^i &= ((h - x_1)x_2x_3, 0, 0)/h^3 \quad \text{in } K, \\ \mathbf{v}^{i+1} &= (0, (h - x_1)x_2x_3, 0)/h^3 \quad \text{in } K. \end{aligned} \tag{9.3.11}$$

Now a direct calculation leads to

$$\begin{aligned} \mathbf{q}^i &= (0, (h - x_1)x_2, (x_1 - h)x_3)/h^3 \quad \text{in } K, \\ \mathbf{q}^{i+1} &= ((x_1 - h)x_2, 0, -x_2x_3)/h^3 \qquad \text{in } K. \end{aligned} \tag{9.3.12}$$

Similarly we obtain $\mathbf{q}^i$ and $\mathbf{q}^{i+1}$ on the other seven elements of $\operatorname{supp} \mathbf{v}^i$.

Suppose further that $\mathbf{y} = (0, h, h) \in \partial\Omega$. For simplicity let $\Omega = (0,1) \times (0,1) \times (0,1)$. In this case, we have $\partial_2 v_3(\mathbf{y}) = \partial_3 v_2(\mathbf{y})$ for any $\mathbf{v} \in V_h$, since $(-1, 0, 0)$ is the exterior unit normal to $\partial\Omega$ at $\mathbf{y}$. Thus by (9.3.11) we find that $\mathbf{v}^i \in V_h$ ($\mathbf{v}^{i+1} \notin V_h$). The corresponding support of $\mathbf{q}^i$ will consist of four elements only and, e.g., $\mathbf{q}^i|_K$ will be given by (9.3.12).

PART IV

On Construction of Strongly Regular Family of Triangulations for Planar Domains

On contents of Part IV

The main topic of Part IV is the construction of a strongly regular family of triangulations for planar domains with piecewise C^2-smooth boundaries.

The contents is based mainly on the own work of the author, see [Korotov, 1997b].

Chapter 10
Strongly regular family of triangulations for planar domains

In Section 2.6 we introduced the conceptions of the standard quasitriangle (see Figure 2.6.3) and the linear homotopy mapping (defined by the relations (2.6.3)).

For conveniency of the reader we reproduce here Figure 2.6.3 again.

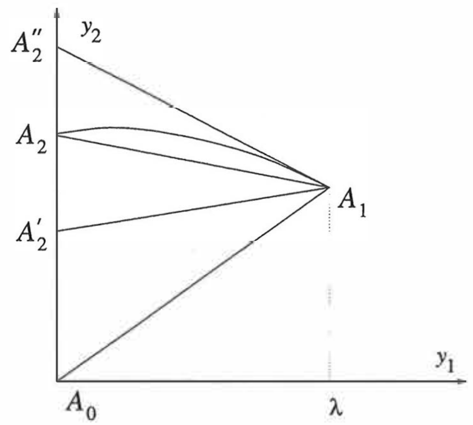

Figure 2.6.3.

These ideas are found to be very convenient for the construction of the strongly regular families of triangulations for planar domains with piecewise C^2-smooth boundaries, see Section 2.7 for definitions.

First, in Section 10.1, we present a constructive proof of an existence of the strongly regular family of triangulations for the standard quasitriangle, see Definition 2.6.1. Further, in Section 10.2, an algorithm of a division of the planar domains into a finite number of the standard quasitriangles is presented. The main result is formulated as Theorem 10.2.2.

In Section 10.3 the similar approach is applied to the convex quadrangle. Section 10.4 presents a short survey of the other approaches to the problem for planar and space domains with smooth boundaries.

10.1. Construction of triangulations for the standard quasitriangle

Throughout this section we use number of denotations from Section 2.6.

From now on we suppose that for $n \to \infty$ and $l \equiv \lambda/n$ (λ is the y_1-coordinate of the point A_1, see Figure 2.6.3) we have

$$0 < C_1 \leq l/h \leq C_2, \tag{10.1.1}$$

where C_1, C_2 are the constants independent of n and h, i.e., the number l can be taken, in fact, as the discretization parameter instead of h.

First, we formulate the base result.

Theorem 10.1.1. *For any standard quasitriangle K' of the order $m \geq 1$ we may find a sequence of polygonal domains $\{K'_h\}$ such that $[A_0A_1] \cup [A_0A_2] \subset \partial K'_h$ and either $K' \subset K'_h$ or $K'_h \subset K'$ for all $h \in (0, h_0)$, where h_0 is a sufficiently small number.*

For this sequence $\{K'_h\}$ we may form a family of triangulations $\mathcal{F} = \{\mathcal{T}_h(K'_h)\}$ such that for any $h \in (0, h_0)$, any triangulation $\mathcal{T}_h \in \mathcal{F}$ and any triangle $K \in \mathcal{T}_h$ the following estimates (cf. (2.7.3)) hold:

$$\varkappa_1 h \leq l_K \leq \varkappa_2 h, \tag{10.1.2}$$

$$0 < \theta \leq \varphi_K \leq \pi - \theta, \tag{10.1.3}$$

where l_K is a length of any edge of K and φ_K is the minimal angle between edges of K. Here, $\varkappa_1$, $\varkappa_2$, θ are the constants independent of h.

Thus, there exists a strongly regular family of decompositions into triangles for the standard quasitriangle of the second order.

Proof. We form the necessary polygonal domains K'_h and their triangulation in the following manner: first, we divide the segment $[A_0A_1]$ by the points $A_{00} \equiv A_0$, $A_{11}, ..., A_{nn} \equiv A_1$ into n equivalent segments (see Figure 10.1.1 with $n = 4$).

Second, we form rays, which are parallel to the segment $[A_0A_2]$ and start from the points $A_{00}, ..., A_{nn}$. On each ray we take a point $A_{rn} = (lr, y^*_{lr})$, $r = 0, ..., n$, such that

$$|y^*_{lr} - f(lr)| \leq C_3 l^2. \tag{10.1.4}$$

We joint all points y^*_{lr} by the straight lines segments and denote the resulting piecewise linear graph by the symbol f^*.

To provide the inclusion $K'_h \supset K'$ we shall take $f^* \geq f$ and to construct $K'_h \subset K'$ we shall take $f^* \leq f$. This graph f^*, together with the segments $[A_0A_1]$ and $[A_0A_2]$, form the boundary $\partial K'_h$.

Further, the segments $[A_{rr}A_{rn}]$ are divided by the points A_{rq} $(r \leq q \leq n)$ into $n - r$ equivalent smaller segments. The points A_{r-1q} and A_{rq}, also A_{rq} and A_{r+1q+1}, are joint by the straight line segments. As a result we get some triangulation of K_h (see Figure 10.1.1).

Now, we prove that the estimates (10.1.2) and (10.1.3) hold.

We denote by the symbol Γ the curved part of the boundary of the quasitriangle K'. Let Γ_0 denote a part of the boundary of the original domain Ω $(K' \subseteq \Omega)$, where the essential boundary condition is given (cf. (1.5.2)). We always suppose that, either $\Gamma \subset \Gamma_0$ or $\Gamma \subset \partial\Omega \setminus \Gamma_0$.

We will present a method of choosing the points A_{rn}; it depends on the position of Γ with respect to Γ_0, and on which of the conditions (2.6.1), (2.6.2) holds in $[0, \lambda]$.

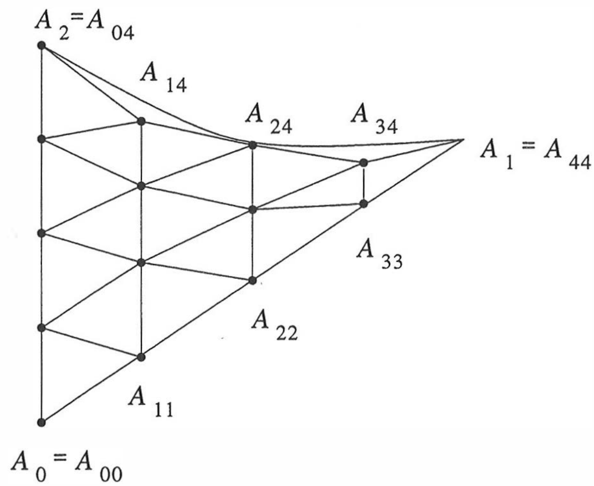

Figure 10.1.1.

Case (i). If $\Gamma \subset \Gamma_0$ and (2.6.2) holds then the points A_{rn} can be taken as the points of an intersection of the defined above rays with Γ.

The same points can be taken when $\Gamma \cap \Gamma_0 = \emptyset$ and (2.6.2) holds.

For both cases the constant C_3 in (10.1.4) is equal to 0.

In fact, for both of the above possibilities the vertices of the resulting triangulations are the images of the vertices of the usual triangulations of the triangle $K = \triangle A_0 A_1 A_2$ under the linear homotopy mapping.

Case (ii). Let $\Gamma \subset \Gamma_0$ and (2.6.1) hold. Consider several subcases. Let n be an even number.

(1) If r is even, then A_{rn} are the points of an intersection of Γ and the rays defined above.

(2) If r is odd, then A_{rn} are the first points of an intersection of the rays with the tangent lines to Γ at the neighbouring points A_{r-1n} and A_{r+1n} (which have been already found).

Doing this procedure we observe that

$$C_3 \leq 0.5 \max_{y_1 \in [0,\lambda]} |f''(y_1)|.$$

Let n be an odd number, i.e., let $n = 2k+1$. Then, for $r = 0, 1, ..., 2k-2$ we use the same rule as in (1) and (2).

Two points A_{rn} with $r = 2k-1$ and $r = 2k$ are left. To define them we use the following procedure: they are taken as the first points of an intersection of the corresponding rays with the tangent lines to Γ at the points $A_{2k-2,n}$ and $A_{2k+1,n}$. Obviously, now,

$$C_3 \leq 2 \max_{y_1 \in [0,\lambda]} |f''(y_1)|.$$

The defined procedure is correct if $f^*(y_1) > ky_1$, where $y_2 = ky_1$ is the equation of the straight line (A_0A_1). This, obviously, holds if

$$l \le \bar{l} < (k - k_1')/C_3, \tag{10.1.5}$$

i.e., $n \ge \lambda/\bar{l}$.

The case, when $\Gamma \cap \Gamma_0 = \emptyset$ and (2.6.1) holds, can be treated in a similar manner.

Now, we will show that for the constructed triangulation the properties (10.1.2) and (10.1.3) are valid. We observe that any vertex A_{rq} of the triangulation of K_h' has the following coordinates in the Descartes coordinate system (y_1, y_2):

$$A_{rq} = \Big(rl, krl + \frac{y_{rl}^* - krl}{n - r}(q - r)\Big).$$

Consider the segment $[A_{rq}, A_{rq+1}]$, its length

$$|A_{rq}A_{rq+1}| = (y_{rl}^* - krl)/(n - r).$$

From the trivial geometrical observations (see Figure 2.6.3) we have ($y_1 \in [0,\ \lambda]$):

$$k(\lambda - y_1) - k_1'(\lambda - y_1) \le f(y_1) - ky_1 \le k(\lambda - y_1) - k_1''(\lambda - y_1).$$

Then, in view of (10.1.4),

$$(k - k_1')l - \frac{C_3 l^2}{n - r} \le |A_{rq}A_{rq+1}| \le (k - k_1'')l + \frac{C_3 l^2}{n - r}.$$

Having (10.1.5), we observe that for sufficiently large n

$$0 < C_1 \le \frac{|A_{rq}A_{rq+1}|}{l} \le C_2. \tag{10.1.6}$$

Further, denote the value $\dfrac{f(rl) - krl}{n - r}$ by s_r, then

$$k - k_1' \le \frac{s_r}{l} \le k - k_1''. \tag{10.1.7}$$

Consider the vector

$$\overrightarrow{A_{r-1,q}A_{rq}} = (l,\ \bar{y}), \tag{10.1.8}$$

where, obviously,

$$\bar{y} = kl + (q - r)\frac{y_{rl}^* - krl}{n - r} - (q - r + 1)\frac{y_{r-1,l}^* - k(r - 1)l}{n - r + 1}. \tag{10.1.9}$$

Obviously, to get estimates similar to (10.1.6), it is enough to evaluate the second coordinate $\bar{y}$ only.

We observe that

$$krl < f(rl) - C_3 l^2 \le y^*_{rl} \le f(rl) + C_3 l^2. \tag{10.1.10}$$

Then

$$\bar{y} \le kl + (q-r)\left(\frac{s_r}{l} + C_3\frac{l}{n-r}\right) - (q-r+1)\left(\frac{s_{r-1}}{l} - C_3\frac{l}{n-r+1}\right).$$

From (10.1.7) we have

$$\frac{\bar{y}}{l} \le k + (q-r)(k-k_1'') - (q-r+1)(k-k_1') + C_3 l\left(\frac{q-r}{n-r} + \frac{q-r+1}{n-r+1}\right).$$

Obviously, $0 \le q - r \le n$. Also, when $r = n$ then the constant C_3 from (10.1.4) can be taken equal to zero. All these facts imply

$$\frac{\bar{y}}{l} \le C_4, \tag{10.1.11}$$

where C_4 is some constant. Analogically,

$$0 < C_5 \le \frac{\bar{y}}{l}. \tag{10.1.12}$$

Further, from (10.1.11), (10.1.12) and (10.1.8) we get

$$1 \le \frac{|A_{r-1q}A_{rq}|}{l} \le C_6. \tag{10.1.13}$$

From (10.1.13) we have

$$\sin \widehat{A_{r-1q-1}A_{r-1q}A_{rq}} \ge 1/C_6. \tag{10.1.14}$$

Now, (10.1.6), (10.1.13) and (10.1.14) imply, obviously, (10.1.2) and (10.1.3). □

Corollary 10.1.2. *From the constructions given in the proof of Theorem 10.1.1 we observe that there exists a bijective piecewise continuous differentiable mapping between points $y \in \Gamma$ and points y^* of the graph f^**

$$y^* = y + \xi(y), \quad \xi(y) = (\xi_1(y), \xi_2(y)), \tag{10.1.15}$$

such that

$$|\xi(y)| \le \varkappa_5 h^2, \quad \left|\frac{\partial \xi_r}{\partial y_l}\right| \le \varkappa_6 h, \ r, l = 1, 2. \tag{10.1.16}$$

10.2. On a division of domains into the standard quasitriangles

Theorem 10.2.1. *Let Ω be a simply connected planar domain with the boundary $\partial\Omega$, which consists of a finite number of arcs γ_r meeting at the interior angles φ, where*

$$0 \le \varphi_0 \le \varphi \le 2\pi - \varphi_0, \tag{10.2.1}$$

and each arc γ_r is twice differentiable and is either convex or concave with respect to Ω. Then there exists a division of $\overline{\Omega}$ into a finite number of the standard quasitriangles of the second order with the curved sides on $\partial\Omega$ only.

Proof. For simplicity, we present only an algorithm for a construction of such a division.

Let h_0 be some fixed number, sufficiently small with respect to the smallest radius of the curvature of any curve γ_r. First, we build an uniform rectangular mesh in $\mathbf{R}^2$ with the step-size equal to h_0.

Further, we define a polygon $P \subset \Omega$ in the following manner. If a square from the uniform mesh belongs to Ω, then we divide this square into two triangles by drawing one of its diagonal and include the both triangles into P. If only three vertices of some square and the triangle, formed by these vertices, belong to Ω, then we also include this triangle into P.

Now, consider a strip between ∂P and $\partial\Omega$. Obviously, connecting the vertices of P with certain points of the boundary $\partial\Omega$ (e.g., along the normal to $\partial\Omega$) or (and) with the point-intersections of the arcs forming $\partial\Omega$, we can divide the strip into a finite number of the standard quasitriangles with the curved side on $\partial\Omega$ only (see Figure 10.2.1). □

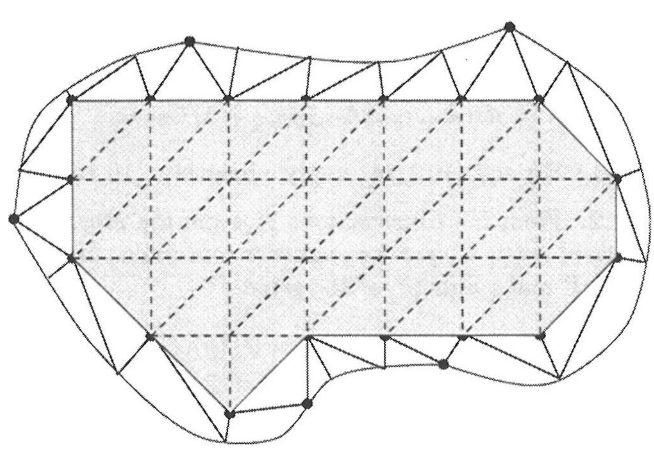

Figure 10.2.1.

Theorem 10.2.2. *Let Ω be a domain described in Theorem 10.2.1. Then there exists a strongly regular family of triangulation Ω into triangles with $h \to 0$.*

Proof. If the division of the given domain Ω into a finite number of the standard quasitriangles from Theorem 10.2.1 is performed, then for any number $n \to \infty$ we form triangulations of each of the quasitriangles as in Theorem 10.1.1. Obviously, all triangulations of all quasitriangles (for fixed n) can be taken together as the global triangulation of the whole domain Ω in view of a way of the construction.

The only obstacle is to provide the nonexistence of common points for any two piecewise linear approximations of two neighbouring arcs, built outside of Ω. But, this requirement is, obviously, fulfilled in view of (10.2.1) for sufficiently large n, i.e., for small h. □

10.3. On triangulation of the convex quadrangle

Let Ω_h be a convex quadrangle (see Figure 10.3.1).

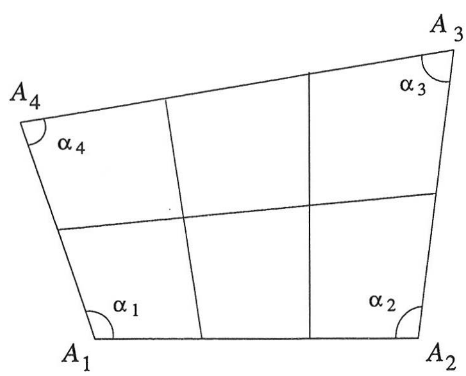

Figure 10.3.1.

We will show how to construct a strongly regular family of triangulations for such a kind of planar domains. This method is used in [Zienkiewicz] for an automatic construction of meshes.

Let the inner angle with the vertex at the point A_r, $r = 1, 2, 3, 4$, be denoted by α_r and the lengths of the sizes $[A_1A_2]$, $[A_2A_3]$, $[A_3A_4]$ and $[A_4A_1]$ by l_1, l_2, l_3 and l_4, respectively.

We divide the segments $[A_1A_2]$ and $[A_4A_3]$ by the points A_0^i and $A_{n_2}^i$, $\imath = 0, ..., n_1$, into n_1 equivalent segments. Here,

$$A_0^0 = A_1, \ A_0^{n_1} = A_2, \ A_{n_2}^0 = A_4, \ A_{n_2}^{n_1} = A_3.$$

We join the points A_0^i and $A_{n_2}^i$, $i = 0, ..., n_1$, by the segments, and each such segment $[A_0^iA_{n_2}^i]$ is divided by the points A_j^i, $j = 0, ..., n_2$, into n_2 equivalent segments; we joint then the neighbouring points A_j^i and A_j^{i+1} by the segments.

Further, each of the resulting quadrangles (the total number of small quadrangles is equal to n_1n_2) is divided by one of the diagonals into two triangles. As a result we obtain some triangulation of the given domain Ω.

Theorem 10.3.1. *Let the discretization parameter h and the numbers n_1, n_2 satisfy the following relations:*

$$0 < \varkappa_1 \leq hn_1 \leq \varkappa_2, \tag{10.3.1}$$

$$0 < \varkappa_3 \leq hn_2 \leq \varkappa_4, \tag{10.3.2}$$

where $\varkappa_i$, $i = 1, ..., 4$, are independent constants, as $n_1, n_2 \to \infty$. Then the algorithm described above leads to a strongly regular family of triangulations of the quadrangle Ω.

Proof. Let the Descartes coordinate system (y_1, y_2) be such that $A_1 = (0, 0)$ and the axes y_1 lie along the segment $[A_1A_2]$.

Further, suppose for simplicity that

$$\alpha_1 + \alpha_2 \geq \pi. \tag{10.3.3}$$

It is not difficult to see that the vector $\overrightarrow{A_0^jA_{n_2}^j}$ has the following coordinates

$$\begin{aligned}\overrightarrow{A_0^jA_{n_2}^j} = \Big(l_4 \cos\alpha_1 + \frac{j}{n_1}(l_3 \cos(\alpha_1 + \alpha_2 - \pi) - l_1), \\ l_4 \sin\alpha_1 + \frac{j}{n_1} l_3 \sin(\alpha_1 + \alpha_2 - \pi)\Big) \equiv (y_1^*, y_2^*).\end{aligned} \tag{10.3.4}$$

Further, we easily find that

$$|y_1^*| \leq \max\{l_4|\cos\alpha_1|, l_2|\cos\alpha_2|\}. \tag{10.3.5}$$

Consider now the quadrangle $Q_i = A_0^iA_0^{i+1}A_{n_2}^{i+1}A_{n_2}^i$ with the vertices on the sides $[A_1A_2]$ and $[A_4A_3]$.

It is easy to check that all inner angles φ of Q_i satisfy the condition

$$0 < \varphi_0 \leq \varphi \leq \pi - \varphi_0,$$

where φ_0 is a constant independent of n_1 as $n_1 \to \infty$.

Indeed, consider, e.g., the angle $\widehat{A_1A_0^iA_{n_2}^i}$. In view of (10.3.4) and (10.3.5) we find that

$$|\cos \widehat{A_1A_0^iA_{n_2}^i}| \leq [1 + \sin^2\alpha_1(\max\{|\cos\alpha_1|; l_2l_4^{-1}|\cos\alpha_2|\})]^{-1/2}. \tag{10.3.6}$$

For the sides of the quadrangle Q_i we have the estimates

$$|A_0^i A_0^{i+1}| = l_1/n_1,$$
$$|A_{n_2}^i A_{n_2}^{i+1}| = l_3/n_1,$$
$$l_4 \sin\alpha_1 \le |A_0^i A_{n_2}^i| \le \max\{l_2 \sin\alpha_4; l_4 \sin\alpha_3\}/\sin\varphi_0.$$

Dividing Q_i into n_2 quadrangles Q_{ij}, we find analogously, that for all inner angles ψ of the quadrangle Q_{ij}

$$0 < \psi_0 \le \psi \le \pi - \psi_0, \tag{10.3.7}$$

where the constant ψ_0 does not depend on n_1 and n_2. Moreover, for the sides of Q_{ij} we have estimates:

$$l_4 \sin\alpha_1/n_2 \le |A_j^i A_{j+1}^i| \le \max\{l_2 \sin\alpha_4; l_4 \sin\alpha_3\}/(n_2 \sin\varphi_0), \tag{10.3.8}$$
$$\min\{l_1; l_2\} \sin\varphi_0/n_1 \le |A_j^i A_j^{i+1}| \le \max\{l_1; l_3\}/(n_1 \sin\psi_0). \tag{10.3.9}$$

Obviously, (10.3.7), (10.3.8) and (10.3.9) prove the theorem. □

10.4. Remark on another approaches

There are several different approaches to the problem when the given domain has C^2-smooth boundary: based on the rectangular mesh, see [Korneev, 1977] and on triangulations of a circle, see [Matsokin, 1975a].

These algorithms can be generalized for three-dimensional smooth domains as well, see [Korneev, 1979b] and [Matsokin, 1975b].

References

Adams, R. A., *Sobolev spaces*, Academic Press, New York, 1975.

Axelsson, O. and Barker, V. A., *Finite element solution of boundary value problems. Theory and computation*, Academic Press, New York, 1984.

Bernardi, C., *Optimal finite-element interpolation on curved domains*, SIAM J. Numer. Anal. **26** (1989), 1212–1240.

Čermák, L. and Zlámal, M., *Transformation of dependent variables and the finite element solution of nonlinear evolution equations*, Internat. J. Numer. Methods Engrg. **15** (1980), 31–40.

Ciarlet, P. G., *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978.

Ciarlet, P. G. and Lions, J. L. (eds.), *Handbook of numerical analysis*, Vol. II, Finite element methods, North-Holland, Amsterdam, 1991.

Ciarlet, P. G. and Raviart, P. A., *The combined effect of curved boundaries and numerical integration in isoparametric finite element method*, The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, (A. K. Aziz, ed.), Academic Press, New York, 1972, pp. 409–474.

Courant, R., *Variational methods for problems of equilibrium and vibration*, Bull. Amer. Math. Soc. **49** (1943), 1–23.

Doktor, P., *On the density of smooth functions in certain subspaces of Sobolev spaces*, Comment. Math. Univ. Carolin. **14** (1973), 609–622.

Douglas, J., Dupont, T. and Serrin, J., *Uniqueness and comparison theorems for nonlinear elliptic equations in divergence form*, Arch. Rational Mech. Anal. **42** (1971), 157 168.

Engels, H., *Numerical quadrature and cubature*, Academic Press, London, 1980.

Feistauer, M., *On the finite element approximation of a cascade flow problem*, Numer. Math. **50** (1987), 655–684.

Feistauer, M., *Mathematical methods in fluid dynamics*, Longman, Harlow, 1993.

Feistauer, M., Felcman, J., Rokyta, R., and Vlášek, Z., *Finite-element solution of flow problems with trailing conditions*, J. Comput. Appl. Math. **44** (1992), 131–165.

Feistauer, M., Křížek, M. and Sobotíková, V., *An analysis of finite element variational crimes for a nonlinear elliptic problem of a nonmonotone type*, East-West J. Numer. Math. **1** (1993), 267–285.

Feistauer, M. and Najzar, K., *Finite element approximation of a problem with a nonlinear Newton boundary condition*, Numer. Math., (to appear).

Feistauer, M. and Sobotíková, V., *Finite element approximation of nonlinear elliptic problems with discontinuous coefficients*, RAIRO Modèl. Math. Anal. Numér. **24** (1990), 457–500.

Feistauer, M. and Ženíšek, A., *Finite element solution of nonlinear elliptic problems*, Numer. Math. **50** (1987), 451–475.

Feistauer, M. and Ženíšek, A., *Compactness method in finite element theory of nonlinear elliptic problems*, Numer. Math. **52** (1988), 147–163.

Fučík, S. and Kufner, A., *Nonlinear differential equations*, Elsevier, Amsterdam, 1980.

George, P. L., *Automatic mesh generation. Application to finite element methods*, John Wiley & Sons, Chichester, 1991.

Gilbarg, D. and Trudinger, N. S., *Elliptic partial differential equations of second order*, Springer, Berlin, 1977.

Girault, V. and Raviart, P. A., *Finite element methods for Navier–Stokes equations*, Springer, Berlin, 1986.

Glowinski, R., *Numerical methods for nonlinear variational problems*, Springer-Verlag, New York, 1984.

Glowinski, R., Lions, J. L. and Trémolières, R., *Numerical analysis of variational inequalities*, North-Holland, Amsterdam, 1981.

Goldberg, M., *Three infinite families of tetrahedral space-fillers*, J. Combin. Theory Ser. A **16** (1974), 348–354.

Hlaváček, I., *A note on the Neumann problem for a quasilinear elliptic problem of a nonmonotone type*, Preprint, Math. Inst., Prague (1996).

Hlaváček, I. and Křížek, M., *Internal finite element approximations in the dual variational methods for second order elliptic problems with curved boundaries*, Apl. Mat. **29** (1984), 52–69.

Hlaváček, I. and Křížek, M., *On a nonpotential and nonmonotone second order elliptic problem with mixed boundary conditions*, Stability Appl. Anal. Contin. Media **3** (1993a), 85–97.

Hlaváček, I. and Křížek, M., *Dual finite element analysis of three-dimensional axisymmetric elliptic problems*, Parts I–II, Numer. Methods Partial Differential Equations **9** (1993b), 507–526, 527–550.

Hlaváček, I., Křížek, M. and Malý, J., *On Galerkin approximations of quasilinear nonpotential elliptic problem of a nonmonotone type*, J. Math. Anal. Appl. **184** (1994), 168–189.

Jensen, R., *The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations*, Arch. Rational Mech. Anal. **101** (1988), 1–27.

Knobloch, P., *Discrete Friedrichs' and Korn's inequalities in two and three dimensions*, East-West J. Numer. Math. **4** (1996), 35–51.

Korneev, V. G., *Schemes of the finite element method for high orders of accuracy*, Leningrad University, Leningrad, 1977, (in Russian).

Korneev, V. G., *Finite element difference schemes of high order of accuracy f second-order elliptic equations in three-dimensional domains, I*, Zh. Vychisl. Mat. i Mat. Fiz. **19/4** (1979a), 937–949, (in Russian).

Korneev, V. G., *Finite element difference schemes of high order of accuracy f second-order elliptic equations in three-dimensional domains, II*, Zh. Vychisl. Mat. i Mat. Fiz. **19/5** (1979b), 1141–1148, (in Russian).

Korotov, S., *On equilibrium finite elements in three-dimensional case*, Appl. Math. **42/3** (1997a), 233–242.

Korotov, S., *On construction of a strongly regular family of triangulations in 2D*, (under preparation) (1997b).

Korotov, S. and Křížek, M., *Finite element analysis of variational crimes for a quasilinear elliptic problem in 3D*, (submitted to Numerische Mathematik).

Křížek, M., *An equilibrium finite element method in three-dimensional elasticity*, Apl. Mat. **27** (1982), 46–75.

Křížek, M., *Conforming equilibrium finite element methods for some elliptic plane problems*, RAIRO Anal. Numér. **17** (1983), 35–65.

Křížek, M., *On semiregular families of triangulations and linear interpolation*, Appl. Math. **36** (1991), 223–232.

Křížek, M. and Lin, Q., *On diagonal dominance of stiffness matrices in 3D*, East-West J. Numer. Math. **3** (1995), 59–69.

Křížek, M. and Liu, L., *On a comparison principle for a quasilinear elliptic boundary value problem of a nonmonotone type*, Applicationes Mathematicae **24** (1996), 97–107.

Křížek, M. and Liu, L., *Finite element approximation of nonlinear heat conduction problems in anisotropic media*, Report 4, Laboratory of Scientific Computing, Jyväskylä (1997), 1–15.

Křížek, M. and Neittaanmäki, P., *Internat. FE-approximation of spaces of divergence-free functions in three-dimensional domains*, Internat. J. Numer. Methods Fluids **6** (1986), 811–817.

Křížek, M. and Neittaanmäki, P., *Finite element approximation of variational problems and applications*, Longman Scientific & Technical, Harlow, 1990.

Křížek, M. and Neittaanmäki, P., *Mathematical and numerical modelling in electrical engineering: theory and applications*, Kluwer Academic Publishers, 1996.

Křížek, M. and Preiningerová, V., *3-d solution of temperature fields in magnetic circuits of large transformers*, Elektrotechn. obzor **76** (1987), 646–652, (in Czech).

Křížek, M. and Preiningerová, V., *Calculation of the 3d temperature field of synchronous and of induction machines by the finite element method*, Elektrotechn. obzor **80** (1991), 78–84, (in Czech).

Křížek, M. and Strouboulis, T., *How to generate local refinements of unstructured tetrahedral meshes satisfying a regularity ball condition?*, Preprint, Math. Inst., Prague (1995), 1–13.

Kufner, A., John, O. and Fučík, S., *Function spaces*, Academia, Prague, 1977.

Ladyzenskaja, O. A. and Ural'ceva, N. N., *Linear and Quasilinear Equations of Elliptic Type*, Academic Press, New York, 1968.

Larsson, S., Thomée, V. and Zhang, N. Y., *Interpolation of coefficients and transformation of the dependent variable in the finite element methods for the nonlinear heat equation*, Math. Methods Appl. Sci. **11** (1989), 105–124.

Lenoir, M., *Optimal isoparametric finite elements and error estimates for domains involving curved boundaries*, SIAM J. Numer. Anal. **23** (1986), 562–580.

Liu, L., *PhD thesis*, Jyväskylä, 1997.

Liu, L. and Křížek, M., *Finite element analysis of a radiation heat transfer problem*, Preprint, Math. Inst., Prague (1995), 1–11.

Liu, L., Křížek, M. and Neittaanmäki, *Higher order finite element approximation of a quasilinear elliptic boundary value problem of a nonmonotone type*, Appl. Math. **41** (1996), 467–478.

Mackerle, J., *Finite element methods. A guide to information sources*, Elsevier, Amsterdam, 1991.

Matsokin, A. M., *Variational-difference method of solution of elliptic equations in a circle*, Preprint of Computer Center, 13, Novosibirsk, 1975a.

Matsokin, A. M., *Authomatization of triangulation for domains with smooth boundaries for solving equations of the elliptic type*, Preprint of Computer Center, 15, Novosibirsk, 1975b.

Mazja, V. G., *Sobolev spaces*, Springer, Berlin, 1985.

Meyers, N. G., *An example of non-uniqueness in the theory of quasi-linear elliptic equations of second order*, Arch. Rational Mech. Anal. **14** (1963), 177–179.

Mikhlin, S. G., *The problem of the minimum of a quadratic functional*, Holden–Den, Amsterdam, 1965.

Mikhlin, S. G., *Mathematical physics, an advanced course*, North–Holland, Amsterdam, 1970.

Mikhailov, V. P., *Partial differential equations*, Mir, Moscow, 1978.

Mitchell, A. R. and Wait, R., *The finite element method in partial differential equations*, John Wiley & Sons, London, 1975.

Nečas, J., *Les méthodes directes en théorie des équations elliptiques*, Academia, Prague, 1967.

Nedelec, J. C., *Eléments finis mixtes incompressibles pour l'equation de Stokes dans* $\mathbf{R}^3$, Numer. Math. **39** (1982), 97–112.

Noor, A. K., *Bibliography of books and monographs on finite element technology*, Appl. Mech. Rev. **44** (1991), 307–311.

Oganesjan, L. A. and Ruhovec, L. A., *Variation-difference methods for solving elliptic equation*, (in Russian), Academy of Sciences, Erevan, 1979.

Preiningerová, V., Křížek, M. and Kahoun, V., *Temperature distribution in large transformer cores*, Proc. of GANZ Conf., (M. Franyó, ed.), Budapest, 1985, pp. 254–261.

Rektorys, K., *Variational methods in mathematics, science and engineering*, D. Reidel, Dordrecht, 1977.

Schulte, E., *Tilings*, Handbook of convex geometry, Vol. B, North-Holland, Amsterdam, 1993.

Sobolev, S. L., *Applications of functional analysis in mathematical physics*, Providence, Rhode Island, 1963.

Sobolev, S. L., *Introduction to the theory of cubature formulae*, Nauka, Moscow, 1974, (in Russian).

Sobotíková, V., *PhD thesis*, Prague, 1993.

Strang, G., *Variational crimes in the finite element method*, The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, (A. K. Aziz, ed.), Academic Press, New York, 1972, pp. 689–710.

Strang, G. and Fix, G., *An analysis of the finite element method*, Prentice-Hall, Englewood Cliffs, New Jersey, 1973.

Synge, J. L., *The hypercircle in mathematical physics*, Cambridge Univ. Press, Cambridge, 1957.

Taylor, A. E., *Introduction to functional analysis*, John Wiley & Sons, New York, 1958.

Temam, R., *Navier–Stokes equations*, North-Holland, Amsterdam, 3rd revised edn, 1984.

Vladimirov, V. S., *Equations of Mathematical Physics.*, Marcel Dekker, New York, 1971.

Vulih, B. S., *Introduction to the functional analysis*, Nauka, Moscow, 1967, (in Russian).

Weatherill, N. P. and Hassan, O., *Efficient three-dimensional Delaunay triangulations with automatic point creation and imposed boundary constraints*, Internat. J. Numer. Methods Engrg. **37** (1994), 2005–2039.

Yosida, K., *Functional analysis*, Springer, Berlin, 1965.

Ženíšek, A., *The finite element method for nonlinear elliptic equations with discontinuous coefficients*, Numer. Math. **58** (1990a), 51–77.

Ženíšek, A., *Nonlinear elliptic and evolution problems and their finite element approximations*, Academic Press, London, 1990b.

Zienkiewicz, O. C., *The finite element method in engineering science*, 3rd edn, McGraw-Hill, London, 1977.

Zlámal, M., *On the finite element method*, Numer. Math. **12** (1968), 394–409.

Zlámal, M., *Curved elements in the finite element method*, SIAM J. Numer. Anal. **10** (1973), 229–240.