The cascadic algorithm for the Dirichlet problem

Gilyova L.V., Shaidurov V.V.

Introduction

Nowadays the finite element method has become one of the most extensively used and efficient tools to solve a variety of problems in mathematical physics and engineering. Its popularity is due to universality and simplicity of its mathematical formulation for a wide class of problems together with flexibility of numerical algorithms that enables to take into account specific properties of a given problem. To not a smaller degree the successful use of this method is caused by the evolution of high-speed computers and by advances in the projective methods and the approximation theory.

Solving a concrete problem of mathematical physics by the finite element method involves the following major stages:

1. variational (generalized) formulation of the problem;
2. triangulation of a geometric domain (i.e., dividing it into small supports of finite elements of a given form) and specifying a finite element space whose basis is formed by functions with a small support;
3. solving a system of algebraic equations.

In spite of a great and constantly increasing number of works concerned with the finite element method, at each of these stages there arise problems which are not sufficiently investigated from the theoretical point of view.

In the present work the questions, related to the two last stages of implementation of the finite element method to solve some elliptic boundary value problems, are considered.

Applying the finite element method to a boundary value problem leads to a system of algebraic equations of large dimension. A number of equations of such system may be as much as several millions, especially for 3D problems.

Therefore the efficiency of the finite element method essentially depends on a method of solving such algebraic system.

Recently multigrid methods have become a very efficient tool for solving systems of algebraic equations obtained by the finite element method. The idea of the classical multigrid method was first suggested by R.P. Fedorenko in [11]. Then in [12] he proved convergence of this method for the finite-difference analogue of the Poisson equation on a square. N.S. Bakhvalov in [1] proved optimality of the method with respect to the number of arithmetical operations to achieve the accuracy of the same order as the discretization error. According to the asymptotic estimates of efficiency, this method outperformed other known iterative algorithms but for a time its advantages had become veiled because of its complicated logic and cumbersome mathematical substantiation.

At a certain stage of development of the finite element method, invoking new mathematical techniques and software essentially simplified realization and justification of the algorithm. Since the mid-seventies the number of publications on multigrid methods has begun to grow rapidly. We point out the monographs of W. Hackbusch [25] and V.V. Shaidurov [32].

The cascadic iterative algorithm can be considered as the simplest version of multigrid methods, without preconditioning or restriction to a coarser grid. In addition, it holds the advantage of the classical multigrid methods, namely, in a finite number of arithmetical operations per one unknown it enables to achieve the accuracy of the same order as the discretization error.

The cascadic algorithm starts on the coarsest grid where a number of equations of the algebraic system is small enough to solve this system directly or by some iterative method with a high accuracy without substantial computational effort. On finer grids the approximate solution is obtained by an appropriate iterative method (smoother) taking the interpolation of the approximate solution from the previous grid as an initial guess.

The cascadic algorithm with the conjugate-gradient method as a smoother was first presented by P Deuflhard in [8] and [9] for solving elliptic boundary value problems. He demonstrated high efficiency of this algorithm numerically. V.V. Shaidurov in [33], [36] proved the optimal complexity of this method for the 2D Dirichlet problem for the H^2–regular second order elliptic equation. In [34], [5], [6], and [39] the optimal complexity was proved for elliptic problems with reduced regularity caused by the fact that a domain has interior angles greater then π. In [39] special nested triangulations were constructed, resulting in the same order of approximation as for the case of a convex polygonal domain.

Besides, in [5] and [6] the estimates of convergence rate are obtained for other iterative smoother for an elliptic boundary value problem in the 2D and 3D cases. In the 2D case among the smoothers being analyzed only the conjugate-gradient method gave the optimal complexity. By [33] and [36] this result holds true for the Jacobi-type iterations with the special iterative parameters as well. In the 3D case several more smoothers provide the optimal complexity.

The convergence of the cascadic algorithm for some nonlinear elliptic problems is established in [40] and [37].

In the present work applying the cascadic algorithm to some 2D problems (the weakly nonlinear elliptic equation, the indefinite-sign elliptic equation, and the plane elasticity problem) and to the 3D elliptic equation is considered.

In multigrid algorithms the "nestedness" of finite elements, i.e., a possibility to express the basis functions on a coarser grid as linear combinations of several ones on a finer grid, is of importance. This property holds on a polygonal domain and provides simple rules for interpolation and restriction from one grid to another. On a domain with a curvilinear boundary, when refining a triangulation, the approximation of the boundary results in that the standard piecewise linear elements do not satisfy this property near the boundary. V.V. Shaidurov in [31, §5.3] proposed a way to construct basis functions, satisfying this property, in a 2D domain with a smooth curvilinear boundary. In [35] he proved the optimal complexity of the cascadic algorithm when such basis functions are used. In the present work this result is extended to a 3D domain with a smooth curvilinear boundary.

Section 1 deals with 2D problems. To obtain a discrete system the standard piecewise linear finite elements on triangles are used. In Section 1.1 the Dirichlet problem for the weakly nonlinear elliptic equation is considered. On each grid we obtain a nonlinear discrete system. For its linearization we use the Newton method with a "frozen" derivative. On a sequence of grid problems we apply the cascadic algorithm with the conjugate-gradient method or the Jacobi-type iterations with special parameters. It is proved that this algorithm has the optimal complexity. In Section 1.2 we investigate the application of the cascadic algorithm to the indefinite-sign elliptic equation. In this case discretization leads to an algebraic system with a matrix being not positive definite. It is proved that the cascadic algorithm with the Jacobi-type iterations with special iterative parameters is accurate with near-optimal complexity. In Section 1.3 the optimal complexity of the cascadic algorithm with the conjugate-gradient method or the Jacobi-type iterations with special parameters is proved for the plane elasticity problem.

Section 2 is concerned with the 3D Dirichlet problem for the elliptic equation [24]. In Section 2.1 we consider this problem on a convex bounded polyhedron. To construct a discrete system we use the piecewise linear elements on tetrahedra. The optimal complexity of the cascadic algorithm with the conjugate-gradient method or the Jacobi-type iterations with special parameters is proved. In Section 2.2 an algorithm of dividing a 3D domain with a smooth curvilinear boundary into tetrahedra is proposed. The algorithm starts on a given coarsest triangulation. Subsequent finer triangulations are constructed by dividing tetrahedra of the previous level into 8 parts with correction of location of vertices lying near the boundary. It is proved that with the appropriate coarsest triangulation this algorithm gives triangulations of good quality no matter how many times the procedure of dividing was repeated. In Section 2.3 we consider the 3D Dirichlet problem for the elliptic equation on a convex bounded domain with a smooth curvilinear boundary. For triangulation refinement we use the algorithm presented in Section 2.2. We propose a way of constructing basis functions, which ensure embedding the finite element subspaces, and prove that the discretization error is of the same order as for the standard piecewise linear elements on a polyhedron. It is proved that the cascadic algorithm has optimal complexity as well as in Section 2.1.

1 The cascadic algorithm for 2D problems

1.1 The weakly nonlinear elliptic equation

1.1.1 Formulation of the differential problem. On a bounded convex polygon $\Omega \subset R^2$ with a boundary Γ we consider the problem

$$\Delta u = f(x,u) \quad \text{in } \Omega, \tag{1.1}$$

$$u = 0 \quad \text{on } \Gamma \tag{1.2}$$

where the function $f(x,u)$ belongs to $C(\bar{\Omega} \times R)$ and satisfies to the following constraints:

$$0 \le \frac{\partial f}{\partial v}(x,v) \le c_1 \quad \forall\, x \in \Omega,\ \forall\, v \in R, \tag{1.3}$$

$$\left| \frac{\partial^2 f}{\partial v^2}(x,v) \right| \le c_2 \quad \forall\, x \in \Omega,\ \ \forall\, v \in R. \tag{1.4}$$

The problem (1.1), (1.2) admits a unique solution in the class $H_2^2(\Omega)$ [27].

We use the common notations of the Sobolev spaces and norms [27]. $L_2(\Omega)$ is the Hilbert space of all Lebesgue measurable functions on Ω with the inner product and the finite norm given by

$$(u,v)_\Omega = \int\limits_\Omega uv dx, \quad \|u\|_{0,\Omega} = (u,u)^{1/2},\ u,v \in L_2(\Omega).$$

$H_2^k(\Omega)$ is the Hilbert space of all functions $u \in L_2(\Omega)$ whose partial derivatives $\partial^{s+r}u/\partial x_1^s \partial x_2^r$ up to order k also belong to $L_2(\Omega)$ and the norm is defined by the equality

$$\|u\|_{k,\Omega} = \left(\sum_{0 \le s+r \le k} \left\| \frac{\partial^{s+r} u}{\partial x_1^s \partial x_2^r} \right\|_0^2 \right)_\Omega^{1/2}.$$

The subspace $H_0^1(\Omega)$ consists of functions $u \in H_2^1(\Omega)$ with the condition $u = 0$ on Γ in the sense of trace space on Γ.

The problem (1.1),(1.2) may be reduced to the generalized formulation:

$$\begin{aligned} &find\ u \in H_0^1(\Omega)\ such\ that \\ &\mathcal{L}(u, v) = 0 \quad \forall\, v \in H_0^1(\Omega) \end{aligned} \tag{1.5}$$

where the bilinear form is given by

$$\mathcal{L}(u, v) = \int_\Omega (\nabla u \cdot \nabla v + f(x, u)v)dx. \tag{1.6}$$

The problem (1.5) admits a unique solution too [32].

1.1.2 Formulation of the discrete problem. In order to construct the Bubnov-Galerkin scheme, we first divide the initial polygon Ω into a small number of closed triangles such that the resulting subdivision is consistent, i.e., any pair of triangles has either no common points or a common vertex only or a common whole side. Let us denote the maximal length of the sides of all triangles by h_0. Put $N_i = 2^i$, $h_i = h_0/N_i$, and for all $i = 1, \ldots, l$ divide each initial triangle into N_i^2 equal triangles. We denote the set of all vertices of the obtained consistent triangulation by $\mathcal{T}_i$ and introduce $\Omega_i = \bar{\Omega}_i \bigcap \Omega$ as well as the number n_i of points of the set Ω_i. At each node $y \in \Omega_i$ we introduce the basis function $\varphi_y^i \in H_0^1(\Omega)$ which is linear on each triangle of triangulation $\mathcal{T}_i$ and equals 1 at the node y but equals 0 at any other node $z \in \bar{\Omega}_i$. Let us denote the linear span of these functions by

$$\mathcal{H}^i = span\{\varphi_y^i\},\ y \in \Omega_i.$$

Restricting (1.5) to the subspace $\mathcal{H}^i \subset H_0^1(\Omega)$, we get the discrete problem:

$$\begin{aligned} &find\ \tilde{z}_i \in \mathcal{H}^i\ such\ that \\ &\mathcal{L}(\tilde{z}_i, v) = 0 \quad \forall\, v \in \mathcal{H}^i. \end{aligned} \tag{1.7}$$

The problem (1.7) admits a unique solution, moreover, the following estimates hold [32]:

$$\|u - \tilde{z}_i\|_{1,\Omega} \le c_3 h_i \|u\|_{2,\Omega}, \tag{1.8}$$
$$\|u - \tilde{z}_i\|_{0,\Omega} \le c_4 h_i^2 \|u\|_{2,\Omega}. \tag{1.9}$$

The problem (1.7) is nonlinear. To linearize it, we use Newton's method with a "frozen" derivative. Fix the number of level $i \in [1, l]$ and consider one iteration of this method. As an initial guess we take a solution $\tilde{u}_{i-1} \in \mathcal{H}^i$ obtained on the previous level by some iterative process (smoothing operator).

On the coarsest grid the number of equations in (1.7) is rather small, so this system may be solved with appropriate computational complexity. Let us denote the obtained solution by $\tilde{u}_0$. This solution is supposed to be obtained with a sufficiently high accuracy, i.e., this accuracy is of the same order as the final one:

$$\|\tilde{u}_0 - \tilde{z}_0\|_{1,\Omega} \le c_5 h_l \|u\|_{2,\Omega}. \tag{1.10}$$

It is this solution that will be used in linearization on a sequence of nested triangulations. As the result we get the problem:

$$\begin{aligned} &\textit{find } \tilde{v}_i \in \mathcal{H}^i \textit{ such that} \\ &\mathcal{L}_i(\tilde{v}_1, v) = (g(\tilde{u}_{i-1}), v) \quad \forall\, v \in \mathcal{H}^i \end{aligned} \tag{1.11}$$

where

$$\mathcal{L}_i(\tilde{v}_i, v) = \int_\Omega \left(\nabla \tilde{v}_i \cdot \nabla v + \frac{\partial f}{\partial u}(x, \tilde{u}_0) \tilde{v}_i v \right) dx, \tag{1.12}$$

$$(g(\tilde{u}_{i-1}), v) = \int_\Omega \left(-f(x, \tilde{u}_{i-1}) v + \frac{\partial f}{\partial u}(x, \tilde{u}_0) \tilde{u}_{i-1} v \right) dx. \tag{1.13}$$

The problem (1.11) admits a unique solution for any initial guess $\tilde{u}_{i-1}$ [32].

Now we introduce the energy norm for functions

$$|\!|\!| \tilde{v} |\!|\!|_\Omega = \mathcal{L}_1(\tilde{v}, \tilde{v})^{1/2} \quad \forall\, \tilde{v} \in H_0^1(\Omega).$$

Due to (1.3) this norm is equivalent to $\|\tilde{v}\|_1$ on the space $H_0^1(\Omega)$:

$$c_6 \|\tilde{v}\|_{1,\Omega} \le |\!|\!| \tilde{v} |\!|\!|_\Omega \le c_7 \|\tilde{v}\|_{1,\Omega} \quad \forall\, \tilde{v} \in H_0^1(\Omega). \tag{1.14}$$

Let M_i be the n_i-dimensional space of vectors w with components $w(x)$, $x \in \Omega_i$. Then the formulation (1.11) is equivalent to the linear system of algebraic equations

$$L_i v_i = f_i \tag{1.15}$$

where L_i is the matrix with the elements given by

$$L_i(x,y) = \mathcal{L}_1(\varphi_x^i, \varphi_y^i), \quad x, y \in \Omega_i; \tag{1.16}$$

$v_i \in M_i$ is the vector of unknowns; $f_i \in M_i$ is defined by

$$f_i(x) = (g(\tilde{u}_{i-1}), \varphi_x^i), \; x \in \Omega_i.$$

Let us introduce the inner products and norms for vectors $v, w \in M_i$:

$$\begin{aligned}
(v,w)_i &= \sum_{x \in \Omega_i} v(x)w(x), \\
[v,w]_i &= (L_i v, w)_i, \\
\|v\|_i &= (v,v)_i^{1/2}, \\
|\!|\!|v|\!|\!|_i &= [v,v]_i^{1/2}.
\end{aligned}$$

Notice that because of (1.16), (1.12), and (1.3) L_i is symmetric and positive definite.

Now we introduce the interpolation operator $I_i : M_i \to M_{i+1}$. Let a rectilinear segment between two neighbouring nodes x', $x'' \in \bar{\Omega}_i$ be an edge of the i-th triangulation. Then, the interpolation $w = I_i v, \; v \in M_i$ is uniquely defined by the formulae

$$\begin{aligned}
w(x') &= v(x'), \\
w(x'') &= v(x''), \\
w((x'+x'')/2) &= (v(x') + v(x''))/2.
\end{aligned}$$

1.1.3 Formulation of the cascadic algorithm. Now let us formulate the cascadic algorithm. The main objective of the proposed algorithm consists in solving the following problem:

$$\begin{aligned}
&\textit{for a given } f_l \in M_l \textit{ find } v_l \in M_l \textit{ such that} \\
&L_l v_l = f_l.
\end{aligned} \tag{1.17}$$

The main feature of the cascadic algorithm consists in successively solving the problems:

$$\begin{aligned}
&\textit{For a given } f_i \in M_i \textit{ find } v_i \in M_i \textit{ such that} \\
&L_i v_i = f_i, \quad i = 0, 1, \ldots, l.
\end{aligned} \tag{1.18}$$

Let us formulate the cascade algorithm for the considered problem with some abstract iterative process S_i (smoothing operator).

1. *Find* u_0 *as an approximate solution of the problem* (1.7) *for* $i = 0$;
2. *for* $i = 1, 2, \ldots, l$ *do*
   $\{$ 2.1. $w_i = I_{i-1} u_{i-1}$;
   2.2. *set* $u_i = S_i(L_i, w_i, f_i); \}$.

As a smoothing operator one may use the conjugate-gradient method or the Jacobi-type one [36].
Conjugate-gradient method (m_i iterations on level i);
Procedure $S_i(L_i, w_i, f_i)$;

3. $y_0 = w_i;\ p_0 = r_0 = f_i - L_i y_0;\ \sigma_0 = (r_0, r_0)_i$;
4. *for* $k = 1, \ldots, m_i$ *do*
   $\{$ *if* $\sigma_{k-1} = 0$ *then* $\{y_{m_i} = y_{k-1};$ *go to* $5\}$;
   $\alpha_{k-1} = \sigma_{k-1} / (p_{k-1}, L_i p_{k-1})_i$;
   $y_k = y_{k-1} + \alpha_{k-1} p_{k-1}$;
   $r_k = r_{k-1} - \alpha_{k-1} L_i p_{k-1}$;
   $\sigma_k = (r_k, r_k)_i;\ \beta_k = \sigma_k / \sigma_{k-1}$;
   $p_k = r_k + \beta_k p_{k-1} \}$;
5. *set* $S_i = y_{m_i}$.

Jacobi-type iterations (m_i iterations on level i);
Procedure $S_i(L_i, w_i, f_i)$;

3. $y_0 = w_i$;
4. *for* $k = 1, \ldots, m_i$ *do*

$$\left\{ \tau_{k-1} = \frac{1}{\lambda_i^*} \cos^{-2} \frac{\pi(2k-1)}{2(2m_i+1)}; \right.$$

$$\left. y_k = y_{k-1} - \tau_{k-1}(L_i\, y_{k-1} - f_i) \right\} ;$$

5. *set* $S_i = y_{m_i}$.

Here λ_i^* is an upper bound of eigenvalues λ of the matrix L_i:

$$L_i \varphi = \lambda \varphi.$$

Usually this value λ_i^* is determined, in practice, from Gerschgorin's Lemma and satisfies the inequalities

$$\max_{\lambda \in Sp(L_i)} \lambda \le \lambda_i^* \le c_9 \max_{\lambda \in Sp(L_i)} \lambda \tag{1.19}$$

with a constant c_9 independent of i.

Let us fix an integer $i \in [1, l]$ and denote by $\delta_0 = v_i - w_i$ the error of the initial guess w_i with respect to exact solution v_i of the problem (1.18) and by $\delta_1 = v_i - u_i$ the error of the final approximation u_i. Both considered iterative processes (the conjugate-gradient method and the Jacobi-type one) define linear "iterative" operators

$$B_i : \delta_0 \to \delta_1 \text{ and } Q_i : \delta_0 \to \delta_1$$

respectively. Moreover, these operators can be represented as polynomials in L_i with coefficients a_k, τ_k [30].

For the operator B_i the following result is valid [30].

Lemma 1. *Under a fixed initial error δ_0, the operator B_i minimizes the error norm $\|\delta_1\|_i$ among all polynomials in L_i with arbitrary coefficients.*

Besides, the following result has been proved in [36].

Lemma 2. *Let the operators L_i be self-adjoint and positive-definite in $(\cdot, \cdot)_i$. Then*

$$\|Q_i w\|_i \le \|w\|_i, \tag{1.20}$$

$$\|Q_i w\|_i \le \frac{\sqrt{\lambda_i^*}}{2m_i + 1} \|w\|_i \tag{1.21}$$

where λ_i^* is the upper bound of eigenvalues of the operator L_i in the space M_i, and m_i is a number of iterations.

1.1.4 Auxiliary estimates.

Let us define the usual isomorphism between vectors $v \in M_i$ and its functional prolongations $\tilde{v} \in \mathcal{H}^i$

$$\tilde{v}(x) = \sum_{y \in \Omega_i} v(y) \varphi_y^i(x), \quad x \in \bar{\Omega}, \tag{1.22}$$

$$v(y) = \tilde{v}(y), \qquad y \in \Omega_i. \tag{1.23}$$

Due to (1.16) and (1.22) we have for an isomorphic pair $v \in M_i$, $\tilde{v} \in \mathcal{H}^i$

$$\|v\|_i = \|\tilde{v}\|_\Omega. \tag{1.24}$$

The norm $\|\tilde{v}\|_{0,\Omega}$ is equivalent to $\|v\|_i$ with the factor h_i [32]:

$$c_8 h_i \|v\|_i \le \|\tilde{v}\|_{0,\Omega} \le c_9 h_i \|v\|_i. \tag{1.25}$$

It is well-known that the convergence of Newton's method is of quadratic order in the norm $\|\!|\cdot|\!\|_\Omega$. A similar estimate is valid for its modification with a "frozen" derivative (see [22] and [23]).

Lemma 3. *For the solution of the problem* (1.11) *the estimate*

$$\|\!|\tilde{z}_i - \tilde{v}_i|\!\|_\Omega \le \frac{c_2}{2} \|\!|\tilde{z}_i - \tilde{u}_{i-1}|\!\|_\Omega^2 + c_2 \|\!|\tilde{z}_i - \tilde{u}_{i-1}|\!\|_\Omega \|\!|\tilde{z}_i - \tilde{u}_0|\!\|_\Omega \tag{1.26}$$

holds. □

In the same way as in [32] one can prove the following lemma.

Lemma 4. *Assume that the condition* (1.3) *is satisfied. Then the maximal eigenvalue* λ_i^* *of the matrix* L_i *with the entries* (1.16) *is bounded in the following sense:*

$$0 < \lambda_i^* \le c_{10} \tag{1.27}$$

with a constant c_{10} *which is independent of* $\tilde{u}_0,\ i,\ l.$ □

1.1.5 Convergence of the cascadic algorithm.

Now we formulate the main results on convergence without proof which can be found in [22] and [23].

Theorem 5. *Assume that* $1 < \alpha < 2$ *and that the cascadic algorithm is applied with* m_i *steps of the conjugate-gradient method or the Jacobi-type one on each level* $i = 1, \dots, l$*, where* m_i *is the smallest integer satisfying*

$$m_i \ge m_l 2^{\alpha(l-i)}. \tag{1.28}$$

Besides, let h_0^2 *be small in comparison with* h_l *and* h_l *be sufficiently small in comparison with 1, i.e.:*

$$h_0^2 \le c_{14} h_l \ll 1, \quad c_{14} \ge 1. \tag{1.29}$$

Then the algorithm yields an approximate solution u_l *on the highest level with the estimate*

$$\|\!|z_l - u_l|\!\|_l \le c_{15} h_l \tag{1.30}$$

where the constant

$$c_{15} = 2c_{14}\|u\|_{2,\Omega} \left(c_5 c_7 + c_{12} c_{14} \|u\|_{2,\Omega} + \frac{c_{13}}{m_l(1 - 2^{1-\alpha})} \right) \tag{1.31}$$

is independent of l, h_l. □

Let us look at the constant (1.31) in the estimate (1.30). In the linear case we have a constant of the form

$$\frac{c}{m_l(1-2^{1-\alpha})}.$$

It implies decreasing the final error proportionally to m_l. In the nonlinear case we have an additional influence of initial and linearization errors. The former is small enough due to (1.10) and the latter is due to (1.29).

Theorem 6. *Let the hypotheses of Theorem 5 be fulfilled. Then the prolongation $\tilde{u}_l \in \mathcal{H}^l$ of the solution u_l obtained by the cascadic algorithm obeys the estimate*

$$|\!|\!| u - \tilde{u}_l |\!|\!|_\Omega \le c_{17} h_l. \tag{1.32}$$

The computational complexity is estimated from above by the value

$$S_l \le (c_{18} m_l + c_{19}) n_l, \tag{1.33}$$

with constants c_{18}, c_{19} independent of m_l and n_l. □

It should be noted that the upper estimate (1.3) often fails for rather simple functions, e.g., $f(x,u) = -e^{-u}$ or $f(x,u) = u^3$. At the same time, it is known that the problem (1.1)–(1.2) has the bounded solution

$$|u(x)| \le c_{20} \qquad \forall\, x \in \overline{\Omega}.$$

Using the cut-off function [28] one can construct a function $f^*(x,v)$ such that

$$f^*(x,v) = f(x,v) \qquad \forall\, x \in \Omega,\ \ \forall\, v \in [-c_{20}, c_{20}].$$

Now the upper estimate in (1.3) is satisfied. Then the modified problem with the right-hand side f^* is uniquely solvable, and its solution coincides with the bounded solution of the original problem.

1.2 The indefinite-sign elliptic problem

1.2.1 Formulation of the differential problem. On a bounded convex polygon $\Omega \subset R^2$ with a boundary Γ we consider the problem

$$Lu \equiv -\sum_{i,j=1}^{2} \partial_i(a_{ij}\partial_j u) + au = f \text{ in } \ \Omega, \tag{1.34}$$

$$u = 0 \text{ on } \Gamma \tag{1.35}$$

where the coefficients and the right-hand side of (1.34) satisfy the conditions

$$\begin{aligned} & f \in L_2(\Omega), \\ & \partial_i a_{ij} \in L_q(\Omega),\ q > 2,\ i, j = 1, 2;\ a_{12} = a_{21} \text{ on } \overline{\Omega}, \\ & \mu \sum_{i=1}^{2} \xi_i^2 \leq \sum_{i,j=1}^{2} a_{ij}\xi_i\xi_j \leq \nu \sum_{i=1}^{2} \xi_i^2 \text{ on } \overline{\Omega} \quad \forall\, \xi_1, \xi_2 \in R,\ \nu \geq \mu > 0. \end{aligned} \tag{1.36}$$

Regarding the coefficient a we assume that

$$a \in C(\bar{\Omega}). \tag{1.37}$$

Since *a can take negative values,* the conditions (1.36)–(1.37) do not insure unambiguous solvability of the problem. Therefore we assume that the operator of the problem (1.34)–(1.35) is nonsingular, i.e., has no eigenvalue equal to 0. On the basis of [27] this results in unambiguous solvability of (1.34)–(1.35) for any right-hand side $f \in L^2(\Omega)$ and in the estimate

$$\|u\|_{2,\Omega} \leq c_1 \|f\|_{0,\Omega}. \tag{1.38}$$

Formulate for (1.34) – (1.35) the generalized problem:

$$\begin{aligned} & \textit{find } u \in H_0^1(\Omega) \textit{ such that} \\ & \mathcal{L}(u, v) = (f, v)_\Omega \qquad \forall\, v \in H_0^1(\Omega) \end{aligned} \tag{1.39}$$

where

$$\mathcal{L}(u, v) = \int\limits_\Omega \left(\sum_{i,j=1}^{2} a_{ij}\partial_j u \partial_i v + auv \right) dx. \tag{1.40}$$

The problem (1.39) has a unique solution as well [27].

Note that the operator of the problem (1.34)–(1.35) is not positive definite, remaining coercive however. To substantiate this fact, we introduce a positive constant a_0 such that

$$a \geq -a_0 \text{ on } \overline{\Omega} \tag{1.41}$$

and point out that the operator $L + a_0$ on functions u satisfying the condition (1.35) is self-adjoint and positive definite on $H_0^1(\Omega)$ in the following sense:

$$\mathcal{L}(u, u) + a_0(u, u)_\Omega \geq c\|u\|_{1,\Omega}^2. \tag{1.42}$$

1.2.2 Formulation of the discrete problem. To construct the Galerkin scheme, we triangulate the polygon Ω and introduce the basis functions φ_y^i as it was described in subsection 1.1.2

Consider the problem (1.39) on the subspace $\mathcal{H}^i \subset H_0^1(\Omega)$. We obtain the discrete problem:

$$\begin{aligned}&find\ \tilde{v}_i \in \mathcal{H}^i\ satisfying\ the\ equality\\ &\mathcal{L}(\tilde{v}_i, v) = (f, v)_\Omega \quad \forall\, v \in \mathcal{H}^i.\end{aligned} \tag{1.43}$$

Let M_i be the n_i-dimensional space involving vectors w with components $w(x)$. Enumerate the nodes $x \in \Omega_i$ from 1 to n_i. From this point of view to any node $x \in \Omega_i$ there corresponds some number. The problem (1.43) is then equivalent to the system of linear algebraic equations

$$L_i v_i = f_i \tag{1.44}$$

where $v_i \in M_i$ is the vector of unknowns with components $v_i(x),\ \ x \in \Omega_i$; $f_i \in M_i$ is the vector with components $f_i(x) = (f, \varphi_x^i)_\Omega,\ x \in \Omega_i$; L_i is the $n_i \times n_i$ matrix with elements

$$L_i(x, y) = \mathcal{L}(\varphi_x^i, \varphi_y^i) \tag{1.45}$$

where x and y are, respectively, the line and column numbers of the element of L_i coinciding with the numbers of the nodes x and y in Ω_i.

It is clear that the matrix L_i is not positive definite in general and therefore is useless for introduction of an energy norm. Because of this, we shall use the diagonal mass matrix D_i that approximates the second term in the positive definite operator $(L + a_0)$. To do so we put

$$D_i(x) = \frac{a_0}{3} \sum_{T \in \mathcal{T}_i, T \ni x} meas(T),\ x \in \Omega_i \tag{1.46}$$

where the summation is taken over all the triangles T of the triangulation $\mathcal{T}_i$ that have the node x as a vertex. The line and column numbers of the element $D_i(x)$ of the matrix D_i agree with the number of the node x in Ω_i.

A vector $v \in M_i$ is put in correspondence with its functional prolongation in $\mathcal{H}^i$:

$$\tilde{v}(x) = \sum_{y \in \Omega_i} v(y)\varphi_y^i(x), \quad x \in \overline{\Omega}_i. \tag{1.47}$$

It is obvious that

$$v(y) = \tilde{v}(y),\ y \in \Omega_i.$$

Thus we have determined the isomorphism between vectors $v \in M_i$ and functions $\tilde{v} \in \mathcal{H}^i$:

The bilinear form

$$\mathcal{L}_1(u,v) = \mathcal{L}(u,v) + a_0(u,v)_\Omega \tag{1.48}$$

is symmetric and positive definite on $H_0^1(\Omega)$ by virtue of (1.42). Thus we introduce the following energy norm for functions from $H_0^1(\Omega)$:

$$|||v|||_\Omega = \mathcal{L}_1(v,v)^{1/2}.$$

Obviously, it is equivalent to the norm $\|\cdot\|_{1,\Omega}$:

$$c_2\|v\|_{1,\Omega} \le |||v|||_\Omega \le c_3\|v\|_{1,\Omega} \quad \forall\, v \in H_0^1(\Omega). \tag{1.49}$$

We also determine the scalar product and the norm for the vectors:

$$(v,w)_i = \sum_{x\in\Omega_i} D_i(x)v(x)w(x) = v^T D_i w,$$

$$\|v\|_i = \left(\sum_{x\in\Omega_i} D_i(x)v^2(x)\right)^{1/2}, \qquad v,w \in M_i,$$

where the sign ^{T} means transposition. Besides, we introduce the matrix

$$A_i = L_i + D_i$$

and determine the energy norm for vectors

$$|||v|||_i = (v^T A_i v)^{1/2}, \; v \in M_i.$$

For any isomorphic couple $v \in M_i$ and $\tilde{v} \in \mathcal{H}^i$, it can be proved that the norm $\|\tilde{v}\|_{0,\Omega}$ is equivalent to $\|v\|_i$ [32], i.e.,

$$c_4\|v\|_i \le \|\tilde{v}\|_{0,\Omega} \le a_0^{-1/2}\|v\|_i, \tag{1.50}$$

as well as the energy norms:

$$c_6|||v|||_i \le |||\tilde{v}|||_\Omega \le |||v|||_i. \tag{1.51}$$

Introduce the interpolation operator $I_i : M_i \to M_{i+1}$ in the same way as in 1.1.2. Note that the functional prolongations $\tilde{v}$ and $\tilde{w}$ of the vectors $v \in M_i$ and $w = I_i v \in M_{i+1}$ coincide, i.e., $\tilde{v} = \tilde{w}$. Thus the operator I_i corresponds to the identity operator on the subspace $\mathcal{H}^i$ with respect to the isomorphism defined above.

Lemma 7. *Assume that the conditions* (1.36)–(1.37) *are satisfied and the problem* (1.34)–(1.35) *has a unique solution. If h_i is sufficiently small then the Galerkin problem* (1.43) *also has a unique solution $\tilde{v}_i$ which obeys the estimate*

$$\|u - \tilde{v}_i\|_\Omega \le c_8 h_i \|f\|_{0,\Omega}. \tag{1.52}$$

Proof. We use the usual way to prove that the finite-dimensional operator is uniquely solvable. First, we suppose that the problem (1.43) has a solution and derive some estimate for it. Then we consider the (homogeneous) problem with the zero right-hand side which indeed has at least one (trivial) solution and demonstrate that there is no other one. On the basis of the general theory, it means that the finite-dimensional operator is regular and the non-homogeneous problem with an arbitrary right-hand side has a unique solution.

So, let us first suppose that the problem (1.43) has a solution $\tilde{v}_i$. Since $\mathcal{H}^i \subset H_0^1(\Omega)$, from (1.39) it follows that

$$\mathcal{L}(u, v) = (f, v)_\Omega \quad \forall\, v \in \mathcal{H}^i.$$

Subtracting from this the equality (1.43) we get

$$\mathcal{L}(u - \tilde{v}_i, v) = 0 \quad \forall\, v \in \mathcal{H}^i. \tag{1.53}$$

Put $v = w - \tilde{v}_i$ where w is an arbitrary function from $\mathcal{H}^i$ and recast the above relation as

$$\mathcal{L}(u - \tilde{v}_i,\ w - \tilde{v}_i) = 0 \quad \forall\, w \in \mathcal{H}^i.$$

Hence we have

$$\mathcal{L}(u - \tilde{v}_i,\ u - \tilde{v}_i) = \mathcal{L}(u - \tilde{v}_i,\ u - w) \qquad \forall\, w \in \mathcal{H}^i. \tag{1.54}$$

Consider the auxiliary problem:

$$\textit{find } z \in H_0^1(\Omega) \textit{ such that}$$
$$\mathcal{L}(v, z) = (g, v)_\Omega \quad \forall\, v \in H_0^1(\Omega)$$

where $g \in L^2(\Omega)$. From (1.38) we have

$$\|z\|_{2,\Omega} \le c_1 \|g\|_{0,\Omega}. \tag{1.55}$$

Put $g = v = u - \tilde{v}_i$. Then

$$\mathcal{L}(u - \tilde{v}_i, z) = \|u - \tilde{v}_i\|_{0,\Omega}^2. \tag{1.56}$$

Subtracting (1.53) from (1.56), we get

$$\|u-\tilde{v}_i\|^2_{0,\Omega} = \mathcal{L}(u-\tilde{v}_i,\ z-v) \quad \forall\, v \in \mathcal{H}^i. \tag{1.57}$$

Using the Cauchy-Bunyakovskii inequality, we can write

$$\|u-\tilde{v}_i\|^2_{0,\Omega} \le d_1\|u-\tilde{v}_i\|_{1,\Omega}\cdot\|z-v\|_{1,\Omega} \quad \forall\, v \in \mathcal{H}^i. \tag{1.58}$$

Put $v = z_I$ where z_I is the interpolant in $\mathcal{H}^i$ of the function z. It obeys the well-known estimate [7]

$$\|z-z_I\|_{1,\Omega} \le d_2 h_i \|z\|_{2,\Omega}. \tag{1.59}$$

From this and (1.55), it follows that

$$\|z-z_I\|_{1,\Omega} \le c_1 d_2 h_i\ \|u-\tilde{v}_i\|_{0,\Omega}. \tag{1.60}$$

Using this estimate with (1.58) yields

$$\|u-\tilde{v}_i\|^2_{0,\Omega} \le c_1 d_1 d_2 h_i \|u-\tilde{v}_i\|_{0,\Omega}\cdot\|u-\tilde{v}_i\|_{1,\Omega}.$$

This estimate and (1.49) give

$$\|u-\tilde{v}_i\|_{0,\Omega} \le d_3 h_i |\!|\!| u-\tilde{v}_i |\!|\!|_{\Omega} \tag{1.61}$$

where $d_3 = c_1 d_1 d_2 / c_2$.

Estimate from below the left-hand side of (1.54). With (1.61) we have

$$\begin{aligned} &\mathcal{L}(u-\tilde{v}_i,\ u-\tilde{v}_i) = |\!|\!| u-\tilde{v}_i |\!|\!|^2_{\Omega} - a_0\|u-\tilde{v}_i\|^2_{0,\Omega} \\ &\ge |\!|\!| u-\tilde{v}_i |\!|\!|^2_{\Omega} - a_0 d_3^2 h_i^2 |\!|\!| u-\tilde{v}_i |\!|\!|^2_{\Omega} \\ &= (1-a_0 d_3^2 h_i^2)|\!|\!| u-\tilde{v}_i |\!|\!|^2_{\Omega}. \end{aligned} \tag{1.62}$$

Using the Cauchy-Bunyakovskii inequality for the right-hand side of (1.54), we have

$$\mathcal{L}(u-\tilde{v}_i, u-w) \le d_4 |\!|\!| u-\tilde{v}_i |\!|\!|_{\Omega}\cdot |\!|\!| u-w |\!|\!|_{\Omega}.$$

Now we use as w the interpolant u_I in $\mathcal{H}^i$ of the function u. Using the estimate similar to (1.59) together with (1.38) and (1.49), we have

$$\mathcal{L}(u-\tilde{v}_i,\ u-u_I) \le d_5 h_i |\!|\!| u-\tilde{v}_i |\!|\!|_{\Omega}\cdot\|f\|_{0,\Omega}. \tag{1.63}$$

Combining (1.62) with (1.63), we have

$$(1-a_0 d_3^2 h_i^2)|\!|\!| u-\tilde{v}_i |\!|\!|_{\Omega} \le d_5 h_i \|f\|_{0,\Omega}.$$

When $h_i \to 0$, the expression in parentheses is positive. Thus the last inequality dictates the validity of the estimate (1.52) with the constant $c_8 = 2d_5$ when h_i is sufficiently small:

$$h_i \le (2a_0 d_3^2)^{-1/2}. \tag{1.64}$$

Now assume that h_i satisfies (1.64). Consider the Galerkin problem (1.44) with the zero right-hand side Θ_i

$$L_i w_i = \Theta_i \tag{1.65}$$

that corresponds to the zero function f in (1.43). The above reasoning gives the following inequality for the interpolant $\tilde{w}_i$:

$$|||\tilde{w}_i|||_\Omega \le 0.$$

This implies $w_i = \Theta_i$. It means that there is no other solution of the problem (1.65) except the trivial one and operator (matrix) L_i is nonsingular. Therefore either of the two problems (1.43) and (1.44) has a unique solution for any right-hand side. $\square$

It is useful to note that the interpolation operator I_{i-1} connects the stiffness matrices on neighbouring levels:

$$L_{i-1} = I_{i-1}^* L_i I_{i-1} \tag{1.66}$$

where the sign * means conjugation. In particular it results in

$$(I_{i-1}v)^T L_i I_{i-1} v = v^T L_{i-1} v \;\forall\, v \in M_{i-1}. \tag{1.67}$$

The lumped mass matrices D_i and D_{i-1} do not satisfy such property but they are connected by an inequality which is sufficient for our purpose.

Lemma 8. *For any vector $v \in M_{i-1}$ the following inequalities hold:*

$$\begin{aligned} ||I_{i-1}v||_i &\le ||v||_{i-1}, \\ |||I_{i-1}v|||_i &\le |||v|||_{i-1}. \end{aligned} \tag{1.68}$$

Proof. Introduce the isomorphic interpolant $\tilde{v} \in \mathcal{H}^{i-1}$ of a vector v and denote three vertices of an arbitrary triangle $T \in \mathcal{T}_{i-1}$ by $a_{1,T}, a_{2,T}, a_{3,T}$. Besides, use the following inequality for two neighbouring vertices $x', x'' \in \Omega_{i-1}$:

$$\tilde{v}^2((x'+x'')/2) = (\tilde{v}(x') + \tilde{v}(x''))^2/4 \le (\tilde{v}^2(x') + \tilde{v}^2(x''))/2.$$

As the result we have

$$
\begin{aligned}
&(I_{i-1}v, I_{i-1}v)_i \\
&= a_0 \sum_{T\in\mathcal{T}_i} (\tilde{v}^2(a_{1,T}) + \tilde{v}^2(a_{2,T}) + \tilde{v}^2(a_{3,T}))meas(T)/3 \\
&= a_0 \sum_{T\in\mathcal{T}_{i-1}} (\tilde{v}^2(a_{1,T}) + \tilde{v}^2(a_{2,T}) + \tilde{v}^2(a_{3,T}) + 3\tilde{v}^2((a_{1,T} + a_{2,T})/2) \\
&+3\tilde{v}^2((a_{1,T} + a_{3,T})/2) + 3\tilde{v}^2((a_{2,T} + a_{3,T})/2))meas(T)/12 \\
&\le a_0 \sum_{T\in\mathcal{T}_{i-1}} (\tilde{v}^2(a_{1,T}) + \tilde{v}^2(a_{2,T}) + \tilde{v}^2(a_{3,T}))meas(T)/3 \\
&= a_0 v^T D_{i-1} v = (v,v)_{i-1}.
\end{aligned}
$$

Thus, we proved the first statement of Lemma. The second one follows from it due to (1.67) and the definition of the vector norms:

$$
\begin{aligned}
|||I_{i-1}v|||_i^2 &= v^T I_{i-1}^* L_i I_{i-1} v + (I_{i-1}v)^T D_i (I_{i-1}v) \\
&= v^T L_{i-1} v + \|I_{i-1}v\|_i \le v^T L_{i-1} v + \|v\|_i = |||v|||_i. \qquad \square
\end{aligned}
$$

1.2.3 Formulation of the cascadic algorithm. On the sequence of grids $\Omega_i,\ i = 0, 1, \dots, l$, we have obtained the sequence of problems:

$$
\begin{aligned}
&for\ the\ given\ f_i \in M_i,\ find\ v_i \in M_i\ such\ that \\
&L_i v_i = f_i. \qquad (1.69)
\end{aligned}
$$

To solve them, we use the cascadic iterative method with the Jacobi-type smoother and special choice of parameters.
The cascadic algorithm:

1. $u_0 = L_0^{-1} f_0$;
2. *for* $i = 1, 2, \dots, l$ *do*
   { 2.1. $w_i = I_{i-1} u_{i-1}$;
   $y_0 = w_i$;
   2.2. *for* $k = 1, 2, \dots, m_i$ *do*
   $y_k = y_{k-1} - \tau_{k-1} D_i^{-1}(L_i y_{k-1} - f_i)$;
   2.3. $u_i = y_{m_i}$; } .

On each level i, the iterative parameters τ_{k-1} are chosen on the basis of the relationship between the upper and the lower bounds of eigenvalues λ of the following spectral problem in the space M_i:

$$L_i \varphi = \lambda D_i \varphi. \tag{1.70}$$

Assume that we have the upper estimate λ_i^* for these eigenvalues λ. In practice it is found by the Gerschgorin lemma [41] for the matrix $D_i^{-1/2} L_i D_i^{-1/2}$ and satisfies the inequality

$$\max_{\lambda \in Sp(D_i^{-1/2} L_i D_i^{-1/2})} \lambda \le \lambda_i^* \le c_9 \max_{\lambda \in Sp(D_i^{-1/2} L_i D_i^{-1/2})} \lambda. \tag{1.71}$$

Put the lower estimate for these eigenvalues equal to -1. Indeed, if $\lambda < -1$ then the matrix $L_i - \lambda D_i = (L_i + D_i) - (\lambda + 1) D_i$ is positive definite and hence nonsingular. By [32, section 3.6] the estimate

$$\lambda_i^* \le c_9 h_i^{-2} \tag{1.72}$$

holds with a constant independent of level i. Introduce the parameter γ_i dependent on λ_i^* and on the number of iterations m_i

$$\gamma_i = \lambda_i^* \mathrm{sh}^2 (\ln(1 + \sqrt{2})/(2m_i + 1))$$

and consider two cases.

1) Let $\gamma_i \ge 1$.

Then we put

$$\tau_{k-1} = \frac{1}{\lambda_i^*} \cos^{-2} \frac{\pi(2k-1)}{2(2m_i+1)}, \ k = 1, \ldots, m_i. \tag{1.73}$$

2) Consider the opposite situation $\gamma_i < 1$.

Then we put $b_i^* = \max\{1, \lambda_i^*\}$ and take

$$\tau_{k-1} = \frac{1}{b_i^*} \cos^{-1} \frac{\pi(2k-1)}{2(m_i+1)}, \ k = 1, \ldots, m_i/2, \tag{1.74}$$

$$\tau_{k-1} = -\frac{1}{b_i^*} \cos^{-1} \frac{\pi(2k - m_i - 1)}{2(m_i+1)}, \ k = m_i/2 + 1, \ldots, m_i, \tag{1.75}$$

supposing m_i to be even.

1.2.4 An auxiliary operator. Concider the polynomial in $x \in R$

$$q_{m_i}(x) = \prod_{k=1}^{m_i}(1 - \tau_{k-1}x) \tag{1.76}$$

with the parameters (1.73) or (1.74)–(1.75). It has been shown in [32, p.124] that this polynomial with the parameters (1.73) satisfies the inequality

$$|q_{m_i}(x)| \leq ln^{-1}(1+\sqrt{2}) \leq \rho = 1.14 \text{ on } [-\gamma_i, \lambda_i^*] \tag{1.77}$$

and it has the least deviation from zero

$$\max_{x\in[-\gamma_i,\lambda_i^*]} |xq_{m_i}^2(x)| = \frac{\lambda_i^*}{(2m_i+1)^2} \tag{1.78}$$

among all polynomials of the form

$$1 + \sum_{k=1}^{m_i} \beta_k x^k \tag{1.79}$$

with real or complex coefficients β_k. If τ_k are chosen by (1.74)–(1.75) with even m_i, the polynomial (1.76) satisfies the inequality [32, p.125]

$$|q_{m_i}(x)| \leq 1 \text{ on } [-b_i^*, b_i^*] \tag{1.80}$$

and has the following deviation from zero:

$$\max_{x\in[-b_i^*,b_i^*]} |xq_{m_i}^2(x)| = \frac{b_i^*}{m_i+1}. \tag{1.81}$$

Now let us introduce the matrix polynomial

$$Q_{m_i} = q_{m_i}(D_i^{-1}L_i) = \prod_{k=1}^{m_i}(I - \tau_{k-1}D_i^{-1}L_i). \tag{1.82}$$

Lemma 9. *Let L_i and D_i be symmetric and D_i be positive definite. Then*

$$|||Q_{m_i}w|||_i \leq \rho|||w|||_i \quad \forall\, w \in M_i. \tag{1.83}$$

In addition, for the parameters (1.73) *we have*

$$|||Q_{m_i}w|||_i \leq \left(\frac{\lambda_i^*}{(2m_i+1)^2} + \rho^2\right)^{1/2} ||w||_i \tag{1.84}$$

and for (1.74)–(1.75) *we have*

$$|||Q_{m_i}w|||_i \leq \left(\frac{b_i^*}{m_i+1} + 1\right)^{1/2} ||w||_i. \tag{1.85}$$

Proof. The spectral problem (1.70) has n_i real eigenvalues

$$-1 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{n_i} \leq \lambda_i^* \tag{1.86}$$

with eigenvectors $\varphi_1, \ldots, \varphi_{n_i}$ which can be taken to be orthogonal in the following sense:

$$(\varphi_j, \varphi_l)_i = \delta_{jl} \tag{1.87}$$

where δ_{jl} is the Kronecker symbol. Let us decompose an arbitrary vector $w \in M_i$ into the sum of the eigenvectors φ_j:

$$w = \sum_{j=1}^{n_i} \alpha_j \varphi_j \text{ with } \alpha_j = (w, \varphi_j)_i. \tag{1.88}$$

Then the norms of w can be represented by

$$\|w\|_i^2 = \sum_{j=1}^{n_i} \alpha_j^2, \tag{1.89}$$

$$|||w|||_i^2 = \sum_{j=1}^{n_i} (\lambda_j + 1)\alpha_j^2. \tag{1.90}$$

On the other hand, we get

$$|||Q_i w|||_i^2 = \sum_{j=1}^{n_i} (\lambda_j + 1) q_{m_i}^2(\lambda_j)\alpha_j^2. \tag{1.91}$$

Because of (1.77) and (1.78), we have

$$\sum_{j=1}^{n_i} (\lambda_j + 1) q_{m_i}^2(\lambda_j)\alpha_j^2 \leq \left(\frac{\lambda_i^*}{(2m_i + 1)^2} + \rho^2 \right) \sum_{j=1}^{n_i} \alpha_j^2.$$

Together with (1.89), it implies (1.84). If we take (1.80) and (1.81), we get

$$\sum_{j=1}^{n_i} (\lambda_j + 1) q_{m_i}^2(\lambda_j)\alpha_j^2 \leq \left(\frac{b_i^*}{m_i + 1} + 1 \right) \sum_{j=1}^{n_i} \alpha_j^2.$$

Together with (1.89), it implies (1.85). Finally, because of (1.90) and either (1.77) or (1.80) we obtain

$$\sum_{j=1}^{n_i} (\lambda_j + 1) q_{m_i}^2(\lambda_j)\alpha_j^2 \leq \sum_{j=1}^{n_i} (\lambda_j + 1)\rho^2 \alpha_j^2 = \rho^2 |||w|||_i^2$$

that results in (1.83). If the iterative parameters τ_{k-1} are defined by (1.74)–(1.75) the estimate (1.83) holds with $\rho = 1$. □

1.2.5 Convergence of the cascadic algorithm.

Lemma 10. *Under the conditions* (1.36) *there exists a constant* $c^* > 0$ *such that the inequality*

$$\|v_i - I_{i-1}v_{i-1}\|_i \leq c^* h_i |\!|\!| v_i - I_{i-1}v_{i-1} |\!|\!|_i \tag{1.92}$$

holds for any $i = 1, 2, \dots, l$.

Proof. Consider the auxiliary problem:

$$\begin{aligned} &find\ w \in H_0^1(\Omega) \ \ such\ that \\ &\mathcal{L}(w, v) = (\tilde{v}_i - \tilde{v}_{i-1}, v)_\Omega \ \ \forall\, v \in H_0^1(\Omega). \end{aligned} \tag{1.93}$$

According to (1.38) we have

$$\|w\|_{2,\Omega} \leq c_1 \|\tilde{v}_i - \tilde{v}_{i-1}\|_{0,\Omega}.$$

Consider the Galerkin scheme for (1.93):

$$\mathcal{L}(\tilde{w}_{i-1}, v) = (\tilde{v}_i - \tilde{v}_{i-1}, v)_\Omega \quad \forall\, v \in \mathcal{H}^{i-1}.$$

By Lemma 7 we have the estimate

$$|\!|\!| w - \tilde{w}_{i-1} |\!|\!|_\Omega \leq c_8 h_{i-1} \|\tilde{v}_i - \tilde{v}_{i-1}\|_{0,\Omega}. \tag{1.94}$$

Putting in (1.93) $v = \tilde{v}_i - \tilde{v}_{i-1}$, we obtain

$$\mathcal{L}(w, \tilde{v}_i - \tilde{v}_{i-1}) = \|\tilde{v}_i - \tilde{v}_{i-1}\|_{0,\Omega}^2. \tag{1.95}$$

Since $\mathcal{H}^{i-1} \subset \mathcal{H}^i$, from (1.43) we have

$$\mathcal{L}(\tilde{v}_i, v) = (f, v)_\Omega \quad \forall\, v \in \mathcal{H}^{i-1}.$$

Subtracting the identity

$$\mathcal{L}(\tilde{v}_{i-1}, v) = (f, v)_\Omega \quad \forall\, v \in \mathcal{H}^{i-1}$$

from the above, we have

$$\mathcal{L}(\tilde{v}_i - \tilde{v}_{i-1}, v) = 0 \quad \forall\, v \in \mathcal{H}^{i-1}. \tag{1.96}$$

Put $v = \tilde{w}_{i-1}$ and, taking into account the symmetry of $\mathcal{L}$, subtract this equality from (1.95):

$$\mathcal{L}(w - \tilde{w}_{i-1}, \tilde{v}_i - \tilde{v}_{i-1}) = \|\tilde{v}_i - \tilde{v}_{i-1}\|_{0,\Omega}^2. \tag{1.97}$$

Using the Cauchy-Bunyakovskii inequality yields

$$\mathcal{L}(w-\tilde{w}_{i-1},\tilde{v}_i-\tilde{v}_{i-1}) \le d_1 \|w-\tilde{w}_{i-1}\|_\Omega \cdot \|\tilde{v}_i-\tilde{v}_{i-1}\|_\Omega. \tag{1.98}$$

On the basis of (1.94) from (1.97) and (1.98) it follows that

$$\|\tilde{v}_i-\tilde{v}_{i-1}\|^2_{0,\Omega} \le d_1 c_8 h_{i-1}\|\tilde{v}_i-\tilde{v}_{i-1}\|_{0,\Omega}\cdot\|\tilde{v}_i-\tilde{v}_{i-1}\|_\Omega.$$

From (1.50) and (1.51) we have

$$\begin{aligned} &c_4\|v_i-I_{i-1}v_{i-1}\|_i \le \|\tilde{v}_i-\tilde{v}_{i-1}\|_{0,\Omega}\\ &\le d_1c_8h_{i-1}\|\tilde{v}_i-\tilde{v}_{i-1}\|_\Omega \le d_1c_8h_{i-1}\|v_i-I_{i-1}v_{i-1}\|_i. \end{aligned} \tag{1.99}$$

Hence the estimate (1.92) follows with the constant $c^* = 2d_1c_8/c_4$. □

Now we are in a position to prove the main estimate.

Theorem 11. *Let* (1.36)–(1.37) *be fulfilled and* h_0 *be small enough. Then*

$$\|v_i-u_i\|_i \le \rho^i\|v_0-u_0\|_0+\sum_{j=1}^{i}\rho^{i-j}(c_1^*\rho h_j+c_2^*(m_j))\|v_j-I_{j-1}v_{j-1}\|_j \tag{1.100}$$

where

$$c_2^*(m_j) = \begin{cases} d_2/(2m_j+1) \text{ if } \gamma_j \ge 1,\\ d_2/\sqrt{m_j+1} \text{ if } \gamma_j < 1, \end{cases} \tag{1.101}$$

and c_1^*, d_2 *are independent of* h_j, m_j, v_j, j.

Proof. Let us denote the error on level j by

$$\varepsilon_j = v_j-u_j \quad \forall\, j=0,1,\ldots,l.$$

In accordance with the formulation of the cascadic algorithm, we have

$$\|\varepsilon_j\|_j = \|Q_j(v_j-w_j)\|_j \le \|Q_j(v_j-I_{j-1}v_{j-1})\|_j+\|Q_jI_{j-1}\varepsilon_{j-1}\|_j. \tag{1.102}$$

Let $\gamma_j \ge 1$, then due to (1.84), (1.72), (1.92) we get

$$\begin{aligned} &\|Q_j(v_j-I_{j-1}v_{j-1})\|_j \le \left(\frac{\lambda_j^*}{(2m_j+1)^2}+\rho^2\right)^{1/2}\|v_j-I_{j-1}v_{j-1}\|_j\\ &\le c^*(c_9(2m_j+1)^{-2}+\rho^2h_j^2)^{1/2}\|v_j-I_{j-1}v_{j-1}\|_j. \end{aligned}$$

Because of inequality $(a^2+b^2)^{1/2} \le a+b$ for $a,b\ge 0$, we have

$$\|Q_j(v_j-I_{j-1}v_{j-1})\|_j \le (c_1^*\rho h_j+c_2^*(m_j))\|v_j-I_{j-1}v_{j-1}\|_j \tag{1.103}$$

with the constants

$$c_1^* = c^* \text{ and } d_2 = c^* c_9^{1/2}. \tag{1.104}$$

For $\gamma_i < 1$ the estimate (1.103) is valid with $\rho = 1$.

To estimate the second term in (1.102), we use (1.83) and (1.68):

$$|\!|\!| Q_j I_{j-1} \varepsilon_{j-1} |\!|\!|_j \le \rho |\!|\!| I_{j-1} \varepsilon_{j-1} |\!|\!|_j \le \rho |\!|\!| \varepsilon_{j-1} |\!|\!|_{j-1}.$$

Using this inequality with (1.103) in (1.102), we get

$$|\!|\!| \varepsilon_j |\!|\!|_j \le (c_1^* \rho h_j + c_2^*(m_j)) |\!|\!| v_j - I_{j-1} v_{j-1} |\!|\!|_j + \rho |\!|\!| \varepsilon_{j-1} |\!|\!|_{j-1}. \tag{1.105}$$

The estimate (1.105) is valid for $\gamma_j < 1$ with $\rho = 1$. Then we use induction on j in order to get (1.100). $\square$

Formulate the estimate (1.100) in a more convenient form. From (1.51) and (1.52) we have

$$\begin{aligned} &|\!|\!| v_j - I_{j-1} v_{j-1} |\!|\!|_j \le c_6^{-1} |\!|\!| \tilde{v}_j - \tilde{v}_{j-1} |\!|\!|_\Omega \\ &\le c_6^{-1} (|\!|\!| \tilde{v}_j - u |\!|\!|_\Omega + |\!|\!| u - \tilde{v}_{j-1} |\!|\!|_\Omega) \le 3 c_6^{-1} c_8 h_j \|f\|_{0,\Omega}. \end{aligned}$$

Then (1.100) can be recast with the constant $d^* = 3c_6^{-1} c_8$ as

$$|\!|\!| v_i - u_i |\!|\!|_i \le \rho^i |\!|\!| u_0 - v_0 |\!|\!|_0 + d^* \sum_{j=1}^{i} \rho^{i-j} (c_1^* \rho h_j + c_2^*(m_j)) h_j \|f\|_{0,\Omega}. \tag{1.106}$$

Among these inequalities, the estimate relating to the finest triangulation $\mathcal{T}_l$ is most useful. It results in the following statement.

Theorem 12. *Assume that the conditions* (1.36)–(1.38), (1.41) *are satisfied for the problem* (1.34)–(1.35) *on a bounded convex polygon* Ω. *Then for sufficiently small* h_0 *the estimate*

$$|\!|\!| v_l - u_l |\!|\!|_l \le \rho^l |\!|\!| u_0 - v_0 |\!|\!|_0 + c_3^* \sum_{j=1}^{l} \rho^{l-j} (c_1^* \rho h_j + c_2^*(m_j)) h_j \tag{1.107}$$

holds where the constant $c_3^* = d^* \|f\|_{0,\Omega}$ *is independent of* h_j, m_j, l. $\square$

1.2.6 Optimization of the number of iterations. By analyzing the sequence of computations in view of the sparsity of the matrices L_i, the upper estimate of the number of arithmetic operations in the cascadic algorithm is established as follows:

$$S_l = d_2^* \sum_{j=1}^{l} (m_j + d_3^*) n_j + d_4^*. \tag{1.108}$$

Here constants d_2^*, d_3^*, d_4^* are independent of n_j and m_j.

Put $m_l = m$ and choose m_i on the preceding levels as the least integer satisfying the inequalities

$$m_i \geq m\rho^{(l-i)/2} 2^{3(l-i)/2}, \quad i = 1, 2, \ldots, l-1, \tag{1.109}$$

then test the validity of the conditions $\gamma_i \geq 1$, $i = 1, \ldots, l$. Because of (1.72) and (1.109) γ_i increases. Therefore for h_l small enough on fine grids we have $\gamma_i \geq 1$ while $\gamma_i < 1$ may be fulfilled only on coarse grids up to the level $k < l$. In this case we choose m_i as the least integer satisfying the inequalities

$$m_i \geq m\rho^{(l-i)/2} 2^{3(l-i)/2}, \quad i = k+1, \ldots, l-1, \tag{1.110}$$

$$m_i \geq mk^2 m 2^{2(l-i)}, \quad i = 1, \ldots, k. \tag{1.111}$$

Taking the number of iteration in this way, we get useful theoretical estimates. Let us introduce the small positive quantity

$$\gamma = \log_2 \rho < 0.19.$$

Theorem 13. *Assume that the conditions* (1.36)–(1.38),(1.41) *are satisfied for the problem* (1.34)–(1.35) *on a bounded convex polygon Ω. Let h_0 and the initial error in the cascadic algorithm be small enough:*

$$h_0 \leq \min\{c_{10} h_l^{(1+\gamma)/(2+\gamma)}, (2a_0 d_3^2)^{-1/2}\}, \tag{1.112}$$

$$|\!|\!| u_0 - v_0 |\!|\!|_0 \leq c_{11} h_l^{1+\gamma} / h_0^{\gamma}. \tag{1.113}$$

Then for any fixed m with the iteration numbers m_j from (1.109) *or* (1.110)–(1.111), *we have the estimates for the differences between the Galerkin exact solution v_l and the cascadic one u_l:*

$$|\!|\!| u_l - v_l |\!|\!|_l \leq d_5^* h_l \|f\|_{0,\Omega} \tag{1.114}$$

and between the functional prolongation $\tilde{u}_l \in \mathcal{H}^l$ of u_l and the exact solution u:

$$|\!|\!| u - \tilde{u}_l |\!|\!|_\Omega = d_6^* h_l \|f\|_{0,\Omega}. \tag{1.115}$$

The number of arithmetic operations is evaluated from above by the value

$$S_l \leq (d_7^* \log_2^3 n_k + d_8^*) m n_l, \tag{1.116}$$

where $k \leq l$ is a number of levels on which $\gamma_i < 1$.

Proof. From the Euler formula for polygons we have $n_{j-1} \leq n_j/4$. Therefore

$$n_j \leq 4^{j-1} n_l. \tag{1.117}$$

According to the construction of grids

$$h_j = 2^{l-j} h_l. \tag{1.118}$$

Let m_j be taken from (1.109) and $\gamma_i \geq 1, \quad i = 1, \ldots, l$. Taking into account (1.113), (1.109) and (1.118), we obtain from (1.107)

$$\begin{aligned} \|\!|v_l - u_l|\!\|_l &\leq 2^{\gamma l} c_{11} h_l (h_l/h_0)^{\gamma} \\ &+ d \sum_{j=1}^{l} 2^{\gamma(l-j)} \left(c_1^* 2^{\gamma} h_j^2 + \frac{d_2}{2m} \frac{2^{l-j}}{2^{(\gamma+3)(l-j)/2}} h_l \right) \|f\|_{0,\Omega} \\ &\leq 2^{\gamma l} c_{11} h_l 2^{-\gamma l} + d^* \|f\|_{0,\Omega} \left(c_1^* h_0^2 2^{\gamma(l+1)} \sum_{j=1}^{l} 2^{-j(\gamma+2)} \right. \\ &\left. + \frac{d_2}{2m} h_l \sum_{j=1}^{l} 2^{(j-l)(1-\gamma)/2} \right). \end{aligned} \tag{1.119}$$

Due to (1.118) and (1.112) we have

$$h_l = 2^{-l} h_0 \leq 2^{-l} c_{10} h_l^{(1+\gamma)/(2+\gamma)},$$

i.e.,

$$h_l^{1/(2+\gamma)} \leq c_{10} 2^{-l}.$$

Using (1.112) and the above inequality yields the estimate

$$2^{\gamma l} h_0^2 \leq 2^{\gamma l} c_{10}^2 h_l^{1 + \frac{\gamma}{(2+\gamma)}} \leq c_{10}^{2+\gamma} h_l. \tag{1.120}$$

Replacing in (1.119) the sums of two geometrical progressions with infinite series sums and taking into account (1.120), we end up with

$$\begin{aligned} \|\!|v_l - u_l|\!\|_l &\leq c_{11} h_l + d^* \|f\|_{0,\Omega} \left(c_1^* c_{10}^{2+\gamma} \frac{2^{\gamma}}{2^{2+\gamma} - 1} h_l \right. \\ &\left. + \frac{d_2}{2m} \frac{\sqrt{2^{1-\gamma}}}{\sqrt{2^{1-\gamma}} - 1} h_l \right), \end{aligned} \tag{1.121}$$

i.e., (1.114) holds with the constant

$$d_5^* = c_{11} \|f\|_{0,\Omega}^{-1} + d^* \left(c_1^* c_{10}^{2+\gamma} \frac{2^{\gamma}}{2^{2+\gamma} - 1} + \frac{d_2}{2m} \frac{\sqrt{2^{1-\gamma}}}{\sqrt{2^{1-\gamma}} - 1} \right).$$

Let us evaluate the number of arithmetic operations. Remember that the number of iterations m_j is chosen as the least integer satisfying (1.109), so

$$m_j \le m2^{(\gamma+3)(l-j)/2} + 1$$

holds. Using the above inequality together with (1.117) in (1.108) yields

$$\begin{aligned} S_l &\le d_2^* \sum_{j=1}^{l} \left(m2^{(\gamma+3)(l-j)/2} + 1 + d_3^* \right) 2^{2(j-l)} n_l + d_4^* \\ &\le d_2^* \left(m n_l \sum_{j=1}^{l} 2^{(j-l)(1-\gamma)/2} + (d_3^* + 1) n_l \sum_{j=1}^{l} 4^{j-l} \right) + d_4^* \qquad (1.122) \\ &\le d_2^* \left(m \frac{\sqrt{2^{1-\gamma}}}{\sqrt{2^{1-\gamma}} - 1} n_l + 4/3(d_3^* + 1) n_l \right) + d_4^*. \end{aligned}$$

Hence the estimate (1.116) holds with the constants

$$d_7^* = 0, \quad d_8^* = d_2^* \left(\frac{\sqrt{2^{1-\gamma}}}{\sqrt{2^{1-\gamma}} - 1} + \frac{4}{3}(d_3^* + 1) \right) + d_4^*.$$

Now consider the case $\gamma_i < 1,\ i = 1, \ldots, k;\ \ \gamma_i \ge 1,\ i = k+1, \ldots, l$ and m_i are chosen from (1.110)–(1.111). In much the same way as (1.122) was obtained, from (1.107) we get the estimate

$$\begin{aligned} \| v_l - u_l \|_l &\le \rho^{l-k} \| u_0 - v_0 \|_0 + c_3^* \left(\sum_{j=1}^{k} (c_1 h_j + c_2^*(m_j)) \right) h_j \\ &+ c_3^* \left(\sum_{j=k+1}^{l} \rho^{l-j} (c_1^* \rho h_j + c_2^*(m_j)) \right) h_j \qquad (1.123) \\ &\le c_{11} h_l + d^* \left(c_1^* c_{10}^{2+\gamma} \frac{2^\gamma}{2^{2+\gamma} - 1} + \frac{d_2}{2m} \frac{\sqrt{2^{1-\gamma}}}{\sqrt{2^{1-\gamma}} - 1} \right) h_l \|f\|_{0,\Omega} \\ &+ c_3^* d_2 \sum_{j=1}^{k} \frac{h_j}{\sqrt{m_j + 1}}. \end{aligned}$$

Due to (1.111) and (1.118) we have

$$\sum_{j=1}^{k} \frac{h_j}{\sqrt{m_j + 1}} \le \sum_{j=1}^{k} \frac{2^{l-j} h_l}{\sqrt{k^2 m 2^{2(l-j)}}} = \frac{h_l}{\sqrt{m}}. \qquad (1.124)$$

Because of (1.123) and (1.124) we conclude that (1.114) is valid with the constant

$$d_5^* = c_{11}\|f\|_{0,\Omega}^{-1} + d^* \left(c_1^* c_{10}^{2+\gamma} \frac{2^\gamma}{2^{2+\gamma}-1} + \frac{d_2}{2m} \frac{\sqrt{2^{1-\gamma}}}{\sqrt{2^{1-\gamma}}-1} + \frac{d_2}{\sqrt{m}} \right).$$

Let us evaluate the number of arithmetic operations. As far as m_i are chosen as the least integer satisfying (1.110),(1.111), the inequalities

$$m_i \le m 2^{(\gamma+3)(l-i)/2} + 1, \quad i = k+1, \dots, l;$$
$$m_i \le k^2 m 2^{2(l-i)} + 1, \quad i = 1, \dots, k$$

hold. Taking into account these inequalities together with (1.117) in (1.108), we get

$$S_l \le d_2^* \left(\sum_{j=1}^{k} k^2 m 2^{2(l-j)} 2^{2(j-l)} n_l \right.$$
$$\left. + \sum_{j=k+1}^{l} m 2^{(\gamma+3)(l-j)/2} 2^{2(j-l)} n_l \right) + d_2^*(d_3^*+2) \sum_{j=1}^{l} 2^{2(j-l)} n_l + d_4^*$$
$$\le d_2^* \left(k^3 + \frac{\sqrt{2^{1-\gamma}}}{\sqrt{2^{1-\gamma}}-1} \right) m n_l + d_2^*(d_3^*+2)\frac{4}{3} n_l + d_4^*$$
$$\le d_2^* \left(\log_2^3 n_k + \frac{\sqrt{2^{1-\gamma}}}{\sqrt{2^{1-\gamma}}-1} + \frac{4}{3}(d_3^*+2) + \frac{d_4^*}{d_2^*} \right) m n_l,$$

i.e., the estimate (1.114) is valid. □

1.3 The plane elasticity problem

1.3.1 Formulation of the differential problem. Consider the plane elasticity problem on a bounded convex polygon $\Omega \subset R^2$ with a boundary Γ:

$$-\mu \Delta \boldsymbol{u} - (\lambda + \mu) \operatorname{grad} \operatorname{div} \boldsymbol{u} = \boldsymbol{f} \quad \text{in} \quad \Omega, \tag{1.125}$$
$$\boldsymbol{u} = 0 \quad \text{on } \Gamma \tag{1.126}$$

where $\lambda, \mu > 0$ are the Lame coefficients, $\boldsymbol{u}$ is the unknown vector-function of displacement

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix},$$

and $\boldsymbol{f}$ is the given vector-function of mass forces with two components

$$\boldsymbol{f} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}.$$

Introduce the inner product and the norm for vector-functions of $(L_2(\Omega))^2$:

$$(\boldsymbol{u}, \boldsymbol{v})_\Omega = \int_\Omega \boldsymbol{u} \cdot \boldsymbol{v} \, dx, \qquad \|\boldsymbol{u}\|_{0,\Omega} = (\boldsymbol{u}, \boldsymbol{u})_\Omega^{1/2}.$$

Assume that

$$f_1, f_2 \in L_2(\Omega). \tag{1.127}$$

Then due to [27] there exists a unique solution of the problem (1.125)–(1.126) such that

$$u_1, u_2 \in H_2^2(\Omega). \tag{1.128}$$

Introduce the norm for vector-functions of $(H_2^m(\Omega))^2$

$$\|\boldsymbol{u}\|_{m,\Omega} = \left(\sum_{i=1}^{2} \|u_i\|_{m,\Omega}^2 \right)^{1/2}.$$

Under the conditions (1.127) the problem (1.125)–(1.126) obeys the estimate

$$\|\boldsymbol{u}\|_{2,\Omega} \le c_1 \|\boldsymbol{f}\|_{0,\Omega}. \tag{1.129}$$

In accordance with [27] we formulate for (1.125)–(1.126) the generalized problem:

$$\begin{aligned} &\textit{find } \boldsymbol{u} \in (H_0^1(\Omega))^2 \textit{ satisfying the equality} \\ &\mathcal{L}(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v})_\Omega \ \forall \boldsymbol{v} \in (H_0^1(\Omega))^2 \end{aligned} \tag{1.130}$$

where the bilinear form $\mathcal{L}$ is defined by the relation

$$\begin{aligned} \mathcal{L}(\boldsymbol{u}, \boldsymbol{v}) = \int_\Omega & \left\{ 2\mu \left(\frac{\partial u_1}{\partial x_1} \frac{\partial v_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} \frac{\partial v_2}{\partial x_2} \right) \right. \\ & + \lambda \left(\frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} \right) \left(\frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2} \right) \\ & \left. + \mu \left(\frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1} \right) \left(\frac{\partial v_1}{\partial x_2} + \frac{\partial v_2}{\partial x_1} \right) \right\} dx. \end{aligned} \tag{1.131}$$

For $\boldsymbol{f} \in (L_2(\Omega))^2$ the problem (1.130) also has a unique solution [13], [27].

1.3.2 Formulation of the discrete problem. To construct the Bubnov-Galerkin scheme, we triangulate the polygon Ω and introduce basis functions in the same way as in Subsection 1.1.2. Denote by $\mathcal{H}^i$ the linear span of functions $\varphi^i_y, \;\; y \in \Omega_i$.

Consider the problem (1.130) on the subspace $\boldsymbol{H}^i = (\mathcal{H}^i)^2 \subset (H^1_0(\Omega))^2$. We obtain the discrete problem:

$$\begin{aligned}&\textit{find } \boldsymbol{v}_i \in \boldsymbol{H}^i \textit{ satisfying the equality}\\ &\mathcal{L}(\boldsymbol{v}_i, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v})_\Omega \;\; \forall \boldsymbol{v} \in \boldsymbol{H}^i.\end{aligned} \tag{1.132}$$

Let M_i be the $2n_i$-dimensional space consisting of vectors W with n_i components $W(x) = \begin{bmatrix} W_1\ (x) \\ W_2\ (x) \end{bmatrix}, \quad x \in \Omega_i$. Then the problem (1.132) is equivalent to the block system of linear algebraic equations

$$L_i V_i = F_i \tag{1.133}$$

where $V_i \in M_i$ is the vector of unknowns with n_i components

$$V_i(x) = \begin{bmatrix} V_{i,1}(x) \\ V_{i,2}(x) \end{bmatrix}, \quad x \in \Omega_i; \; F_i \in M_i$$

has n_i components

$$F_i(x) = \begin{bmatrix} F_{i,1}(x) \\ F_{i,2}(x) \end{bmatrix} = \begin{bmatrix} (f_1,\varphi^i_x) \\ (f_2,\varphi^i_x) \end{bmatrix}, \quad x \in \Omega_i;$$

L_i is the block matrix of dimension $2n_i \times 2n_i$ with 2×2 blocks

$$L_{i,x,y} = \begin{pmatrix} \mathcal{L}(\varphi^i_x,0;\varphi^i_y,0) & \mathcal{L}(0,\varphi^i_x,\varphi^i_y,0) \\ \mathcal{L}(\varphi^i_x,0;\,0,\,\varphi^i_y) & \mathcal{L}(0,\varphi^i_x;\,0,\,\varphi^i_y) \end{pmatrix}, \quad x, y \in \Omega_i. \tag{1.134}$$

From (1.131) it is evident that L_i is symmetric. Using bilinearity of the functional $\mathcal{L}$ and its positive definiteness [31], we can show that L_i is positive definite and, hence, nonsingular.

A vector $V \in M_i$ is put in correspondence with its functional prolongation in $\boldsymbol{H}^i$:

$$\boldsymbol{v}(x) = \sum_{y \in \Omega_i} \begin{bmatrix} V_1\ (y) \\ V_2\ (y) \end{bmatrix} \varphi^i_y(x), \; x \in \overline{\Omega}. \tag{1.135}$$

It is obvious that

$$V(y) = \boldsymbol{v}(y), \; y \in \Omega_i.$$

Thus, we have determined the isomorphism between vectors $V \in M_i$ and vector-functions $\boldsymbol{v} \in \boldsymbol{H}^i$.

Introduce the energy norm for vector-functions

$$|||\boldsymbol{v}|||_{\Omega} = \mathcal{L}(\boldsymbol{v}, \boldsymbol{v})^{1/2}, \ \boldsymbol{v} \in (H_0^1(\Omega))^2,$$

as well as the scalar product and the norms for vectors

$$(V, W)_i = \sum_{x \in \Omega_i} V(x)W(x) = \sum_{x \in \Omega_i} V_1(x)W_1(x) + \sum_{x \in \Omega_i} V_2(x)W_2(x),$$

$$\|V\|_i = \left\{ \sum_{x \in \Omega_i} |V(x)|^2 \right\}^{1/2} = \left\{ \sum_{x \in \Omega_i} (V_1(x))^2 + \sum_{x \in \Omega_i} (V_2(x))^2 \right\}^{1/2},$$

$$|||V|||_i = (L_i V, V)_i^{1/2}, \quad V, W \in M_i.$$

Taking into account (1.134), (1.135) and bilinearity of $\mathcal{L}$, for an isomorphic couple $V \in M_i, \ \boldsymbol{v} \in \boldsymbol{H}^i$ we have

$$(L_i V, V)_i = \left(\begin{bmatrix} \mathcal{L}(\sum\limits_{x \in \Omega_i} V_1(x)\varphi_x^i, \sum\limits_{x \in \Omega_i} V_2(x)\varphi_x^i; \varphi_y^i, 0) \\ \mathcal{L}(\sum\limits_{x \in \Omega_i} V_1(x)\varphi_x^i, \sum\limits_{x \in \Omega_i} V_2(x)\varphi_x^i; 0, \varphi_y^i) \end{bmatrix}, \begin{bmatrix} v^1(y) \\ v^2(y) \end{bmatrix} \right)_i$$

$$= \mathcal{L}(\sum_{x \in \Omega_i} V_1(x)\varphi_x^i, \sum_{x \in \Omega_i} V_2(x)\varphi_x^i; \sum_{y \in \Omega_i} V_1(y)\varphi_y^i, \sum_{y \in \Omega_i} V_2(y)\varphi_y^i) = \mathcal{L}(\boldsymbol{v}, \boldsymbol{v}),$$

i.e.,

$$|||V|||_i = |||\boldsymbol{v}|||_{\Omega}. \tag{1.136}$$

When studying one elliptic equation in [36], for a vector Z of dimension n_i with elements $Z(x), \ x \in \Omega_i$ the norm

$$\|Z\|_i = (\sum_{x \in \Omega_i} (Z(x))^2)^{1/2}$$

was introduced. It is equivalent with the multiplier h_i to the norm $\|z\|_{0,\Omega}$ of the functional prolongation $z \in H^i$ [31], i.e.,

$$d_1 h_i \|Z\|_i \leq \|z\|_{0,\Omega} \leq d_2 h_i \|Z\|_i.$$

Taking into account the equality

$$\|\boldsymbol{v}\|_{0,\Omega}^2 = \sum_{j=1}^{2} \|v_j\|_{0,\Omega}^2,$$

we see that for an isomorphic couple $V \in M_i$ and $\boldsymbol{v} \in \boldsymbol{H}^i$ the norm $\|V\|_i$ is also equivalent to the norm $\|\boldsymbol{v}\|_{0,\Omega}$ with the multiplier h_i:

$$c_2 h_i \|V\|_i \leq \|\boldsymbol{v}\|_{0,\Omega} \leq c_3 h_i \|V\|_i. \tag{1.137}$$

Introduce the interpolation operator $I_i : M_i \to M_{i+1}$ as follows. Let $x', x'' \in \Omega_i$ be two neighboring nodes of the triangulation $\mathcal{T}_i$. Then the interpolation $W = I_i V,\ V \in M_i$ is defined by the formulae

$$W(x') = V(x'), \quad W(x'') = V(x''),$$
$$W\left(\frac{x' + x''}{2}\right) = \frac{V(x') + V(x'')}{2}.$$

Note that the prolongations of the vectors V and W coincide, i.e., $\boldsymbol{v} = \boldsymbol{w}$. Thus, the operator I_i corresponds to the identity operator on the subspace $\boldsymbol{H}^i$ with respect to the isomorphism defined above.

The convergence of the Bubnov-Galerkin solution to the exact one has been proved in [17] and [18].

Lemma 14. *For $\boldsymbol{f} \in (L_2(\Omega))^2$ the problem* (1.132) *has a unique solution. It obeys the estimate*

$$|||\boldsymbol{u} - \boldsymbol{v}_i|||_\Omega \le c_4 h_i \|\boldsymbol{f}\|_{0,\Omega}. \qquad \square \tag{1.138}$$

Note that the eigenvalues of L_i obey the estimate

$$0 < \lambda_i^* \le c_6 \tag{1.139}$$

where λ_i^* is the maximum eigenvalue of L_i [32].

Let us sum up. On the sequence of grids $\Omega_i,\ i = 0, \ldots, l$ we obtained the sequence of problems:

for given $F_i \in M_i$ find $V_i \in M_i$ such that

$$L_i V_i = F_i. \tag{1.140}$$

For their sequentially solving we use the cascadic iterative method.

1.3.3 Formulation of the cascadic algorithm. We first formulate the cascadic algorithm with some abstract iterative process S_i (smoothing operator).

The cascadic algorithm:

1. $U_0 = L_0^{-1} F_0$;
2. *for* $i = 1, 2, \ldots, l$ *do*
   { 2.1. $W_i = I_{i-1} U_{i-1}$;
   2.2. *Set* $U_i = S_i(L_i, W_i, F_i)$; } .

We consider two iterative processes as smoothing operators.
The conjugate-gradient method (m_i iterations);
procedure $S_i(L_i, W_i, F_i)$:

$$\begin{aligned}
&3.\ Y_0 = W_i;\quad P_0 = R_0 = F_i - L_iY_0;\quad \sigma_0 = (R_0, R_0)_i;\\
&4.\ \textit{for } k = 1, 2, \ldots, m_i \textit{ do}\\
&\qquad \{\ \textit{if } \sigma_{k-1} = 0,\ \textit{then } \{\ Y_{m_i} = Y_{k-1};\ \textit{go to}\ \ 5;\ \}\\
&\qquad\ \ \alpha_{k-1} = \sigma_{k-1}/(P_{k-1}, L_iP_{k-1})_i;\\
&\qquad\ \ Y_k = Y_{k-1} + \alpha_{k-1}P_{k-1};\\
&\qquad\ \ R_k = R_{k-1} - \alpha_{k-1}L_iP_{k-1};\\
&\qquad\ \ \sigma_k = (R_k, R_k)_i;\ \beta_k = \sigma_k/\sigma_{k-1};\\
&\qquad\ \ P_k = R_k + \beta_kP_{k-1};\};\\
&5.\ \textit{set } S_i = Y_{m_i}.
\end{aligned} \tag{1.141}$$

The Jacobi-type method (m_i iterations);
procedure $S_i(L_i, W_i, F_i)$:

$$\begin{aligned}
&3.\ Y_0 = W_i;\\
&4.\ \textit{for } k = 1, 2, \ldots, m_i \textit{ do}\\
&\qquad \left\{ \tau_{k-1} = \frac{1}{\Lambda_i^*}\cos^{-2}\frac{\pi(2k-1)}{2(2m_i+1)}\,; \right.\\
&\qquad\qquad \left. Y_k = Y_{k-1} - \tau_{k-1}(L_iY_{k-1} - F_i); \right\};\\
&5.\ \textit{set } S_i = Y_{m_i}.
\end{aligned} \tag{1.142}$$

Here Λ_i^* is the upper bound of eigenvalues λ of the operator L_i in the space M_i: $L_i\Phi = \lambda\Phi$. In the Jacobi-type method the numerical value of this quantity is required. It can be found by Gerschgorin's Lemma [41] and satisfies the inequality

$$\lambda_i^* = \max_{\lambda\in Sp(L_i)} \lambda \le \Lambda_i^* \le c_1 \max_{\lambda\in Sp(L_i)} \lambda = c_1\lambda_i^*. \tag{1.143}$$

In the conjugate-gradient method it is supposed to equal λ_i^*, i.e., (1.143) is fulfilled with the constant $c_1 = 1$.

1.3.4 Convergence of the cascadic algorithm and optimization of the number of iterations. In [17] and [18] the estimate

$$\|\!|V_i - U_i|\!\|_i \le c^* c_4 \sum_{j=1}^{i} \frac{1}{2m_j + 1} h_{j-1} \|\boldsymbol{f}\|_{0,\Omega} \tag{1.144}$$

has been proved where V_i is the exact solution of the algebraic problem (1.133), U_i is its approximation obtained by the cascadic algorithm, and m_j is the number of iterations in the conjugate-gradient method (1.141) or in the Jacobi-type one (1.142) on level j. These two methods differ from each other only in the value of the constant c^*.

Among these inequalities for $i = 1, \ldots, l$, the estimate relating to the finest triangulation $\mathcal{T}_l$ is most useful.

Theorem 15. *Assume that for the problem* (1.125)–(1.126) *on a bounded convex polygon Ω the condition* (1.127) *holds. Then for the solution U_l of the cascadic algorithm with one of the iterative smoothers* (1.141) *or* (1.142) *on each level $j = 1, \ldots, l$, we have the estimate*

$$\|\!|V_l - U_l|\!\|_l \le d_1 \sum_{j=1}^{l} \frac{h_{j-1}}{2m_j + 1} \tag{1.145}$$

where $d_1 = c^* c_4 \|\boldsymbol{f}\|_{0,\Omega}$. □

By analyzing the sequence of computations in view of the sparsity of the matrices L_i, the upper estimate of the number of arithmetic operations in the cascadic algorithm is established as follows:

$$S_l = d_2 \sum_{j=1}^{l} (m_j + d_3) n_j + d_4. \tag{1.146}$$

Here the constants d_2, d_3, d_4 are independent of n_j and m_j but different for iterative processes (1.141) and (1.142). It is obvious that these constants are smaller for the latter process.

We now propose to choose the number of iterations $m_1, \ldots, m_{l-1}$ to minimize S_l as a function of $m_1, \ldots, m_{l-1}$ when the right-hand side of (1.145) is fixed. Applying the Lagrange multiplier method gives

$$2m_i + 1 = (2m_l + 1)\sqrt{n_l h_{i-1} / n_i h_{l-1}}.$$

This equality gives noninteger m_i. Therefore we put $m_l = m$ and choose m_i on the preceding levels as the least integer satisfying the inequality

$$2m_i + 1 \ge (2m + 1)\sqrt{n_l h_{i-1} / n_i h_{l-1}}. \tag{1.147}$$

Then the following result is valid (see [17] and [18]).

Theorem 16. *Under the hypotheses of Theorem 15 the error of the cascadic algorithm with the conjugate-gradient method* (1.141) *or the Jacobi-type one* (1.142) *is estimated as*

$$\|U_l - V_l\|_l \leq \frac{c_3 h_l}{2m+1} \|\boldsymbol{f}\|_{0,\Omega}. \tag{1.148}$$

The functional prolongation $\boldsymbol{u}_l \in \boldsymbol{H}^l$ *of the vector* U_l *obeys the estimate*

$$\|\boldsymbol{u}_l - \boldsymbol{u}\|_\Omega \leq h_l \left(c_4 + \frac{c_3}{2m+1} \right) \|\boldsymbol{f}\|_{0,\Omega}. \tag{1.149}$$

The number of arithmetic operations is estimated from above by the value

$$S_l \leq (c_5 m + c_6) n_l \tag{1.150}$$

with constants $c_3 - c_6$ *independent of* m *and* n_l. □

It is apparent that the number of iterations m on the highest level should be chosen according to the condition $c_4 \approx c_3/(2m+1)$. Although the constants c_3 and c_4 are unknown, it is seen that m is independent of the number of levels and the number of unknowns. Therefore (1.148)–(1.150) characterize the following property: in a finite number of arithmetic operations per one unknown, the error of the iterative process is of the same order as the discretization error.

2 The cascadic algorithm for the 3D Dirichlet problem

2.1 The 3D Dirichlet problem on a polyhedron

2.1.1 Formulation of the differential problem. Let us consider the Dirichlet problem on a bounded convex polyhedron $\Omega \subset R^3$ with a boundary Γ:

$$-\sum_{i,j=1}^{3} \partial_i (a_{ij} \partial_j u) + au = f \ \text{ in } \ \Omega, \tag{2.1}$$

$$u = 0 \ \text{ on } \ \Gamma. \tag{2.2}$$

The coefficients and the right-hand side of the equation satisfy the conditions

$$
\begin{gathered}
\partial_i a_{ij} \in L_3(\Omega), \;\; i,j = 1,2,3; \\
a_{ij} = a_{ji} \;\text{ on }\; \overline{\Omega}, \;\; i,j = 1,2,3; \\
\mu \sum_{i=1}^{3} \xi_i^2 \leq \sum_{i,j=1}^{3} a_{ij}\xi_i\xi_j \leq \nu \sum_{i=1}^{3} \xi_i^2 \;\text{ on }\; \overline{\Omega} \;\; \forall\, \xi_i \in R, \;\; \nu \geq \mu > 0; \\
a, f \in L_2(\Omega), \;\; a \geq 0 \;\text{ on }\; \overline{\Omega}.
\end{gathered} \tag{2.3}
$$

Under these conditions the problem (2.1)–(2.2) is uniquely solvable in $H_2^2(\Omega)$ and obeys the estimate [27]

$$
\|u\|_{2,\Omega} \leq c_1 \|f\|_{0,\Omega}. \tag{2.4}
$$

Formulate for (2.1)–(2.2) the generalized problem:

$$
\begin{aligned}
&\textit{find } u \in H_0^1(\Omega) \;\; \textit{satisfying the equality} \\
&\mathcal{L}(u,v) = (f,v)_\Omega \qquad \forall\, v \in H_0^1(\Omega),
\end{aligned} \tag{2.5}
$$

where the bilinear form is defined by the relation

$$
\mathcal{L}(u,v) = \int_\Omega \Big(\sum_{i,j=1}^{3} a_{ij}\partial_j u \partial_i v + auv\Big)dx. \tag{2.6}
$$

The conditions (2.3) are sufficient for the problem (2.5) to be uniquely solvable.

2.1.2 Formulation of the discrete problem. To construct the Bubnov-Galerkin scheme, we carry out a spatial triangulation of the polyhedron Ω [31]. First we divide it into a small number of *closed* regular tetrahedra. We call this subdivision consistent if each two tetrahedra are mutually disjointed or have either one common vertex, or common entire edge, or common entire face. Denote the maximal length of the edges by h_0. Put $N_i = 2^i$ and for all $i = 1, \ldots, l$ divide every initial tetrahedron into N_i^3 equivolumed tetrahedra. For this purpose, complete every initial tetrahedron with two tetrahedra of the same volume to obtain a trihedral prism (Fig. 1). Divide each of its edges into N_i equal parts and pass, through the obtained points, 4 families of planes parallel to the faces. As the result, we get N_i^3 prisms similar to the initial one with the coefficient $1/N_i$. Each of them is further divided into 3 tetrahedra, just like the initial prism. Of

$3N_i^3$ tetrahedra obtained, we retain only N_i^3 tetrahedra forming the initial tetrahedron. Finally, we have 3 kinds of tetrahedra of equal volume but with the edges of different length. The maximum edge length is not over $h_i = \sqrt{2}h_0/N_i$. Assume that the initial subdivsion is consistent. Then the obtained subdivision is also consistent.

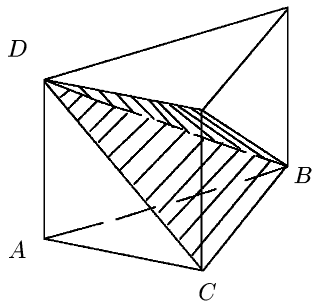


Fig. 1. The complement of a tetrahedron $ABCD$ up to a prism.

Denote by $\overline{\Omega}_i$ the union of all vertices of tetrahedra obtained by dividing the edges of the initial tetrahedra into N_i parts and introduce $\Omega_i = \overline{\Omega}_i \bigcap \Omega$ as well as the number n_i of points of the set Ω_i. For every point $y \in \Omega_i$ we introduce the basis function $\varphi_y^i \in H_0^1(\Omega)$ which equals 1 at the node y, 0 at the other nodes of $\overline{\Omega}_i$ and is linear on each tetrahedron of the i-th subdivision. Denote by $\mathcal{H}^i$ the linear span of the functions φ_y^i, $y \in \Omega_i$.

Restricting (2.5) to the subspace $\mathcal{H}^i \subset H_0^1(\Omega)$, we get the discrete problem:

$$\begin{aligned}&\textit{find}\ \ \tilde{v}_i \in \mathcal{H}^i \ \ \textit{such that}\\ &\mathcal{L}(\tilde{v}_i, v) = (f, v)_\Omega \quad \forall\, v \in \mathcal{H}^i.\end{aligned} \tag{2.7}$$

Let M_i be the n_i-dimensional space of vectors w with components $w(x)$, $x \in \Omega_i$. Then (2.7) is equivalent to the linear system of algebraic equations

$$L_i v_i = f_i \tag{2.8}$$

where $v_i \in M_i$ is the vector of unknowns with components $v_i(y)$, $y \in \Omega_i$; $f_i \in M_i$ is defined by $f_i(x) = (f, \varphi_x^i)_\Omega$, $x \in \Omega_i$; L_i is the $n_i \times n_i$ matrix with the elements given by

$$L_i(x, y) = \mathcal{L}(\varphi_x^i, \varphi_y^i). \tag{2.9}$$

Because of (2.6) the matrix L_i is self-adjoint. As far as the bilinear form $\mathcal{L}$ is positive definite [31], L_i is also positive definite and, hence, nonsingular.

A vector $v \in M_i$ is put in correspondence with its prolongation in $\mathcal{H}^i$ by the following way:

$$\tilde{v}(x) = \sum_{y \in \Omega_i} v(y)\varphi_y^i(x), \quad x \in \overline{\Omega}. \tag{2.10}$$

It is obvious that

$$v(y) = \tilde{v}(y), \quad y \in \Omega_i. \tag{2.11}$$

Thus, we have determined the usual isomorphism between vectors $v \in M_i$ and functions $\tilde{v} \in \mathcal{H}^i$.

Introduce the energy norm for functions

$$|\!|\!| v |\!|\!|_\Omega = \mathcal{L}(v,v)^{1/2}, \ v \in \ H_0^1(\Omega)$$

as well as the scalar product and the norms for vectors

$$(v,w)_i = \sum_{x \in \Omega_i} v(x)w(x),$$

$$\|v\|_i = \left(\sum_{x \in \Omega_i} v^2(x) \right)^{1/2},$$

$$|\!|\!| v |\!|\!|_i = (L_i v, v)^{1/2}, \ v, w \in M_i.$$

Due to (2.10) and (2.11) we have for an isomorphic couple $v \in M_i$, $\tilde{v} \in \mathcal{H}^i$

$$|\!|\!| v |\!|\!|_i = |\!|\!| \tilde{v} |\!|\!|_\Omega. \tag{2.12}$$

The norm $\|\tilde{v}\|_{0,\Omega}$ is equivalent to the norm $\|v\|_i$ with the factor $h_i^{3/2}$ [31]:

$$c_2 h_i^{3/2} \|v\|_i \le \|\tilde{v}\|_{0,\Omega} \le c_3 h_i^{3/2} \|v\|_i. \tag{2.13}$$

Let us introduce the interpolation operator $I_i : M_i \to M_{i+1}$ in the same way as it was done in the subsection 1.1.2.

The convergence of the Bubnov-Galerkin solution to the exact one is proved in [7].

Lemma 17. *Under the conditions* (2.3) *the solution of* (2.7) *does exist, is uniquely determined, and obeys the estimate*

$$|\!|\!| u - \tilde{v}_i |\!|\!|_\Omega \le c_4 h_i \|f\|_{0,\Omega}. \qquad \square \tag{2.14}$$

Note that the maximum eigenvalue λ_i^* of the matrix L_i under the conditions (2.3) satisfies the estimate [31]

$$0 < \lambda_i^* \le c_5 h_i \quad \forall\, i = 1, \dots, l. \tag{2.15}$$

Thus, on the sequence of grids Ω_i, $i = 0, 1, \dots, l$, we obtained the sequence of problems:

$$\begin{aligned} &\textit{for given } f_i \in M_i \textit{ find } v_i \in M_i \textit{ such that} \\ &\qquad L_i v_i = f_i. \end{aligned} \tag{2.16}$$

To solve them we use the cascadic iterative method.

2.1.3 Formulation of the cascadic algorithm.

First we formulate the cascadic algorithm with some abstract iterative process S_i (smoother).

The cascadic algorithm:

$$\begin{aligned} &1.\ u_0 = L_0^{-1} f_0; \\ &2.\ \textit{for } i = 1, 2, \dots, l \textit{ do} \\ &\qquad \{\ 2.1.\ w_i = I_{i-1} u_{i-1}; \\ &\qquad 2.2.\ \textit{set } u_i = S_i(L_i, w_i, f_i); \}. \end{aligned}$$

We consider two iterative processes as smoothers:

The conjugate-gradient method (m_i iteration steps on level i);

Procedure $S_i(L_i, w_i, f_i)$;

$$\begin{aligned} &3.\ y_0 = w_i;\ p_0 = r_0 = f_i - L_i y_0;\ \sigma_0 = (r_0, r_0)_i; \\ &4.\ \textit{for } k = 1, \dots, m_i \textit{ do} \\ &\qquad \{\ \textit{if } \sigma_{k-1} = 0 \textit{ then } \{y_{m_i} = y_{k-1};\ \textit{go to } 5\}; \\ &\qquad \alpha_{k-1} = \sigma_{k-1}/(p_{k-1}, L_i p_{k-1})_i; \\ &\qquad y_k = y_{k-1} + \alpha_{k-1} p_{k-1}; \\ &\qquad r_k = r_{k-1} - \alpha_{k-1} L_i p_{k-1}; \\ &\qquad \sigma_k = (r_k, r_k)_i;\ \beta_k = \sigma_k/\sigma_{k-1}; \\ &\qquad p_k = r_k + \beta_k p_{k-1} \}; \\ &5.\ \textit{set } S_i = y_{m_i}. \end{aligned} \tag{2.17}$$

The Jacobi-type iterations (m_i iteration steps on level i);
Procedure $S_i(L_i, w_i, f_i)$;

$$\begin{aligned}
&3.\ y_0 = w_i;\\
&4.\ \textit{for}\ k = 1, \dots, m_i\ \textit{do}\\
&\qquad \left\{ \tau_{k-1} = \frac{1}{\Lambda_i^*} \cos^{-2} \frac{\pi(2k-1)}{2(2m_i+1)}; \right.\\
&\qquad \left. \quad y_k = y_{k-1} - \tau_{k-1}(L_i\, y_{k-1} - f_i) \right\};\\
&5.\ \textit{set}\ S_i = y_{m_i}.
\end{aligned} \tag{2.18}$$

Here Λ_i^* is the upper estimate of eigenvalues of the operator L_i in the space M_i, i.e., $L_i\varphi = \lambda\varphi$. In the Jacobi-type iterations the explicit value of this quantity is required. It is found from Gerschgorin's Lemma [41] and satisfies the inequality

$$\lambda_i^* = \max_{\lambda \in Sp(L_i)} \lambda \le \Lambda_i^* \le c_6 \max_{\lambda \in Sp(L_i)} \lambda = c_6 \lambda_i^*. \tag{2.19}$$

In the conjugate-gradient method Λ_i^* is supposed to equal λ_i^*, i.e., (2.19) is fulfilled with the constant $c_6 = 1$.

2.1.4 Convergence of the cascadic algorithm.

In [33] the convergence criterion has been formulated as follows:

$$\textit{there exists a constant } c^* > 0 \textit{ such that } \forall\, i = 1, \dots, l$$
$$\|v_i - I_{i-1}v_{i-1}\|_i \le \frac{c^*}{\sqrt{\Lambda_i^*}} |\!|\!| v_i - I_{i-1}v_{i-1} |\!|\!|_i \tag{2.20}$$

where Λ_i^* is the introduced above upper estimate of the eigenvalues of L_i.

Lemma 18. *Under the conditions* (2.3) *the criterion* (2.20) *is fulfilled with the constant* $c^* = 2c_4\sqrt{c_5}/c_2$.

Proof. Let us consider the auxiliary problem:

$$\begin{aligned}
&\textit{find}\ \ w \in H_0^1(\Omega)\ \ \textit{such that}\\
&\mathcal{L}(w, v) = (\tilde{v}_i - \tilde{v}_{i-1}, v)_\Omega \quad \forall\, v \in H_0^1(\Omega).
\end{aligned} \tag{2.21}$$

According to (2.4) we have

$$\|w\|_{2,\Omega} \le c_1 \|\tilde{v}_i - \tilde{v}_{i-1}\|_{0,\Omega}.$$

Consider the Bubnov-Galerkin problem:

$$\begin{aligned}&\textit{find}\ \ \tilde{w}_{i-1} \in \mathcal{H}^{i-1}\ \ \textit{such that}\\ &\mathcal{L}(\tilde{w}_{i-1}, v) = (\tilde{v}_i - \tilde{v}_{i-1}, v)_\Omega \ \ \forall\, v \in \mathcal{H}^{i-1}.\end{aligned}$$

From Lemma 17 we have the estimate

$$|||w - \tilde{w}_{i-1}|||_\Omega \le c_4 h_{i-1} \|\tilde{v}_i - \tilde{v}_{i-1}\|_{0,\Omega}. \tag{2.22}$$

Setting in (2.21) $v = \tilde{v}_i - \tilde{v}_{i-1}$ yields

$$\mathcal{L}(w, \tilde{v}_i - \tilde{v}_{i-1}) = \|\tilde{v}_i - \tilde{v}_{i-1}\|^2_{0,\Omega}. \tag{2.23}$$

Since each function from $\mathcal{H}^{i-1}$ is contained in $\mathcal{H}^i$, from (2.7) we have

$$\mathcal{L}(\tilde{v}_i, v) = (f, v)_\Omega \qquad \forall\, v \in \mathcal{H}^{i-1}.$$

Subtracting from the above equality the identity

$$\mathcal{L}(\tilde{v}_{i-1}, v) = (f, v)_\Omega \qquad \forall\, v \in \mathcal{H}^{i-1},$$

we get

$$\mathcal{L}(\tilde{v}_i - \tilde{v}_{i-1}, v) = 0 \qquad \forall\, v \in \mathcal{H}^{i-1}. \tag{2.24}$$

Setting $v = \tilde{w}_{i-1}$ and taking into account the symmetry of $\mathcal{L}$, we obtain by subtracting this expression from (2.23):

$$\mathcal{L}(w - \tilde{w}_{i-1}, \tilde{v}_i - \tilde{v}_{i-1}) = \|\tilde{v}_i - \tilde{v}_{i-1}\|^2_{0,\Omega}.$$

Applying the Cauchy-Bunyakovskii inequality and (2.22) we get

$$\begin{aligned}&\|\tilde{v}_i - \tilde{v}_{i-1}\|^2_{0,\Omega} \le |||w - \tilde{w}_{i-1}|||_\Omega \cdot |||\tilde{v}_i - \tilde{v}_{i-1}|||_\Omega\\ &\le c_4 h_{i-1} \|\tilde{v}_i - \tilde{v}_{i-1}\|_{0,\Omega} \cdot |||\tilde{v}_i - \tilde{v}_{i-1}|||_\Omega.\end{aligned}$$

From (2.12) and (2.13) it follows that

$$\begin{aligned}&c_2 h_i^{3/2} \|v_i - I_{i-1} v_{i-1}\|_i \le \|\tilde{v}_i - \tilde{v}_{i-1}\|_{0,\Omega}\\ &\le c_4 h_{i-1} |||\tilde{v}_i - \tilde{v}_{i-1}|||_\Omega = c_4 h_{i-1} |||v_i - I_{i-1} v_{i-1}|||_i.\end{aligned}$$

Using (2.19) together with (2.15) yields the inequality

$$\|v_i - I_{i-1} v_{i-1}\|_i \le 2 \frac{c_4}{c_2} \sqrt{\frac{c_5}{\Lambda_i^*}}\, |||v_i - I_{i-1} v_{i-1}|||_i. \qquad \square$$

Since the operators L_i are self-adjoint and positive definite in the scalar product $(\cdot,\cdot)_i$ and

$$L_{i-1} = I_{i-1}^* L_i I_{i-1},$$

from [33] (Theorem 4.1) the following estimate holds:

$$\|v_i - u_i\|_i \le c^* \sum_{j=1}^{i} \frac{1}{2m_j + 1} \|v_j - I_{j-1} v_{j-1}\|_j. \tag{2.25}$$

Here v_i is the exact solution of the algebraic system (2.8), u_i is its approximation obtained by the cascadic algorithm, m_j is the number of iterations in the conjugate-gradient method (2.17) or in the Jacobi-type one (2.18) on the j-th level. These two methods differ from one another only in the value of the constant c^*.

To derive a more convenient formulation of (2.25), let us prove the inequality

$$\|\tilde{v}_j - \tilde{v}_{j-1}\|_\Omega \le \|u - \tilde{v}_{j-1}\|_\Omega. \tag{2.26}$$

Setting $v = \tilde{v}_i - \tilde{v}_{i-1}$ in (2.5) and (2.7), we have

$$\mathcal{L}(u, \tilde{v}_i - \tilde{v}_{i-1}) = \mathcal{L}(\tilde{v}_i, \tilde{v}_i - \tilde{v}_{i-1}).$$

Taking into account the symmetry of $\mathcal{L}$ and setting $v = \tilde{v}_{i-1}$ in (2.24), subtract (2.24) from the both sides of the above equality:

$$\mathcal{L}(u - \tilde{v}_{i-1}, \tilde{v}_i - \tilde{v}_{i-1}) = \mathcal{L}(\tilde{v}_i - \tilde{v}_{i-1}, \tilde{v}_i - \tilde{v}_{i-1}).$$

Apply the Cauchy-Bunyakovskii inequality to the left-hand side:

$$\|\tilde{v}_i - \tilde{v}_{i-1}\|_\Omega^2 \le \|u - \tilde{v}_{i-1}\|_\Omega \cdot \|\tilde{v}_i - \tilde{v}_{i-1}\|_\Omega.$$

From this we have (2.26), so (2.25) may be rewritten as

$$\|v_i - u_i\|_i \le c^* \sum_{j=1}^{i} \frac{1}{2m_j + 1} \|u - \tilde{v}_{j-1}\|_\Omega$$

or due to (2.14) as

$$\|v_i - u_i\|_i \le c^* c_4 \sum_{j=1}^{i} \frac{h_{j-1}}{2m_j + 1} \|f\|_{0,\Omega}. \tag{2.27}$$

Among these inequalities the estimate related to the finest grid is the most useful. We formulate it as a theorem.

Theorem 19. *Let the conditions* (2.3) *be fulfilled for the problem* (2.1)–(2.2) *on a bounded convex polyhedron Ω. Then for the solution u_l of the cascadic algorithm with the smoother* (2.17) *or* (2.18) *on every level $j = 1, \ldots, l$ the estimate*

$$|||v_l - u_l|||_l \le d_1 \sum_{j=1}^{l} \frac{h_j}{2m_j + 1} \tag{2.28}$$

holds where $d_1 = c^* c_4 \|f\|_{0,\Omega}$. □

2.1.5 Optimization of the number of iterations. By analyzing the sequence of computations in view of sparsity of the matrices L_i, the upper bound of the number of arithmetic operations in the cascadic algorithm is determined as

$$S_l = d_2 \sum_{j=1}^{l} (m_j + d_3) n_j + d_4 \tag{2.29}$$

with the constants d_2, d_3, d_4 independent of n_j and m_j but different for the iterative processes (2.17) and (2.18). These constants are certainly smaller for the latter one.

Choose the number of iterations $m_1, \ldots, m_{l-1}$ to minimize S_l as a function of $m_1, \ldots, m_{l-1}$ for a fixed right-hand side of the inequality (2.28). Applying the Lagrange multiplier method, we obtain

$$2m_i + 1 = (2m_l + 1)\sqrt{n_l h_{i-1}/n_i h_{l-1}}.$$

The exact solution of this equality gives non-integer m_i. Therefore we put $m_l = m$ and choose m_i on the preceding levels as the least integer satisfying the inequality

$$2m_i + 1 \ge (2m + 1)\sqrt{n_l h_{i-1}/n_i h_{l-1}}. \tag{2.30}$$

Theorem 20. *Under the hypotheses of Theorem 19 the error of the cascadic algorithm with the conjugate-gradient method* (2.17) *or the Jacobi-type iterations* (2.18) *as a smoother is estimated as*

$$|||u_l - v_l|||_l \le \frac{c_7 h_l}{2m + 1} \|f\|_{0,\Omega}. \tag{2.31}$$

The prolongation $\tilde{u}_l \in \mathcal{H}^l$ of the vector u_l obeys the estimate

$$|||\tilde{u}_l - u|||_\Omega \le h_l \left(c_4 + \frac{c_7}{2m + 1} \right) \|f\|_{0,\Omega}. \tag{2.32}$$

The number of arithmetic operations is evaluated from above by

$$S_l \leq (c_8 m + c_9) n_l \tag{2.33}$$

with constants c_4, c_7–c_9 *independent of* m *and* n_l.

Proof. For the Dirichlet problem we have $n_{j-1} \leq n_j/8$. Therefore

$$n_j \leq 8^{j-l} n_l. \tag{2.34}$$

According to the mesh refinement

$$h_j = 2^{l-j} h_l. \tag{2.35}$$

Using these relations together with (2.30) in (2.28) we get

$$\begin{aligned} \| v_l - u_l \|_l &\leq 2 d_1 h_1 \frac{1}{2m+1} \sum_{j=1}^{l} \sqrt{\frac{n_j h_{j-1}}{n_l h_{l-1}}} \\ &\leq 2 d_1 h_l \frac{1}{2m+1} \sum_{j=1}^{l} 2^{(j-l)} \leq \frac{4 h_l d_1}{2m+1}. \end{aligned}$$

Therefore (2.31) holds with the constant $c_7 = 4c^* c_4$.

Due to the triangle inequality, the equivalence of the norms (2.12), and the estimate (2.14),

$$\begin{aligned} \| \tilde{u}_l - u \|_\Omega &\leq \| \tilde{u}_l - \tilde{v}_l \|_\Omega + \| \tilde{v}_l - u \|_\Omega \\ &\leq \| u_l - v_l \|_l + c_4 h_l \| f \|_{0,\Omega}. \end{aligned}$$

Together with the inequality (2.31) already proved it gives (2.32).

To evaluate the number of arithmetic operations we remind that m_i is chosen as the least integer satisfying the condition (2.30). Therefore

$$2(m_i - 1) + 1 \leq (2m+1)\sqrt{n_l h_{i-1} / n_i h_{l-1}},$$

hence

$$m_i \leq (m + 1/2)\sqrt{n_l h_{i-1} / n_i h_{l-1}} + 1/2. \tag{2.36}$$

Applying this inequality to (2.29) gives

$$S_l \leq d_2 \sum_{j=1}^{l} ((m+1/2) n_l \sqrt{n_j h_{j-1} / n_l h_{l-1}} + (d_3 + 1/2) 8^{j-l} n_l) + d_4.$$

Taking into account (2.34) and (2.35) we get

$$S_l \leq d_2 \sum_{j=1}^{l} ((m+1/2)n_l 2^{(j-l)} + (d_3+1/2)8^{j-l} n_l) + d_4.$$

Replace the sums of two geometrical progressions with the infinite series sums:

$$S_l \leq d_2 \left(2(m+1/2) + (d_3+1/2)\frac{8}{7} \right) n_l + d_2.$$

Hence (2.33) holds with the constants

$$c_8 = 2d_2 \qquad \text{and} \qquad c_9 = d_2 + 4(2d_3+1)/7 + d_4. \quad \square$$

2.2 Asymptotic stability of the algorithm of triangulation refinement for a 3D domain

2.2.1 The algorithm of dividing. The considered algorithm is an extension of the well-known algorithm of dividing a 2D bounded domain with a curvilinear boundary (see, for example, [31]). Being often used for 3D problems, it gives triangulations of good quality when an initial triangulation is appropriate. But up to now, as we know, it has not been proved that obtained triangulations remain of good quality for an arbitrary number of recurrent application of this algorithm.

To begin with, we consider the procedure *Div* of dividing a tetrahedron with rectilinear edges (see Fig. 2.a). Two edges of a polyhedron are said to be incidental if they have only one common vertex and nonincidental if they have no common vertex. In an initial tetrahedron (of level 0) we mark two nonincidental edges which have the shortest (among three possible cases) distance between their midpoints (edges 2 and 6 in Fig. 2.a). Then the procedure *Div* consists of five steps.

1. We draw four triangular faces through the midpoints of each triple of incidental edges. These faces cut off four tetrahedra similar to the initial one with the coefficient 1/2. They are shown in Fig.2.c on a scale of 2:1.

Note that the edges marked with the same numbers in Figs. 2, 3, and 4 are parallel to each other and their lengths are in the ratio of 1:2 or 1:4 depending on the number of primes. For example, the edge $4'$ in Fig. 2.c is parallel to the edge 4 in Fig. 2.a and twice shorter. Respectively the edge $4''$ in Fig. 3.c is parallel to the edges 4, $4'$ and four times shorter then the edge 4.

2. Each cut off tetrahedron has exactly one edge which is one-half of the marked edge of the tetrahedron being divided. We mark this edge in the new tetrahedron together with the nonincidental one.

3. After cutting-off four tetrahedra the octahedron with six vertices remains. In Fig. 2.b it is shown on a scale of 2:1. Its vertices may be connected by three diagonals not coincident with the edges. The choice of a diagonal is of principal significance for quality of the obtaining tetrahedra. We draw the diagonal connecting the midpoints of the marked edges ($7'$ in Fig. 2.c).

4. Through this diagonal and each of four vertices lying outside of it we draw triangular faces cutting the octahedron into 4 tetrahedra. Thus we obtain two pairs of equal tetrahedra.

5. In each obtained tetrahedron we mark the common edge, being the diagonal of the octahedron, and the nonincidental one.

These tetrahedra of level 1 are shown in Figs. 2.d and 2.e on a scale of 2:1. In a general case these tetrahedra are not equal to ones initially cut off although they all are of equal volume. As the result we have 3 different kinds of tetrahedra unlike the 2D case when all triangles of the same level are equal.

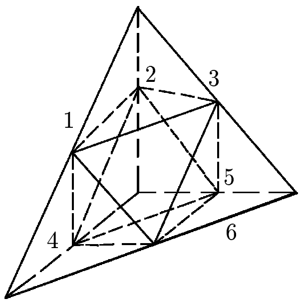
Thus, in each of 8 obtained tetrahedra there are two marked edges. Therefore we can repeat the procedure Div (steps 1–5) for these tetrahedra (Figs. 3–4). As a result of its recurrent applying we obtain a set of 8, 64, 512 etc. tetrahedra of the levels 1, 2, 3. Their volumes are 1/8, 1/64, 1/512 of initial one respectively.

Lemma 21. *On each level of dividing an initial tetrahedron with rectilinear edges, the procedure Div gives tetrahedra only of three different kinds, i.e., for any number of recurrent applying this procedure we obtain only three kinds of similar tetrahedra.*

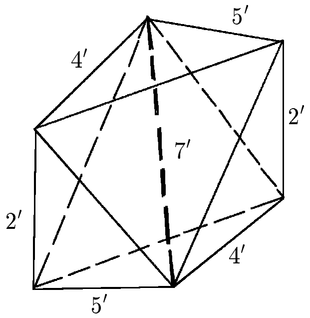
Proof. The application of the procedure Div to tetrahedra of two other kinds (of level 1) is shown in Figs. 3 and 4. Using the well-known theorems on similarity of triangles and on parallelism of straight lines one can easily see that on level 2 new kinds of tetrahedra do not arise. □

Now let us consider a bounded domain Ω in R^3 with a boundary Γ of the class C^2. Further we will intensively use different aspects of smoothness of a boundary. There exist many equivalent definitions of C^2 (see, for example, [27], [7], [14]). We choose the definition similar to the one presented in [27]. It is suitable for theoretical investigations and therefore is often used in literature. In practice, however, specifying a boundary either as level surfaces of some functions, or as a result of some set-theoretic operations, or as a parametric representation etc. is preferred.

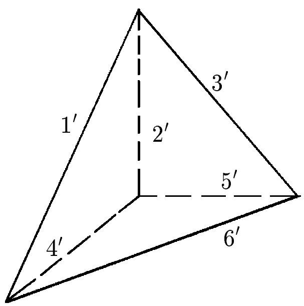
The coordinates $y = (y_1, y_2, y_3)$ obtained by rotation of the given Cartesian ones (x_1, x_2, x_3) and translation of the origin into the point $\overline{x} = (\overline{x}_1, \overline{x}_2, \overline{x}_3)$,



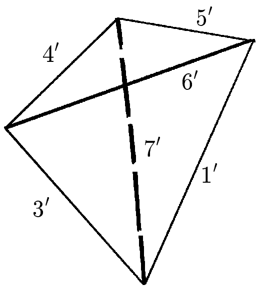
a) the initial tetrahedron
(of the first kind) on a scale of 1:1



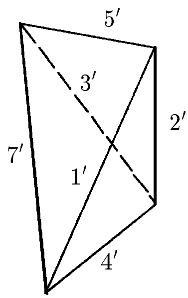
b) the octahedron of the first kind
with the main diagonal 7′ (2:1)



c) four tetrahedra of the first kind
with the similarity ratio 1/2 (2:1)

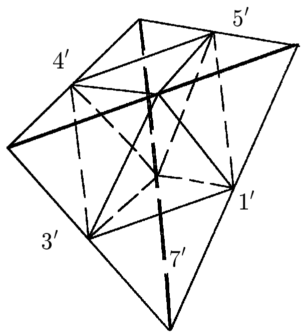


d) two equal tetrahedra
of the second kind (2:1)

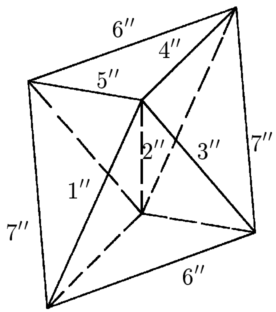


e) two equal tetrahedra
of the third kind (2:1)

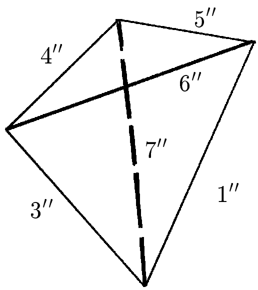
Fig. 2. The subdivision of the initial tetrahedron with rectilinear edges into 8 parts.



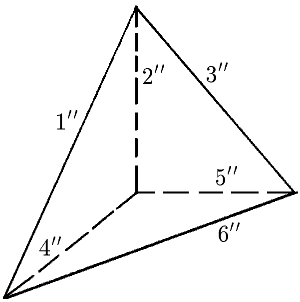
a) the tetrahedron of the second kind on a scale of 2:1



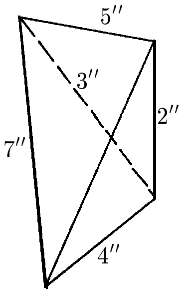
b) the octahedron of the second kind with the main diagonal 2″ (4:1)



c) four tetrahedra of the second kind (4:1)

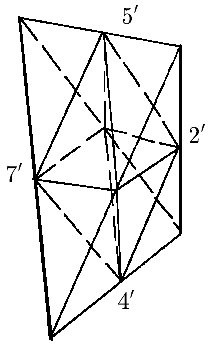


d) two equal tetrahedra of the first kind (4:1)

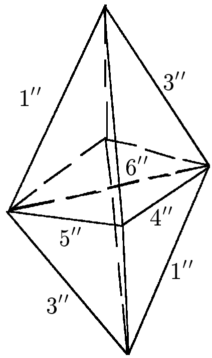


e) two equal tetrahedra of the third kind (4:1)

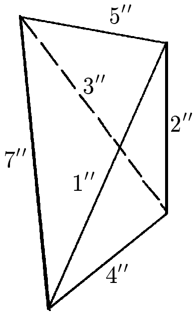
Fig. 3. The subdivision of the tetrahedron of the second kind into 8 parts.



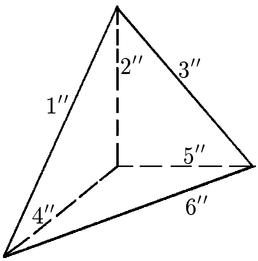
a) the tetrahedron of the third kind on a scale of 2:1



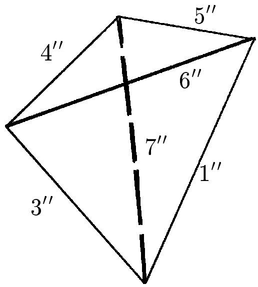
b) the octahedron of the third kind with the main diagonal 6″ (4:1)



c) four tetrahedra of the third kind (4:1)



d) two equal tetrahedra of the first kind (4:1)



e) two equal tetrahedra of the second kind (4:1)

Fig. 4. The subdivision of the tetrahedron of the third kind into 8 parts.

i.e., $y_k = \sum\limits_{l=1}^{3} a_{kl}(x_l - \overline{x}_l)$ with the orthogonal matrix $(a_{kl})_{k,l=1}^{3}$, are called the local coordinates at the point $\overline{x}$.

Denote by $K_r \subset R^2$ the open circle with the radius r and the center at the origin:

$$K_r = \{(y_1, y_2) : y_1^2 + y_2^2 \leq r^2\}$$

and by $C_{r,a} \subset R^3$ the open cylinder $K_r \times (-a, a)$.

The domain Ω is said to be of the class C^2 if there exist three positives r, a, b such that at each point $x_0 \in \Gamma$ the local coordinates $y = (y_1, y_2, y_3)$ can be introduced and the intersection of the boundary Γ with the closed cylinder $\overline{C}_{r,a}$ in y-coordinates is given by the equation

$$y_3 = \omega(y_1, y_2) \tag{2.37}$$

where $\omega \in C^2(\overline{K}_r)$ and

$$\|\omega\|_{C^2(\overline{K}_r)} \leq b, \tag{2.38}$$

moreover,

$$\overline{\Omega} \bigcap \overline{C}_{r,a} = \{y : \, y_1^2 + y_2^2 \leq r, \; \omega(y_1, y_2) \leq y_3 \leq a\}. \tag{2.39}$$

We assume that the initial coarsest triangulation $\mathcal{T}_0$ is given and satisfies the following properties for $i = 0$.

1. A triangulation $\mathcal{T}_i$ is a set of closed nondegenerate tetrahedra. Their union is a polyhedron with only triangular faces and vertices lying on Γ.
2. A triangulation $\mathcal{T}_i$ is consistent, i.e., each two tetrahedra are mutually disjointed or have either one common vertex, or common entire edge, or common entire face.
3. The maximal edge length of a triangulation $\mathcal{T}_i$ denoted by h_i satisfies the inequality
$$h_i \leq \min\{r, a\} \tag{2.40}$$
with constants r, a from (2.37)–(2.39).
4. Each tetrahedron of $\mathcal{T}_i$ has at least one vertex inside Ω.

A set of all vertices of tetrahedra will be called the nodes of the triangulation. For subsequent refinement of the triangulation we generalize the notion of an edge and introduce new nodes by the following rule. If at least one of two nodes connected by some edge lies inside Ω then, as before, an edge of the triangulation is the straight segment connecting these nodes. In this case we take the midpoint of the edge as a new node of the finer triangulation. If both nodes connected by some edge belong to Γ, we introduce at one of them the local coordinates satisfying the conditions (2.37)–(2.39). The distance between this two nodes is less then r and a, therefore there exists a point $\overline{y} = (\overline{y}_1, \overline{y}_2, \overline{y}_3)$ corresponding to the second node in this coordinates such that $\overline{y}_3 = \omega(\overline{y}_1, \overline{y}_2)$. Then as an edge we put the curvilinear segment on Γ, connecting these nodes:

$$y_1 = t\overline{y}_1, \; y_2 = t\overline{y}_2, \; y_3 = \omega(t\overline{y}_1, t\overline{y}_2), \; t \in [0, 1].$$

As a new node we take the point $\hat{y} = (\hat{y}_1, \hat{y}_2, \hat{y}_3) \in \Gamma$ with coordinates

$$\hat{y}_1 = \overline{y}_1/2, \; \hat{y}_2 = \overline{y}_2/2, \; \hat{y}_3 = \omega(\hat{y}_1, \hat{y}_2). \tag{2.41}$$

Thus, if a tetrahedron has at most one vertex on Γ, we take the midpoints of its edges as new nodes. Such tetrahedron is divided by the procedure *Div* into 8 tetrahedra of equal volume but not equal among themselves. In general we obtain tetrahedra of three different kinds: four of the first kind and two pairs of the other kinds. Simultaneously the edges are marked (it is required to apply the procedure *Div* latter).

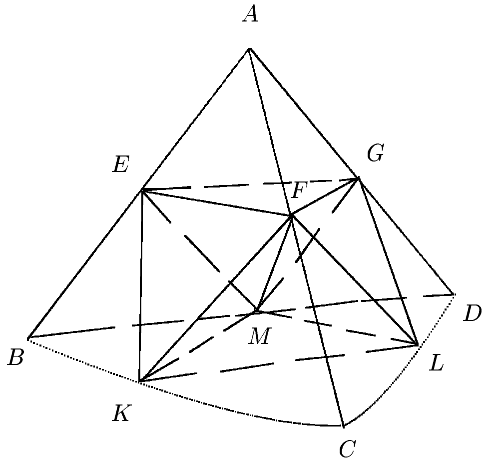


Fig. 5. Dividing a curvilinear "tetrahedron"

Now we consider dividing a tetrahedron which has more then one vertex on Γ. As an example we take the tetrahedron $ABCD$ (Fig. 5) whose three vertices B, C, D belong to Γ and the vertex A lies inside Ω according to the property 4. In a general case the new nodes K, L, M may lie outside of the plane faces of the tetrahedron $ABCD$. Therefore the union of the smaller tetrahedra may not coincide with the initial one. In this connection we realize refinement of a triangulation by analogue of the procedure Div using the generalized terms "edge" and "midpoint of an edge". The generalized term "tetrahedron" we assign to a set of four nodes and connecting them edges (some edges may be curvilinear). Then a union of three mutually incidental edges, sequentially connecting three vertices, is said to be a triangular "face".

To begin with, we mark two nonincidental edges of the initial curvilinear "tetrahedron" (of level 0) which have the shortest distance between their midpoints. Then we realize five steps of the procedure Div using generalized terms "edge", "face" and "polyhedron" ("tetrahedron", "octahedron"). The tetrahedron $AEFG$ is cut off from $ABCD$ by the plane passing through the nodes E, F, G. The "tetrahedra" $BEKM$, $CFKL$, and $DGLM$ are formed by the new edges, drawn through the midpoints of the corresponding edges of $ABCD$, and halves of the old ones. After this we divide the octahedron $EFGMKL$ into 4 tetrahedra. We draw the straight segment connecting the midpoints of the two marked "edges" (the segment FM in Fig. 5). Through this segment and each of four remaining vertices we draw the triangular "faces" dividing the octahedron into 4 tetrahedra. Using four vertices of $ABCD$ together with six new nodes we obtain 8 smaller tetrahedra. Observe that each of them has at least one vertex inside Ω.

If only two vertices of an initial tetrahedron belong to Γ, we apply the same subdivision scheme with some clear simplification.

The triangulation $\mathcal{T}_i$ obtained by dividing all tetrahedra of a consistent triangulation $\mathcal{T}_{i-1}$ is consistent as well.

The considered procedure of simultaneously dividing all tetrahedra can be repeated over and over again. We want this to be done l times.

We denote by $\mathcal{F}_i^1$ the set of tetrahedra of the triangulation $\mathcal{T}_i$ which have at most one vertex belonging to Γ and by $\mathcal{F}_i^2$ the set of remaining ones.

Remark 1. To demonstrate the importance of the choice of the main diagonal in an octahedron we consider the following example. Let us take the tetrahedron with the vertices (0, 0, 0), (1, 0, 0), (0, 1, 0), and (0, 0, 1) and repeat a few times the subdivision procedure similar to Div but choosing the *largest* diagonal of octahedra. Sequential degeneration of a tetrahedron into an arrow, as a result of application of this procedure, is shown in Fig. 6.

Remark 2. Algorithmically the choice of the shortest diagonal in octahedra (as it has been done on level 0) seems more attractive. It gives, in principle, the same quality of tetrahedra but is of local character and does not connect two levels with each other. But in the case of equal diagonals in octahedra with rectilinear edges as well as in the case of approximately equal

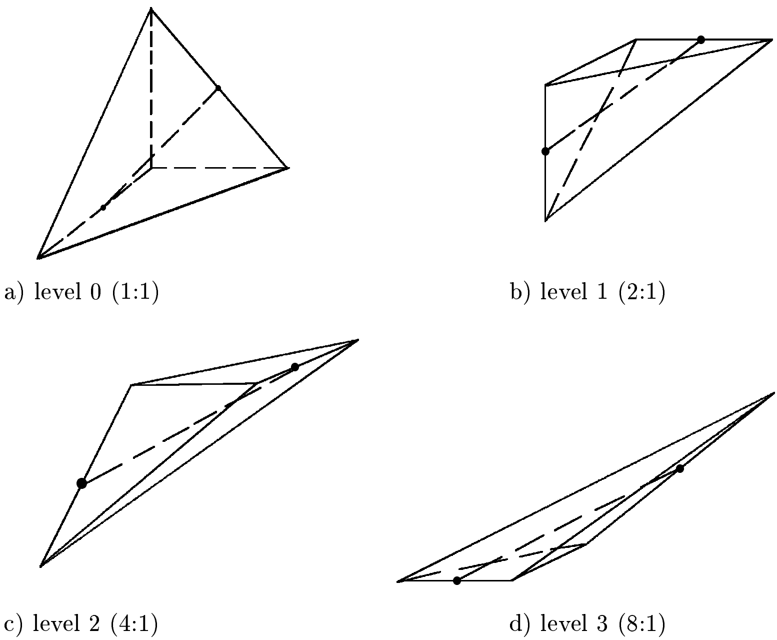


Fig. 6. Sequential degeneration of a tetrahedron (in parentheses the scale is indicated)

diagonals in octahedra with curvilinear edges, a “jump” is possible from one kind of diagonals to another one, resulting in irregular structure of the final triangulation. The rule assumed in the procedure *Div* is rather more complicated but gives a global regular subdivision. For rectilinear initial tetrahedra it gives the same regular triangulations as the global procedure in [31]. For curvilinear tetrahedra the final triangulation is regular within a smooth mapping.

Remark 3. The refinement algorithm for rectilinear tetrahedra, that also gives only three different kinds of tetrahedra on each level of dividing, has been proposed by J. Bey [4]. Without knoledge of this, the authors developed the algorithm presented above. We point out some distinctions between these two approaches. In the algorithm of J. Bey not edges but vertices of tetrahedra are marked. The proof of Lemma 21 is of a radically different kind from that in [4]. Furthemore, in [4] the diagonal in the octahedron on the initial level of dividing is chosen in an arbitrary way. We choose the shortest diagonal to minimize degeneracy of the obtained tetrahedra.

2.2.2 Criteria of quality of a triangulation. If an initial tetrahedron belongs to $\mathcal{F}_0^1$ then on the i-th step of the triangulation refinement it is divided into 2^{3i} tetrahedra of equal volume. Some of them are similar to the initial one with the coefficient $1/2^i$. Remaining tetrahedra form two groups of equal tetrahedra. In addition, tetrahedra of each group are similar to some tetrahedra of the triangulation $\mathcal{T}_1$, obtained by the refinement of the initial one, with the coefficient $1/2^{i-1}$.

If an initial tetrahedron belongs to $\mathcal{F}_0^2$ then, in general, tetrahedra of the triangulation $\mathcal{T}_i$ obtained by its dividing are not similar to the initial one. As a result, there arises a question concerned with the potential possibility of deterioration of quality of a triangulation, for example, of asymptotic degeneration of a tetrahedron into an arrow, a plate, etc.

Some criteria of quality of a triangulation for rectilinear tetrahedra are known from the literature (see, for example, [7] and [31, §2.6]). The most frequently encountered criterion is

$$\varkappa_1(\mathcal{T}_i) = \max_{T\in\mathcal{T}_i} diam(T)/\rho(T) \tag{2.42}$$

where $\rho(T)$ is the largest radius of the balls inscribed into a tetrahedron T and $diam(T)$ is a diameter of T coinciding with the largest length of its edges.

Recall that the inequality [7]

$$\varkappa_1(\mathcal{T}_i) \le c_1 < \infty \tag{2.43}$$

is one of the sufficient conditions providing the usual degree of approximation of functions by finite elements on tetrahedra.

Another widely known criterion of quality is the quantity

$$\varkappa_2(\mathcal{T}_i) = h_i / \min_{T\in\mathcal{T}_i} diam(T), \tag{2.44}$$

$$\text{where } \; h_i = \max_{T\in\mathcal{T}_i} diam(T). \tag{2.45}$$

For example, the inequality [7]

$$\varkappa_2(\mathcal{T}_i) \le c_2 < \infty, \tag{2.46}$$

which is called "the opposite assumption", together with (2.43) is a part of the sufficient conditions providing convergence of an approximate solution in some norms weaker then the energy one.

We consider the criterion

$$\varkappa_3(\mathcal{T}_i) = h_i^3 / \min_{T\in\mathcal{T}_i} meas(T) \tag{2.47}$$

where $meas(T)$ is the volume of a tetrahedron T. For rectilinear tetrahedra this criterion is stronger then either of (2.42) and (2.44) in the following sense.

Lemma 22. *Let the inequality*

$$\varkappa_3(\mathcal{T}_i) \le c_3 < \infty \tag{2.48}$$

be fulfilled. Then (2.43) *and* (2.46) *holds with the constants*

$$c_1 \le 2c_3/3 \qquad \text{and} \qquad c_2 \le (c_3/6)^{1/3} \tag{2.49}$$

respectively.

Proof. First we prove the inequality (2.46). Assume that a tetrahedron $T' \in \mathcal{T}_i$ has the smallest diameter. Then from (2.48) and the obvious inequality

$$meas(T) \le \frac{1}{6} diam^3(T) \tag{2.50}$$

it follows that

$$\begin{aligned} h_i^3 &\le c_3 \min_{T\in\mathcal{T}_i} meas(T) \\ &\le c_3 meas(T') \le \frac{c_3}{6} diam^3(T') \le \frac{c_3}{6} \min_{T\in\mathcal{T}_i} diam^3(T). \end{aligned}$$

Dividing the obtained inequality by min $diam^3(T)$ and extracting the root result in (2.46) with the constant c_2 from (2.49).

To prove the inequality (2.43) we apply another obvious relation

$$meas(T) = \frac{1}{3}\rho(T)(S_1 + S_2 + S_3 + S_4) \le \frac{2}{3}\rho(T) diam^2(T), \tag{2.51}$$

where S_i are the areas of four faces of the tetrahedron. From this relation and (2.48) we obtain the sequence of the inequalities

$$\begin{aligned} diam^3(T') &\le h_i^3 \le c_3 \min_{T\in\mathcal{T}_i} meas(T) \\ &\le c_3 meas(T') \le \frac{2c_3}{3}\rho(T') diam^2(T') \end{aligned}$$

for an arbitrary tetrahedron $T' \in \mathcal{T}_i$. Let us divide the both parts of the obtained inequality by $\rho(T')\, diam^2(T')$ and take maximum among $T' \in \mathcal{T}_i$. We end up with (2.43) with the constant c_1 from (2.49). □

For a curvilinear tetrahedron T the quantities $\rho(T)$, $diam(T)$ and $meas(T)$ are considered as the same ones corresponding to the rectilinear tetrahedron T' with the same vertices.

2.2.3 Estimation of quality of the triangulation. Thus we turn our attention to the criterion (2.47). We intend to show that its value increases only slightly, no matter how many times the procedure of dividing have been applied.

Theorem 23. *Let $\Omega \subset R^3$ be a bounded domain with a boundary Γ of the class C^2 and the initial triangulation $\mathcal{T}_0$ satisfies the properties 1–4. Assume that the inequality*

$$\varkappa_3(\mathcal{T}_0) \leq c_4' < \infty \tag{2.52}$$

holds and h_0 is small enough. Then after applying the procedure Div to all tetrahedra of $\mathcal{T}_0$ we obtain the triangulation $\mathcal{T}_1$ satisfying the properties 1–4, and for $\mathcal{T}_1$ the following estimate holds:

$$\varkappa_3(\mathcal{T}_1) \leq c_4 < \infty \quad where \ c_4 \leq 16c_4'. \tag{2.53}$$

Further, if the inequality (2.53) holds and h_1 is small enough then each triangulation $\mathcal{T}_i$ of level $i = 2, 3, \dots, l$ obtained by the procedure Div satisfies the properties 1–4. Moreover,

$$\varkappa_3(\mathcal{T}_i) \leq \sigma c_4 \quad where \ \sqrt{2} < \sigma \leq 2. \tag{2.54}$$

Proof. First we prove the more fine inequality (2.54) assuming that the properties 1–4 for the triangulation $\mathcal{T}_1$ have been proved, $\mathcal{T}_1$ is consistent, and the first inequality in (2.53) is valid. Assume that

$$h_1 \leq \min\left\{ \frac{2(2^{1/6}-1)}{2^{1/3}b^2}, \frac{\sigma - 2^{1/2}}{3 \cdot 2^{2/3}\sigma c_4 b^2} \right\}. \tag{2.55}$$

Let us consider a tetrahedron T_1 of $\mathcal{F}_1^1$. For each tetrahedron $T_2 \in \mathcal{T}_2$ of eight ones obtained by dividing T_1 the equality

$$meas(T_2) = meas(T_1)/8 \tag{2.56}$$

is valid. As we already know, on the second level of dividing a rectilinear tetrahedron new groups of similarity do not arise. Therefore for each $T_2 \in \mathcal{T}_2$ there exists a tetrahedron of $\mathcal{T}_1$ similar to T_2 with the coefficient 2. It follows that there exists a tetrahedron $T_1' \in \mathcal{T}_1$ satisfying

$$\max_{T_2 \in \mathcal{T}_1} diam(T_2) = diam(T_1')/2. \tag{2.57}$$

Now we derive similar relations for the tetrahedra obtained by dividing an initial one of $\mathcal{F}_1^2$. The transformation of the face ABC of a tetrahedron $T_{i-1} \in \mathcal{F}_{i-1}^2$ with the vertex A lying inside Ω into four faces of tetrahedra of $\mathcal{T}_i$ is shown in Fig. 7. Note that in general the new node K lies outside of the plane passing through 3 nodes A, B, C. Let us introduce the local coordinates at the point B. Taking into account the way of specifying the node K we can represent the curve BKC as follows

$$y_1 = t\overline{y}_1,\ y_2 = t\overline{y}_2,\ y_3 = \omega(t\overline{y}_1, t\overline{y}_2),\quad t \in [0,1], \tag{2.58}$$

where $(\overline{y}_1, \overline{y}_2, \overline{y}_3)$ are the local coordinates of the point C and the values $t = 0,\ 1/2,\ 1$ correspond to the nodes B, K, C respectively. Analyzing the remainder term of the Lagrange interpolational polynomial of degree 1 we get the estimate for deviation of the point K from the midpoint N of the rectilinear segment BC:

$$|NK| \le \frac{b^2}{8} \cdot |BC|^2 \tag{2.59}$$

with the constant b from the condition (2.38).

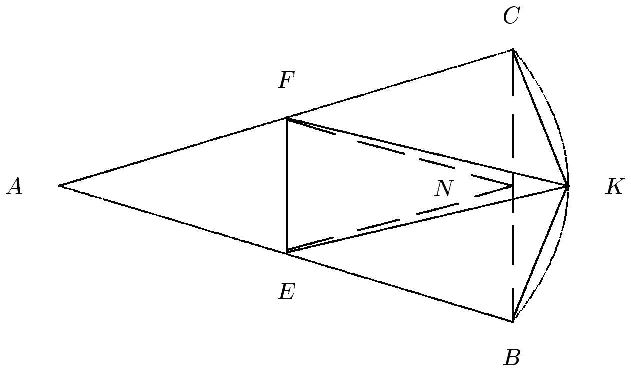


Fig. 7. Transformation of the face ABC of a tetrahedron of $\mathcal{F}^2_{i-1}$ into four ones of tetrahedra of level i.

Then the differences between the lengths of the segments FK, CK, EK, BK and the halves of the lengths of the segments AB, BC, AC, and BC respectively are less then the right-hand side of (2.59). We can see from here that the length a_i of an arbitrary edge of a tetrahedron of the triangulation $\mathcal{T}_i$ is estimated as

$$a_i \le \frac{1}{2} a_{i-1} + c_5 b^2_{i-1},\quad c_5 = b^2/8,$$

where a_{i-1} is the length of the corresponding edge of the tetrahedron T_{i-1} and b_{i-1} is the length of the edge of T_{i-1} that connects the vertices belonging to the boundary. Taking the maximum of the both sides of the inequality and using (2.57), we conclude that

$$h_i \le \frac{1}{2} h_{i-1} + c_5 h_{i-1}^2, \quad i = 2, 3, \ldots, l. \tag{2.60}$$

Because of (2.55) for the triangulation $\mathcal{T}_1$ we have

$$h_1 \le 2(\sqrt[6]{2} - 1)/\sqrt[3]{2} b^2. \tag{2.61}$$

Let us prove the estimate

$$h_i \le \sqrt[6]{2} h_1 / 2^{i-1}. \tag{2.62}$$

It follows from (2.60) and (2.61) that (2.62) is valid for $i = 2$. Assume that it holds for $i - 1$. The estimate (2.60) implies

$$h_i \le \frac{1}{2^{i-1}} h_1 + c_5 \sum_{j=2}^{i} \frac{1}{2^{i-j}} h_{j-1}^2.$$

We apply (2.62) to estimate h_{j-1}^2 for $j \le i$ and replace the sum of the geometrical progression with the infinite series sum:

$$h_i \le \frac{h_1}{2^{i-1}} \left(1 + \sqrt[3]{2} c_5 h_1 \sum_{j=2}^{i} \frac{1}{2^{j-3}} \right) \le \frac{h_1}{2^{i-1}} \left(1 + \frac{\sqrt[3]{2}}{2} b^2 h_1 \right). \tag{2.63}$$

Hence under the condition (2.61) the estimate (2.62) holds for all i.

Now we consider the "tetrahedron" $AB'C'D'$ of the triangulation $\mathcal{T}_i$ (Fig. 8) with the vertex A lying inside Ω. If the initial tetrahedron of the triangulation $\mathcal{T}_{i-1}$ belonged to $\mathcal{F}_{i-1}^1$ we would obtain the usual tetrahedron $ABCD$ instead of $AB'C'D'$. To compare their volumes we consider the determinant

$$\mathcal{D}_i = \begin{vmatrix} B_1 - A_1 & B_2 - A_2 & B_3 - A_3 \\ C_1 - A_1 & C_2 - A_2 & C_3 - A_3 \\ D_1 - A_1 & D_2 - A_2 & D_3 - A_3 \end{vmatrix},$$

where $A_j,\ B_j,\ C_j,\ D_j,\ \ j = 1, 2, 3$ are the j-the coordinates of the points $A,\ B,\ C,\ D$ respectively. By [26] we have

$$|\mathcal{D}_i| = 6\, meas\ (ABCD).$$

In a similar way

$$|\mathcal{D}_i'| = 6\, meas\ (AB'C'D')$$

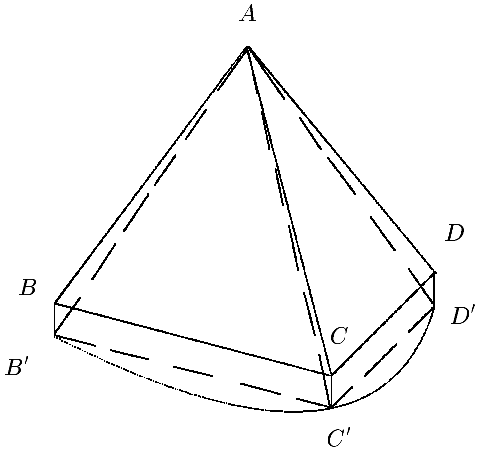


Fig. 8. Rectilinear tetrahedron $ABCD$ and curvilinear "tetrahedron" $AB'C'D'$

where

$$\mathcal{D}_i' = \begin{vmatrix} B_1' - A_1 & B_2' - A_2 & B_3' - A_3 \\ C_1' - A_1 & C_2' - A_2 & C_3' - A_3 \\ D_1' - A_1 & D_2' - A_2 & D_3' - A_3 \end{vmatrix}.$$

A determinant of the third order can be represented as a polynomial function $\mathcal{D} : R^9 \to R$ whose arguments are the elements of the matrix of the determinant. As far as this function is differentiable, by the Lagrange formula we have

$$\mathcal{D}(z + \Delta z) - \mathcal{D}(z) = \sum_{k,j=1}^{3} \frac{\partial \mathcal{D}}{\partial z_{kj}}(\theta)\Delta z_{kj} \tag{2.64}$$

where

$$\theta = (\theta_{kj})_{k,j=1}^3, \quad \theta_{kj} \in [z_{kj},\, z_{kj} + \Delta z_{kj}] \quad k,j = 1,\ldots,3,$$
$$(z_{1j},\, z_{2j},\, z_{3j})_{j=1}^3 = (A_j - B_j,\, A_j - C_j,\, A_j - D_j)_{j=1}^3$$
$$(z_{1j} + \Delta z_{1j},\, z_{2j} + \Delta z_{2j},\, z_{3j} + \Delta z_{3,j})_{j=1}^3 = (A_j - B_j',\, A_j - C_j',\, A_j - D_j')_{j=1}^3.$$

Taking into account the reasoning adduced in the beginning of the proof and applying the inequality $h_{i-1} \le 2h_i$, we get

$$|\Delta z_{kj}| \le c_5 h_{i-1}^2 \le 4c_5 h_i^2, \quad k,j = 1,\ldots,3. \tag{2.65}$$

The quantity $\partial\mathcal{D}/\partial z_{kj}$ is a determinant of the second order. Each element of its matrix does not exceed h_i in modulus. Therefore

$$\left|\frac{\partial \mathcal{D}}{\partial z_{kj}}(\theta)\right| \le 2h_i^2,\ k,j = 1,\ldots,3. \tag{2.66}$$

From (2.64), (2.65), and (2.66) it follows that

$$|\mathcal{D}_i| - |\mathcal{D}'_i| \le |\mathcal{D}_i - \mathcal{D}'_i| \le 9b^2 h_i^4. \tag{2.67}$$

Let us consider the tetrahedron of the triangulation $\mathcal{T}_{i-1}$ that was divided to obtain the "tetrahedron" $AB'C'D'$. Denote by $\mathcal{D}'_{i-1}$ the determinant used for the computation of the volume of this tetrahedron. According to the constructing $ABCD$ we have

$$|\mathcal{D}_i| = \frac{1}{8}|\mathcal{D}'_{i-1}|. \tag{2.68}$$

For $|\mathcal{D}'_{i-1}|$ we can obtain the estimate similar to (2.67). Together with (2.68) this allows to estimate $|\mathcal{D}'_i|$ in term of $|\mathcal{D}_{i-1}|$. Continuing in a similar way, we obtain from (2.67)

$$|\mathcal{D}'_i| \ge \frac{1}{8^{i-1}}|\mathcal{D}_1| - 9b^2 \sum_{j=1}^{i} \frac{1}{8^{j-1}} h_{i-j+1}^4 \tag{2.69}$$

where $\mathcal{D}_1$ is the determinant characterizing the volume of the initial tetrahedron of the triangulation $\mathcal{T}_1$. Taking into account (2.62) and estimating the sum of the geometrical progression, we get from (2.69) the sequence of the inequalities

$$|\mathcal{D}'_i| \ge \frac{1}{8^{i-1}}|\mathcal{D}_1| - 9\sqrt[3]{4}b^2 h_1^4 \sum_{j=1}^{i} \frac{1}{2^{4i-j-3}} \ge \frac{1}{8^{i-1}}(|\mathcal{D}_1| - 18\sqrt[3]{4}b^2 h_1^4),$$

which is valid for tetrahedra of $\mathcal{F}_1^1$ due to (2.56). With (2.53) it leads to

$$|\mathcal{D}'_i| \ge \frac{1}{8^{i-1}}(6h_1^3/c_4 - 18\sqrt[3]{4}b^2 h_1^4) \ge \frac{h_i^3}{\sqrt{2}}(6/c_4 - 18\sqrt[3]{4}b^2 h_1). \tag{2.70}$$

From (2.55) we have

$$h_1 \le \frac{\sigma - 2^{1/2}}{3 \cdot 2^{2/3}\sigma c_4 b^2}.$$

Applying this inequality to (2.70) results in (2.54). The inequality (2.54) guarantees, in particular, nondegeneracy of all tetrahedra of $\mathcal{T}_i$.

From the latter inequality in (2.67) it follows that the determinants $\mathcal{D}'_i$ and $\mathcal{D}_i$ are of the same signs. It means that $ABCD$ and $AB'C'D'$ are identically oriented. Similarly one can prove that the determinants of the second order, characterizing the area and the orientation of its faces (not generalized but plane), are of the same signs and only slightly differ from each other. Basing on obvious reasoning we conclude that the triangulation $\mathcal{T}_i$ satisfies the properties 1–4.

Now we assume that the inequality (2.52) holds and prove (2.53). The initial triangulation is assumed to satisfy the condition

$$h_0 \le \min\{4(2-\sqrt{2})/b^2,\ 1/(24b^2c_4')\}. \tag{2.71}$$

Consider the tetrahedron $T_1 \in \mathcal{T}_1$ of the least volume. It is obtained by dividing the initial one $T_0 \in \mathcal{T}_0$. According to (2.67)

$$\min_{T\in\mathcal{T}_1} meas(T) = meas(T_1) \ge meas(T_0)/8 - 3b^2h_1^4/2. \tag{2.72}$$

From (2.52) it follows that

$$\min_{T\in\mathcal{T}_0} meas(T) \ge h_0^3/c_4'. \tag{2.73}$$

For tetrahedra of the triangulation $\mathcal{T}_1$ the inequality

$$h_1 \le \frac{\sqrt{2}h_0}{2} + c_5h_0^2 \tag{2.74}$$

holds. It can be proved similarly to (2.60) but differs from the latter one by the additional factor $\sqrt{2}$ of the first term in the right-hand side. For rectilinear tetrahedra this factor arises since the maximal length of the edges of the triangulation may increase when the diagonals of octahedra are taken as the edges (the edges $7'$ and $7''$ in Figs. 2–4). To estimate the lengths of this edges one can apply simple consequences of parallelism of corresponding straight lines and planes and also the equality of the sum of squared diagonals of a parallelogram to the sum of its squared sides. Passing on to the curvilinear "tetrahedra", in the same way as in (2.60) we obtain the second term in the right-hand side of (2.74).

Because of (2.71) $h_0 \le 4(2-\sqrt{2})/b^2$. Then (2.74) leads to

$$h_1 \le h_0. \tag{2.75}$$

Apply this inequality together with (2.73) to (2.72):

$$\min_{T\in\mathcal{T}_1} meas(T) \ge \frac{h_0^3}{8c_4'} - \frac{3b^2h_1^4}{2} \ge \frac{h_1^3}{8c_4'}(1-12c_4'b^2h_1). \tag{2.76}$$

From (2.71) and (2.75) we have $h_1 \le 1/(24b^2c_4')$. Using this inequality in the right-hand side of (2.76) we obtain

$$\min_{T\in\mathcal{T}_1} meas(T) \ge h_1^3/(16c_4').$$

From here, taking into consideration the definition of $\varkappa_3(\mathcal{T}_1)$, we conclude that (2.53) holds with the constant $c_4 \le 16c_4'$. It provides, in particular, nondegeneracy of all tetrahedra of $\mathcal{T}_i$.

Further, using the coincidence of signs of the determinants $\mathcal{D}_1'$ and $\mathcal{D}_1$ as well as of the mutually corresponding determinants of the second order which characterize the area and the orientation of the faces of $ABCD$ and $AB'C'D'$, we make a conclusion that the considered tetrahedra as well as their faces are identically oriented. Hence the properties 1–4 are valid for the triangulation $\mathcal{T}_1$. □

Note should be taken of a distinction between the 3D and 2D cases. In the 2D case dividing a triangle into four equal triangles similar to the initial one with the coefficient 1/2 does not change the values of criteria analogous to $\varkappa_1$–$\varkappa_3$. In the 3D case dividing a rectilinear initial tetrahedron into eight ones gives three different groups of similarity. Moreover, the values of the criteria $\varkappa_1$–$\varkappa_3$ may be changed. For example, the value of $\varkappa_3$ for the *reference* initial tetrahedron equals $6\sqrt{2}$, but for some of eight tetrahedra obtained by its dividing it is $2\sqrt{2}$ times greater and equals 24. If the situation described above was repeated the criterion $\varkappa_3$ would be worse exponentially as is shown in Fig. 6 for the unsuccessful strategy of the choice of the main diagonal in octahedra. But in the procedure *Div* such deterioration may occur only once. On subsequent dividing three groups of similar tetrahedra remain and the value of the criterion does not change.

When a curvilinear tetrahedron is divided a similar situation takes place. The more detailed analysis of the constants in the proof of Theorem 23 shows that the less h_1 is, the less the value of $\varkappa_3$ may increase, i.e., the constant σ in (2.54) tends to 1 as h_1 decreases. If we compare $\mathcal{T}_i$ not with $\mathcal{T}_1$ but with $\mathcal{T}_0$ then a jump $\varkappa_3(\mathcal{T}_i)/\varkappa_3(\mathcal{T}_0)$ depends on the geometric form of initial tetrahedra of $\mathcal{T}_0$ and may differ from 1 no matter how small h_0 is.

2.3 The cascadic algorithm for a domain with a smooth curvilinear boundary

2.3.1 Formulation of the differential problem. Consider the Dirichlet problem on a *convex* bounded domain $\Omega \subset R^3$ with a boundary Γ of the class C^2:

$$-\sum_{i,j=1}^{3} \partial_i(a_{ij}\partial_j u) + au = f \ \text{ in } \ \Omega, \tag{2.77}$$

$$u = 0 \ \text{ on } \ \Gamma. \tag{2.78}$$

The coefficients and the right-hand side of the equation satisfy the conditions

$$\begin{gathered}
\partial_i a_{ij} \in L_3(\Omega),\ \ i,j = 1,2,3; \\
a_{ij} = a_{ji} \ \text{ on } \ \overline{\Omega}, i,j = 1,2,3; \\
\mu \sum_{i=1}^{3} \xi_i^2 \le \sum_{i,j=1}^{3} a_{ij}\xi_i\xi_j \le \nu \sum_{i=1}^{3} \xi_i^2 \ \text{ on } \ \overline{\Omega} \forall\, \xi_i \in R, \quad \nu \ge \mu > 0; \\
a, f \in L_2(\Omega), \quad a \ge 0 \ \text{ on } \ \overline{\Omega}.
\end{gathered} \tag{2.79}$$

The boundary Γ is assumed to be the union of intersecting pieces of surfaces $\Gamma_i,\ i = 1, \ldots, k$ represented in the parametric form:

$$x_j = g_{i,j}(u,v), \quad j = 1,2,3, \quad u,v \in R. \tag{2.80}$$

Under this conditions the problem (2.77)–(2.78) admits a unique solution in the class $H_2^2(\Omega)$; moreover the estimate

$$\|u\|_{2,\Omega} \le c_1 \|f\|_{0,\Omega} \tag{2.81}$$

holds [27].

Formulate for (2.77)–(2.78) the generalized problem:

$$\begin{aligned}
&\textit{find} \ \ u \in H_0^1(\Omega) \ \ \textit{such that} \\
&\mathcal{L}(u,v) = (f,v)_\Omega \quad \forall\, v \in H_0^1(\Omega)
\end{aligned} \tag{2.82}$$

where the bilinear form is given by

$$\mathcal{L}(u,v) = \int_\Omega \left(\sum_{i,j=1}^{3} a_{ij}\partial_j u \partial_i v + auv \right) dx. \tag{2.83}$$

Due to (2.79) the problem (2.82) admits a unique solution as well.

2.3.2 Formulation of the discrete problem. Let us construct a spatial triangulation of Ω. First we construct a polyhedron inscribed into Ω, which has only triangular faces and vertices lying on Γ. Then we divide it into a small number of tetrahedra such that the obtained triangulation satisfies the properties 1–4 from the preceding section. Denote by h_0 the maximal length of the edges of the tetrahedra. Thus we obtained the initial triangulation $\mathcal{T}_0$.

For triangulation refinement we apply the procedure *Div* proposed in the preceding section. We introduce new nodes in the following way. If at least one of two nodes connected by some edge lies inside Ω then we take the midpoint of the edge as a new node of the finer triangulation. If an edge connects two nodes belonging to Γ then two cases are possible. If each of the two nodes belong to the same piece of a surface Γ_i then we take as a new node the midpoint of the segment in the parametric variables (2.80) that connects these nodes. In the opposite case we take as a new node the point on Γ nearest to the midpoint of the edge.

The procedure of simultaneously dividing all the tetrahedra may be repeated over and over again. We want this to be done l times.

The triangulation $\mathcal{T}_i$ obtained by dividing all the tetrahedra of $\mathcal{T}_{i-1}$ is certainly consistent as well. Denote by h_i the maximal length of the edges of the triangulation $\mathcal{T}_i$.

Denote by $\mathcal{F}_i^1$ the set of tetrahedra of $\mathcal{T}_i$ which have at most one vertex belonging to Γ and by $\mathcal{F}_i^2$ the set of remaining ones.

Denote by $\overline{\Omega}_i$ the set of all nodes of $\mathcal{T}_i$ and introduce $\Omega_i = \overline{\Omega}_i \bigcap \Omega$ as well as the number n_i of points of Ω_i. Tetrahedra of the finest triangulation will be called elementary.

For each node $y \in \Omega_l$ we introduce the usual basis function $\varphi_y^l \in H_0^1(\Omega)$ which equals 1 at the node y, 0 at the other nodes of $\overline{\Omega}_l$, is linear on each elementary tetrahedron, and vanishes on the rest of $\overline{\Omega}$. Basis functions on coarser grids will be formed as linear combinations of basis functions on finer grids. Suppose that the basis functions φ_y^{i+1} have been constructed for all $y \in \Omega_{i+1}$. Take an arbitrary node $y \in \Omega_i$. Several edges of tetrahedra of the triangulation $\mathcal{T}_i$ issue out of it. Their midpoints are the nodes $y_1, \ldots, y_m$ of the triangulation $\mathcal{T}_{i+1}$. Put

$$\varphi_y^i(x) = \varphi_y^{i+1}(x) + \frac{1}{2}\sum_{j=1}^{m} \varphi_{y_j}^{i+1}(x). \qquad (2.84)$$

Denote by $\mathcal{H}^i$ the linear span of the functions $\varphi_y^i,\ y \in \Omega_i$.

Observe that on the tetrahedra of $\mathcal{F}_i^1$ the basis functions φ_y^i are linear. On the part of $\overline{\Omega}$ lying between the boundary and the union of all elementary tetrahedra the basis functions equal zero. On the remaining part of $\overline{\Omega}$ φ_y^i are defined not on the tetrahedra of $\mathcal{F}_i^2$ but on the polyhedra, which are formed by joining the elementary tetrahedra obtained by dividing the tetrahedra of $\mathcal{F}_i^2$ and does not coincide with the initial tetrahedra. These functions are composed of pieces of functions being linear on each elementary tetrahedron only.

Considering (2.82) on the subspace $\mathcal{H}^i \subset H_0^1(\Omega)$ we obtain the discrete problem:

$$\begin{aligned} &\textit{find } \tilde{v}_i \in \mathcal{H}^i \textit{ such that} \\ &\mathcal{L}(\tilde{v}_i, v) = (f, v)_\Omega \quad \forall\, v \in \mathcal{H}^i. \end{aligned} \tag{2.85}$$

Let M_i be the n_i-dimensional space which consists of vectors w with components $w(x),\ x \in \Omega_i$. Then the problem (2.85) is equivalent to the system of linear algebraic equations

$$L_i v_i = f_i. \tag{2.86}$$

Here $v_i \in M_i$ is the vector of unknowns with components $v_i(y),\ y \in \Omega_i$; the vector $f_i \in M_i$ is defined by $f_i(x) = (f, \varphi_x^i)_\Omega,\ x \in \Omega_i$; L_i is the $n_i \times n_i$ matrix with elements

$$L_i(x, y) = \mathcal{L}(\varphi_x^i, \varphi_y^i). \tag{2.87}$$

Because of (2.83) L_i is self-adjoint. Under the conditions (2.79) it is positive-definite and, hence, nonsingular.

To a vector $v \in M_i$ we put into correspondence its prolongation in $\mathcal{H}^i$:

$$\tilde{v}(x) = \sum_{y \in \Omega_i} v(y)\varphi_y^i(x), \quad x \in \overline{\Omega}. \tag{2.88}$$

It is obvious that

$$v(y) = \tilde{v}(y), \quad y \in \Omega_i. \tag{2.89}$$

Thus, we defined the usual isomorphism between vectors $v \in M_i$ and functions $\tilde{v} \in \mathcal{H}^i$.

Now we introduce the energy norm for functions

$$|\!|\!| v |\!|\!|_\Omega = \mathcal{L}(v, v)^{1/2}, \quad v \in H_0^1(\Omega),$$

as well as the inner product and the norms for vectors

$$\begin{aligned} (v, w)_i &= \sum_{x \in \Omega_i} v(x)w(x), \\ \|v\|_i &= \Big(\sum_{x \in \Omega_i} v^2(x)\Big)^{1/2}, \\ |\!|\!| v |\!|\!|_i &= (L_i v, v)^{1/2}, \quad v, w \in M_i. \end{aligned}$$

Due to (2.88) and (2.89) we have for an isomorphic pair $v \in M_i,\ \tilde{v} \in \mathcal{H}^i$

$$|\!|\!| v |\!|\!|_i = |\!|\!| \tilde{v} |\!|\!|_\Omega. \tag{2.90}$$

We introduce the interpolation operator $I_i : M_i \to M_{i+1}$ in the usual way.

2.3.3 An auxiliary result.

Lemma 24. *Let Ω be a convex bounded domain with a boundary Γ of the class C^2, which is the union of intersecting pieces of surfaces Γ_i, $i = 1, \dots, k$, and ω^δ be an adjoining the boundary interior layer of thickness δ with sufficiently small δ. Then for any arbitrary function $v \in H_2^1(\Omega)$ the estimate*

$$\int\limits_{\omega^\delta} v^2 dx \le c_3 \delta \|v\|^2_{1,\Omega} \tag{2.91}$$

holds.

Proof. Let us consider a surface Γ_i given by the parametric representation $x_j = g_j(\xi_1, \xi_2)$, $j = 1, 2, 3$ where $0 \le \xi_1 \le T_1$, $0 \le \xi_2 \le T_2$. At the boundary points of Γ_i we erect the normal vectors to Γ directed into Ω. They cut off a part of ω^δ adjoining Γ_i. We denote it by ω_i^δ. Let us introduce in ω_i^δ the coordinates (ξ_1, ξ_2, ξ_3) where ξ_3 is a distance between a point and the surface Γ_i. Denote by $N(\xi_1, \xi_2) = (n_1, n_2, n_3)$ the unit normal vector at the point $(\xi_1, \xi_2) \in \Gamma_i$ directed into Ω, i.e.,

$$N = -\left(\frac{\partial G}{\partial \xi_1} \times \frac{\partial G}{\partial \xi_2}\right) \Bigg/ \left|\frac{\partial G}{\partial \xi_1} \times \frac{\partial G}{\partial \xi_2}\right|$$

where $G(\xi_1, \xi_2)$ is the vector-function with components g_j, $j = 1, 2, 3$ and the sign '$\times$' denotes the vector product. Then at a point of ω_i^δ the coordinates $(x_j)_{j=1}^3$ and $(\xi_j)_{j=1}^3$ are connected by the relations

$$x_j = g_j(\xi_1, \xi_2) + n_j(\xi_1, \xi_2)\xi_3, \quad j = 1, 2, 3. \tag{2.92}$$

Let us consider the Jacobian

$$J = \frac{D(x_1, x_2, x_3)}{D(\xi_1, \xi_2, \xi_3)}. \tag{2.93}$$

Denote by X the vector with the components x_j, $j = 1, 2, 3$. Then (2.93) can be represented as the mixed product [10]

$$J = \left(\frac{\partial X}{\partial \xi_1} \times \frac{\partial X}{\partial \xi_2}\right) \cdot \frac{\partial X}{\partial \xi_3}.$$

With (2.92) this yields

$$J = \left(\frac{\partial G}{\partial \xi_1} + \frac{\partial N}{\partial \xi_1}\xi_3\right) \times \left(\frac{\partial G}{\partial \xi_2} + \frac{\partial N}{\partial \xi_2}\xi_3\right) \cdot N. \tag{2.94}$$

We omit tedious calculations and adduce the final result for (2.94):

$$J = |D|^{1/2}(1 - 2H\xi_3 + K\xi_3^2) \tag{2.95}$$

where $|D| = \left|\frac{\partial G}{\partial \xi_1} \times \frac{\partial G}{\partial \xi_2}\right|^2$ is the determinant of the matrix of the first quadratic form for the surface Γ_i, $H = (k_2 + k_2)/2$ is the mean curvature of Γ_i, $K = k_1 k_2$ is the Gaussian curvature of Γ_i, k_1 and k_2 are the values of the principal curvature. Note that because of $\Gamma_i \in C^2$ these quantities are bounded, moreover, $|D| > 0$.

Let us prove that for sufficiently small δ the Jacobian J does not vanish. For definiteness we assume that $|k_1| > |k_2|$. The expression in parentheses in the right-hand side of (2.95) equals zero for $\xi_3 = -1/k_1$ and $\xi_3 = -1/k_2$. Hence for $0 \le \xi_3 \le 1/|k_1|$ we have

$$|1 - 2H\xi_3 + K\xi_3^2| > 0,$$

i.e., for $\delta \le 1/|k_1| - \varepsilon$ with small $\varepsilon > 0$ the estimate

$$0 < c_4 \le |J| \le c_5 \tag{2.96}$$

holds.

Note that under a stronger restriction on δ we can estimate the Jacobian with the constants in an explicit form. Setting $\delta \le 1/4|k_1|$ and taking into account the inequality $\xi_3 \le \delta$, we obtain

$$|2H\xi_3 - K\xi_3^2| \le |(k_1 + k_2)\xi_3| + |k_1 k_2 \xi_3^2| \le 2|k_1|\xi_3 + k_1^2\xi_3^2 \le \frac{9}{16}.$$

This results in the estimate

$$0 < \frac{7}{16}|D|^{1/2} \le J \le \frac{25}{16}|D|^{1/2}. \tag{2.97}$$

Now we are in a position to prove (2.91). Consider the identity

$$v(\xi_1, \xi_2, 0) = v(\xi_1, \xi_2, \xi_3) - \int_0^{\xi_3} \frac{\partial v(\xi_1, \xi_2, \tau)}{\partial \tau} d\tau. \tag{2.98}$$

Square it and integrate over ω_i^δ:

$$\begin{aligned} &\int_0^{T_1}\int_0^{T_2}\int_0^{\delta} v^2(\xi_1, \xi_2, 0) d\xi_3 d\xi_2 d\xi_1 \\ &= \int_0^{T_1}\int_0^{T_2}\int_0^{\delta} \left[v(\xi_1, \xi_2, \xi_3) - \int_0^{\xi_3} \frac{\partial v(\xi_1, \xi_2, \tau)}{\partial \tau} d\tau \right]^2 d\xi_3 d\xi_2 d\xi_1. \end{aligned} \tag{2.99}$$

For the expression in brackets we have

$$\left[v(\xi_1,\xi_2,\xi_3) - \int\limits_0^{\xi_3} \frac{\partial v(\xi_1,\xi_2,\tau)}{\partial\tau} d\tau\right]^2 \le 2v^2(\xi_1,\xi_2,\xi_3) + 2\left(\int\limits_0^{\xi_3} \frac{\partial v(\xi_1,\xi_2,\tau)}{\partial\tau} d\tau\right)^2. \tag{2.100}$$

Applying the Cauchy-Bunyakovskii inequality and taking into account the inequality $0 \le \xi_3 \le \delta$, we obtain the estimate for the second term in the right-hand side of (2.100):

$$\left(\int\limits_0^{\xi_3} \frac{\partial v(\xi_1,\xi_2,\tau)}{\partial\tau} d\tau\right)^2 \le \int\limits_0^{\xi_3} d\tau \int\limits_0^{\xi_3} \left(\frac{\partial v(\xi_1,\xi_2,\tau)}{\partial\tau}\right)^2 d\tau = \xi_3 \int\limits_0^{\xi_3} \left(\frac{\partial v(\xi_1,\xi_2,\tau)}{\partial\tau}\right)^2 d\tau \le \delta \int\limits_0^{\delta} \left(\frac{\partial v(\xi_1,\xi_2,\xi_3)}{\partial\tau}\right)^2 d\xi_3. \tag{2.101}$$

Because of (2.100) and (2.101), from (2.99) we have the inequality

$$\delta \int\limits_0^{T_1}\int\limits_0^{T_2} v^2(\xi_1,\xi_2,0)d\xi_2 d\xi_1 \le 2\int\limits_0^{T_1}\int\limits_0^{T_2}\int\limits_0^{\delta} v^2(\xi_1,\xi_2,\xi_3)d\xi_3 d\xi_2 d\xi_1 + 2\delta^2 \int\limits_0^{T_1}\int\limits_0^{T_2}\int\limits_0^{\delta} \left(\frac{\partial v(\xi_1,\xi_2,\xi_3)}{\partial\xi_3}\right)^2 d\xi_3 d\xi_2 d\xi_1.$$

Dividing this inequality by δ, we get

$$\int\limits_{\Gamma_i} v^2 d\xi_1 d\xi_2 \le \frac{2}{\delta}\int\limits_0^{T_1}\int\limits_0^{T_2}\int\limits_0^{\delta} v^2(\xi_1,\xi_2,\xi_3)d\xi_3 d\xi_2 d\xi_1 + 2\delta \int\limits_0^{T_1}\int\limits_0^{T_2}\int\limits_0^{\delta} \left(\frac{\partial v(\xi_1,\xi_2,\xi_3)}{\partial\xi_3}\right)^2 d\xi_3 d\xi_2 d\xi_1. \tag{2.102}$$

Next we go over from the variables $(\xi_j)_{j=1}^3$ to $(x_j)_{j=1}^3$. Applying (2.96) we estimate the first term in the right-hand side of (2.102):

$$\int_0^{T_1}\int_0^{T_2}\int_0^{\delta} v^2 d\xi_3 d\xi_2 d\xi_1 = \int_{\omega_i^\delta} v^2 |J|^{-1} dx \le \frac{1}{c_4}\int_{\omega_i^\delta} v^2 dx. \tag{2.103}$$

Then we obtain the estimate for the second term. It follows from (2.92) that

$$\frac{\partial x_j}{\partial \xi_3} = n_j, \quad j = 1, 2, 3.$$

Denote by V_x the vector with the components $\partial v/\partial x_j,\ j = 1, 2, 3$. Then

$$\frac{\partial v}{\partial \xi_3} = \frac{\partial v}{\partial x_1}\frac{\partial x_1}{\partial \xi_3} + \frac{\partial v}{\partial x_2}\frac{\partial x_2}{\partial \xi_3} + \frac{\partial v}{\partial x_3}\frac{\partial x_3}{\partial \xi_3} = V_x \cdot N.$$

Since N is the unit vector, we have

$$\left(\frac{\partial v}{\partial \xi_3}\right)^2 = (V_x \cdot N)^2 \le (V_x \cdot V_x)(N \cdot N) = V_x \cdot V_x = \sum_{j=1}^3 \left(\frac{\partial v}{\partial x_j}\right)^2.$$

Together with (2.96) this yields

$$\begin{aligned}&\int_0^{T_1}\int_0^{T_2}\int_0^{\delta} \left(\frac{\partial v}{\partial \xi_3}\right)^2 d\xi_3 d\xi_2 d\xi_1 \\ &\le \int_{\omega_i^\delta} \sum_{j=1}^3 \left(\frac{\partial v}{\partial x_j}\right)^2 |J|^{-1} dx \le \frac{1}{c_4}\int_{\omega_i^\delta} \sum_{j=1}^3 \left(\frac{\partial v}{\partial x_j}\right)^2 dx.\end{aligned} \tag{2.104}$$

From (2.102), (2.103), and (2.104) it follows that

$$\int_{\Gamma_i} v^2 d\xi_1 d\xi_2 \le c_6 \left[\frac{2}{\delta}\int_{\omega_i^\delta} v^2 dx + 2\delta \int_{\omega_i^\delta} \sum_{j=1}^3 \left(\frac{\partial v}{\partial x_j}\right)^2 dx\right].$$

The same inequalities are valid also for the remaining pieces of the surfaces forming Γ. Combining them and extending the domain of integration we get the estimate

$$\int_{\Gamma} v^2 d\xi_1 d\xi_2 \le c_7 \|v\|_{1,\Omega}^2 \tag{2.105}$$

where a constant c_7 depends on properties of the boundary Γ but does not depend on v.

Rewrite (2.98) as

$$v(\xi_1, \xi_2, \xi_3) = v(\xi_1, \xi_2, 0) + \int_0^{\xi_3} \frac{\partial v(\xi_1, \xi_2, \tau)}{\partial \tau} d\tau.$$

Square it and integrate over ω_i^δ. In the same way as (2.105) was proved we obtain the inequality

$$\int_0^{T_1}\int_0^{T_2}\int_0^{\delta} v^2 d\xi_3 d\xi_2 d\xi_1 \le 2\delta^2 \int_0^{T_1}\int_0^{T_2}\int_0^{\delta} \left(\frac{\partial v}{\partial \xi_3}\right)^2 d\xi_3 d\xi_2 d\xi_1 + 2\delta \int_{\Gamma_i} v^2 d\xi_1 d\xi_2. \tag{2.106}$$

Let us go over to the variables $(x_j)_{j=1}^3$. Due to (2.96) we obtain the lower estimate for the left-hand side of (2.106):

$$\int_0^{T_1}\int_0^{T_2}\int_0^{\delta} v^2 d\xi_3 d\xi_2 d\xi_1 = \int_{\omega_i^\delta} v^2 |J|^{-1} dx \ge \frac{1}{c_5} \int_{\omega_i^\delta} v^2 dx.$$

With (2.106) and (2.104) this yields

$$\int_{\omega_i^\delta} v^2 dx \le c_8 \left(2\delta^2 \int_{\omega_i^\delta} \left(\frac{\partial v}{\partial x_j}\right)^2 dx + 2\delta \int_{\Gamma_i} v^2 d\xi_1 d\xi_2 \right).$$

The same inequalities are valid for the remaining parts of ω^δ. Combining them and applying (2.105), we end up with the estimate (2.91). □

2.3.4 Convergence of the Galerkin solution. To prove the convergence of the solution of (2.85) to the exact solution of the problem (2.77)–(2.78) the estimate of approximation of the latter one by functions of $\mathcal{H}^i$ is required. Let us formulate and prove it.

Lemma 25. *Let $u \in H_2^2(\Omega)$ and $u = 0$ on Γ. Then for each $i = 0, 1, \dots, l$ there exists a function $\hat{u}_i \in \mathcal{H}^i$ such that*

$$||u - \hat{u}_i||_{1,\Omega} \le c_9 h_i ||u||_{2,\Omega}. \tag{2.107}$$

Proof. Put

$$\hat{u}_i(x) = \sum_{y\in\Omega_i} u(y)\varphi_y^i(x). \tag{2.108}$$

Denote by ω^δ the part of $\overline{\Omega}$ lying outside of the union of all elementary tetrahedra. Because of (2.59) the maximal thickness δ of the layer ω^δ satisfies the inequality

$$\delta \le c_{10}h_l^2.$$

As far as $\hat{u}_i \equiv 0$, by Lemma 24 we have the estimate

$$\|u - \hat{u}_i\|_{1,\omega^\delta} \le c_{11}h_i\|u\|_{2,\Omega} \tag{2.109}$$

for all $i = 0, 1, \dots, l$.

Since on the finest grid Ω_l the basis functions are of the standard form, the estimate (2.107) for $i = l$ is proved in a number of books (see, for example, [7]). Now we use induction on i. Assume that (2.107) is valid for some i and prove it for $i-1$.

Denote by Ω' the greater part of Ω which is the union of all tetrahedra of $\mathcal{F}_{i-1}^1$. On Ω' the basis functions are standard and hence

$$\|u - \hat{u}_{i-1}\|_{1,\Omega'} \le c_{12}h_{i-1}\|u\|_{2,\Omega}. \tag{2.110}$$

Thus it remains to prove (2.107) on $\Omega'' = (\Omega \setminus \Omega') \setminus \omega^\delta$.

Consider an arbitrary tetrahedron of $\mathcal{F}_{i-1}^2$, for example, $ABCD$ shown in Fig. 5. Denote by $\widetilde{ABCD}$ the union of elementary tetrahedra obtained by dividing $ABCD$. Due to (2.101) and (2.84) we have on $\widetilde{ABCD}$

$$\hat{u}_{i-1}(x) = u(A)\varphi_A^i(x) + \frac{1}{2}u(A)(\varphi_E^i(x) + \varphi_F^i(x) + \varphi_G^i(x)),$$
$$\hat{u}_i(x) = u(A)\varphi_A^i(x) + u(E)\varphi_E^i(x) + u(F)\varphi_F^i(x) + u(G)\varphi_G^i(x).$$

Then

$$\hat{u}_i(x) - \hat{u}_{i-1}(x) = \left(\frac{1}{2}u(A) - u(E)\right)\varphi_E^i(x) + \left(\frac{1}{2}u(A) - u(F)\right)\varphi_F^i(x) + \left(\frac{1}{2}u(A) - u(G)\right)\varphi_G^i(x).$$

Applying the Cauchy algebraic inequality we obtain the estimate

$$\left|\frac{\partial}{\partial x_1}(\hat{u}_i - \hat{u}_{i-1})\right|^2 \le \left[\left(\frac{1}{2}u(A) - u(E)\right)^2 + \left(\frac{1}{2}u(A) - u(F)\right)^2 + \left(\frac{1}{2}u(A) - u(G)\right)^2\right] \cdot \left[\left(\frac{\partial\varphi_E^i}{\partial x_1}\right)^2 + \left(\frac{\partial\varphi_F^i}{\partial x_1}\right)^2 + \left(\frac{\partial\varphi_G^i}{\partial x_1}\right)^2\right]. \tag{2.111}$$

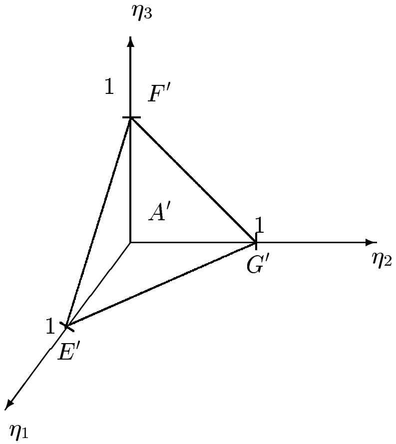


Fig. 9. The *reference* tetrahedron

Some of smaller tetrahedra obtained by dividing $ABCD$ belong to $\mathcal{F}_i^1$, for example, the tetrahedron $AEFG$ in Fig.5. Consider the linear mapping which transforms the *reference* tetrahedron $A'E'F'G'$ with the vertices (0, 0, 0), (1, 0, 0), (0, 1, 0), and (0, 0, 1) (see Fig.9) into $AEFG$. The Jacobian of this mapping can be represented as

$$J = \frac{D(x_1, x_2, x_3)}{D(\eta_1, \eta_2, \eta_3)} = \begin{vmatrix} E_1 - A_1 & G_1 - A_1 & F_1 - A_1 \\ E_2 - A_2 & G_2 - A_2 & F_2 - A_2 \\ E_3 - A_3 & G_3 - A_3 & F_3 - A_3 \end{vmatrix} \tag{2.112}$$

where $A_j, E_j, F_j, G_j,\ j = 1, 2, 3$ are the j-th coordinates of the points $A,\ E,\ F,\ G$. Introduce the basis functions in the variables η

$$\Phi^i_{y'}(\eta) = \varphi^i_y(x(\eta)).$$

Then

$$\int\limits_{AEFG} \left(\frac{\partial \varphi^i_E}{\partial x_1} \right)^2 dx = \int\limits_{A'E'F'G'} \left(\sum_{j=1}^{3} \frac{\partial}{\partial \eta_j} \Phi^i_{E'}(\eta) \frac{\partial \eta_j}{\partial x_1} \right)^2 |J| d\eta. \tag{2.113}$$

By definition of the basis functions we have on $A'E'F'G'$

$$\left| \frac{\partial}{\partial \eta_j} \Phi^i_{E'}(\eta) \right| \leq 1, \quad j = 1, 2, 3.$$

Therefore from (2.113) it follows that

$$\int_{AEFG} \left(\frac{\partial \varphi_E^i}{\partial x_1}\right)^2 dx \le 3 \int_{A'E'F'G'} \sum_{j=1}^{3} \left(\frac{\partial \eta_j}{\partial x_1}\right)^2 |J| d\eta. \tag{2.114}$$

Each element of the matrix in (2.112) does not exceed h_i in modulus. The quantities $\partial \eta_k / \partial x_j$ are the elements of the inverse matrix. They are expressed in terms of the determinant J and the minors of the second order which does not exceed $c_{13} h_i^2$ in modulus. Hence

$$\left| \frac{\partial \eta_k}{\partial x_j} \right| \le \frac{c_{13} h_i^2}{|J|}, \quad k, j = 1, 2, 3. \tag{2.115}$$

Due to [26]

$$|J| = 6 \; meas \; AEFG.$$

Then because of Theorem 23

$$|J| \ge c_{14} h_i^3. \tag{2.116}$$

From (2.114) together with (2.115) and (2.116) we obtain the following estimate:

$$\int_{AEFG} \left(\frac{\partial \varphi_E^i}{\partial x_1}\right)^2 dx \le c_{15} h_i. \tag{2.117}$$

The estimate (2.117) is valid not only on $AEFG$ but also on all tetrahedra of $\mathcal{F}_i^1$ obtained by dividing $ABCD$.

Further we divide the remaining tetrahedra of $\mathcal{F}_i^2$ into smaller ones and obtain on them the estimates for the standard basis functions. This procedure will be repeated up to level l.

Denote by Δ_{i+k} a tetrahedron of $\mathcal{F}_{i+k}^1$ obtained by dividing a tetrahedron of $\mathcal{F}_{i+k-1}^2$ such that the function φ_E^i is not equal to zero on Δ_{i+k}. We denote the vertices of Δ_{i+k} by z_j, $j = 1, \ldots, 4$. From (2.84) it follows that on Δ_{i+k} either

$$\varphi_E^i(x) = \frac{1}{2^k} \sum_{j=1}^{4} \varphi_{z_j}^{i+k}(x) \tag{2.118}$$

if no vertex of z_j, $j = 1, \ldots, 4$ is a node of $\mathcal{T}_{i+k-1}$ or

$$\varphi_E^i(x) = \frac{1}{2^{k-1}} \varphi_{z_1}^{i+k}(x) + \frac{1}{2^k} \sum_{j=2}^{4} \varphi_{z_j}^{i+k}(x) \tag{2.119}$$

if one of the vertices (z_1 in this case) is a node of $\mathcal{T}_{i+k-1}$. On Δ_{i+k} the basis functions $\varphi_{z_j}^{i+k}$ from the right-hand sides of (2.118) and (2.119) are of the standard form therefore on Δ_{i+k} the estimate similar to (2.117) is valid for them. Hence

$$\int\limits_{\Delta_{i+k}} \left(\frac{\partial \varphi_E^i}{\partial x_1}\right)^2 dx \le \frac{c_{16} h_{i+k}}{2^{2(k-1)}}. \tag{2.120}$$

The estimates similar to (2.117) and (2.120) hold for the functions φ_F^i and φ_G^i as well.

Denote by m_{i+k} the number of tetrahedra of $\mathcal{F}_{i+k}^2$, obtained by dividing $ABCD$, and estimate it. As a result of dividing, the face BCD is transformed into the surface composed of smaller triangles. The number of these triangles equals $2^{2(k+1)}$. Each of them is a base of a polyhedron composed of three tetrahedra of $\mathcal{T}_{i+k}$. For example, in Fig. 5 the triangle BKM is a base of the polyhedron composed of the tetrahedra $EBKM$, $EFKM$, and $EFGM$.

It is evident that all tetrahedra of $\mathcal{F}_{i+k}^2$ obtained by dividing $ABCD$ are contained in the layer which is the union of such polyhedra, i.e.,

$$m_{i+k} \le 3 \cdot 2^{2(k+1)}. \tag{2.121}$$

The number of tetrahedra of $\mathcal{F}_{i+k}^1$, on which the estimates similar to (2.120) hold for the basis functions φ_E^i, φ_F^i, and φ_G^i, is evaluated from above by the quantity $c_{17} m_{i+k-1}$. Then (2.117), (2.120), and (2.121) result in

$$\begin{gathered}\int\limits_{\widetilde{ABCD}} \left(\left(\frac{\partial \varphi_E^i}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_F^i}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_G^i}{\partial x_1}\right)^2\right) dx \\ \le c_{18} h_i + \sum_{k=1}^{l-i} c_{17} m_{i+k-1} \frac{c_{19} h_{i+k}}{2^{2(k-1)}} \le c_{18} h_i + 12 c_{17} c_{19} \sum_{k=1}^{l-i} h_{i+k}.\end{gathered} \tag{2.122}$$

Let us introduce the quantity $\tilde{h}_j = \sqrt{2} h_0 / 2^j$. Because of (2.62) and (2.75) the estimate

$$h_j \le 2 \tilde{h}_j \tag{2.123}$$

holds. Applying (2.123) and evaluating the sum of the geometrical progression we obtain from (2.122)

$$\begin{gathered}\int\limits_{\widetilde{ABCD}} \left(\left(\frac{\partial \varphi_E^i}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_F^i}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_G^i}{\partial x_1}\right)^2\right) dx \\ \le 2 c_{18} \tilde{h}_i + 12 c_{17} c_{19} \tilde{h}_i \sum_{k=1}^{l-i} \frac{1}{2^{k-1}} \le 2 c_{18} \tilde{h}_i + 24 c_{17} c_{19} \tilde{h}_i \le c_{20} h_{i-1}.\end{gathered} \tag{2.124}$$

Since $u = 0$ on Γ, we have

$$\left|\frac{1}{2}u(A) - u(E)\right| \leq \left|\frac{1}{2}u(A) - \frac{1}{2}u(E)\right| + \left|\frac{1}{2}u(E) - \frac{1}{2}u(B)\right|. \tag{2.125}$$

According to the theorem of embedding $H_2^2(\widetilde{ABCD})$ into $C^{1/2}(\widetilde{ABCD})$ and the definition of the norm in $C^{1/2}(\widetilde{ABCD})$ we have

$$|u(A) - u(E)| = |AE|^{1/2}\frac{|u(A) - u(E)|}{|AE|^{1/2}} \leq c_{21}h_i^{1/2}\|u\|_{2,\widetilde{ABCD}}.$$

The same estimate holds also for the second term in the right-hand side of (2.125). Hence

$$\left|\frac{1}{2}u(A) - u(E)\right| \leq c_{22}h_{i-1}^{1/2}\|u\|_{2,\widetilde{ABCD}}.$$

In the same way we estimate $\left|\frac{1}{2}u(A) - u(F)\right|$ and $\left|\frac{1}{2}u(A) - u(G)\right|$. As a result we obtain

$$\begin{aligned}&\left(\frac{1}{2}u(A) - u(E)\right)^2 + \left(\frac{1}{2}u(A) - u(F)\right)^2 + \left(\frac{1}{2}u(A) - u(G)\right)^2 \\ &\leq c_{23}h_{i-1}\|u\|^2_{2,\widetilde{ABCD}}.\end{aligned} \tag{2.126}$$

From (2.111), (2.124), and (2.126) it follows that

$$\int\limits_{\widetilde{ABCD}} \left(\frac{\partial}{\partial x_1}(\hat{u}_i - \hat{u}_{i-1})\right)^2 dx \leq c_{24}h_{i-1}^2\|u\|^2_{2,\widetilde{ABCD}}. \tag{2.127}$$

The similar inequalities are valid for the derivatives with respect to x_2 and x_3 as well. Since the tetrahedron $ABCD$ is arbitrary, (2.127) holds on the union of elementary tetrahedra obtained by dividing each tetrahedron of $\mathcal{F}_{i-1}^2$. Combining this inequalities we obtain the estimate

$$\|\hat{u}_i - \hat{u}_{i-1}\|^2_{1,\Omega''} \leq c_{25}h_{i-1}^2\|u\|^2_{2,\Omega}. \tag{2.128}$$

By the triangle inequality we have

$$\begin{aligned}&\|u - \hat{u}_{i-1}\|_{1,\Omega''} \leq \|u - \hat{u}_i\|_{1,\Omega''} + \|\hat{u}_i - \hat{u}_{i-1}\|_{1,\Omega''} \\ &\leq \|u - \hat{u}_i\|_{1,\Omega} + \|\hat{u}_i - \hat{u}_{i-1}\|_{1,\Omega''}.\end{aligned} \tag{2.129}$$

The estimate (2.107) is valid for the first term in the right-hand side of (2.129) by the induction hypothesis. To estimate the second term, we use (2.128):

$$\|u - \hat{u}_{i-1}\|_{1,\Omega''} \le c_{26} h_{i-1} \|u\|_{2,\Omega}. \tag{2.130}$$

Combining (2.109), (2.110), and (2.130) we conclude that (2.107) holds for $i-1$. □

Observe that the norms $\|v\|_{1,\Omega}$ and $\|\!|v|\!\|_{\Omega}$ are equivalent, i.e.,

$$c_{27}\|v\|_{1,\Omega} \le \|\!|v|\!\|_{\Omega} \le c_{28}\|v\|_{1,\Omega} \quad \forall\, v \in H_0^1(\Omega). \tag{2.131}$$

By Lemma 25, (2.124), and (2.131) the convergence of the Galerkin solution to the exact one is proved in the usual way (see, for example, [7]).

Lemma 26. *Under the conditions* (2.79) *there exists a unique solution of* (2.85). *It obeys the estimate*

$$\|\!|u - \tilde{v}_i|\!\|_{\Omega} \le c_{29} h_i \|f\|_{0,\Omega}. \tag{2.132}$$

□

2.3.5 The estimate of the eigenvalues of the matrix of the discrete problem.

Lemma 27. *Under the conditions* (2.79) *the eigenvalues* λ *of the matrix* L_i *satisfy the inequality*

$$0 < \lambda \le c_{30} h_i. \tag{2.133}$$

Proof. Consider an arbitrary vector v with components $v(x)$, $x \in \Omega_i$, define it to be zero at the nodes of $\overline{\Omega}_i$ belonging to Γ, and construct the prolongation $\tilde{v} \in \mathcal{H}^i$. We have the equality

$$v^T L_i v = \mathcal{L}(\tilde{v}, \tilde{v}). \tag{2.134}$$

From (2.131) it follows that

$$\mathcal{L}(\tilde{v}, \tilde{v}) \le c_{31}\|v\|^2_{1,\Omega}. \tag{2.135}$$

As in the proof of Lemma 25 we divide Ω into three parts: Ω', Ω'', and ω^δ. Consider an arbitrary tetrahedron $ABCD$ of $\mathcal{F}_i^1$. We have

$$\begin{aligned}\int\limits_{ABCD}\left(\frac{\partial \tilde{v}}{\partial x_1}\right)^2 dx &= \int\limits_{ABCD}\left(v(A)\frac{\partial \varphi_A^i}{\partial x_1} + v(B)\frac{\partial \varphi_B^i}{\partial x_1} + v(C)\frac{\partial \varphi_C^i}{\partial x_1} + v(D)\frac{\partial \varphi_D^i}{\partial x_1}\right)^2 dx \\ &\le \left(v^2(A) + v^2(B) + v^2(C) + v^2(D)\right) \\ &\times \int\limits_{ABCD}\left(\left(\frac{\partial \varphi_A^i}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_B^i}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_C^i}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_D^i}{\partial x_1}\right)^2\right) dx.\end{aligned} \tag{2.136}$$

On $ABCD$ the estimate similar to (2.117) is valid:

$$\int\limits_{ABCD} \left(\frac{\partial \varphi_A^i}{\partial x_1}\right)^2 dx \le c_{32} h_i.$$

Together with (2.136) this yields

$$\int\limits_{ABCD} \left(\frac{\partial \tilde{v}}{\partial x_1}\right)^2 dx \le c_{33} h_i (v^2(A) + v^2(B) + v^2(C) + v^2(D)).$$

The same inequalities are valid for the derivatives with respect to x_2 and x_3 as well. Sum up them over all tetrahedra of $\mathcal{F}_i^1$. Since each node is a vertex of only a few tetrahedra, we get

$$\|\tilde{v}\|^2_{1,\Omega'} \le c_{34} h_i \sum_{y \in \Omega_i} v^2(y). \tag{2.137}$$

Then consider a tetrahedron $EFGH$ of $\mathcal{F}_i^2$. The vertices F, G, and H are assumed to belong Γ. Denote by K, L, and M the midpoints of the edges EF, EG, and EH respectively and by $\widetilde{EFGH}$ the union of elementary tetrahedra obtained by dividing $EFGH$. From (2.84) it follows that

$$\begin{aligned}
&\int\limits_{\widetilde{EFGH}} \left(\frac{\partial \tilde{v}}{\partial x_1}\right)^2 dx = v^2(E) \int\limits_{\widetilde{EFGH}} \left(\frac{\partial \varphi_E^i}{\partial x_1}\right)^2 dx \\
&= v^2(E) \int\limits_{\widetilde{EFGH}} \left(\frac{\partial \varphi_E^{i+1}}{\partial x_1} + \frac{1}{2}\left(\frac{\partial \varphi_K^{i+1}}{\partial x_1} + \frac{\partial \varphi_L^{i+1}}{\partial x_1} + \frac{\partial \varphi_M^{i+1}}{\partial x_1}\right)\right)^2 dx \qquad (2.138) \\
&\le \frac{7}{4} v^2(E) \int\limits_{\widetilde{EFGH}} \left(\left(\frac{\partial \varphi_E^{i+1}}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_K^{i+1}}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_L^{i+1}}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_M^{i+1}}{\partial x_1}\right)^2\right) dx.
\end{aligned}$$

Observe that $\left(\partial \varphi_E^{i+1}/\partial x_1\right)^2$ is not equal to zero on the tetrahedron $ELKM$ of $\mathcal{F}_{i+1}^1$ and vanishes on the remaining part of $\widetilde{EFGH}$. From (2.117) we have

$$\int\limits_{ELKM} \left(\frac{\partial \varphi_E^{i+1}}{\partial x_1}\right)^2 dx \le c_{35} h_i. \tag{2.139}$$

The estimate (2.124) leads to

$$\int\limits_{\widetilde{EFGH}} \left(\left(\frac{\partial \varphi_K^{i+1}}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_L^{i+1}}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_M^{i+1}}{\partial x_1}\right)^2\right) dx \le c_{36} h_i. \tag{2.140}$$

Combine (2.138), (2.139), and (2.140):

$$\int\limits_{\widetilde{EFGH}} \left(\frac{\partial \tilde{v}}{\partial x_1}\right)^2 dx \le c_{37} h_i v^2(E). \tag{2.141}$$

If two vertices of the tetrahedron, namely E and F, lie inside Ω, the estimate (2.141) would be of the form

$$\int\limits_{\widetilde{EFGH}} \left(\frac{\partial \tilde{v}}{\partial x_1}\right)^2 dx \le c_{38} h_i (v^2(E) + v^2(F)).$$

As the result we obtain on Ω'' the estimate

$$\|\tilde{v}\|^2_{1,\Omega''} \le c_{39} h_i \sum_{y \in \Omega_i} v^2(y). \tag{2.142}$$

Since $\tilde{v}$ vanishes on ω^δ, (2.137) with (2.142) yields

$$\|\tilde{v}\|^2_{1,\Omega} \le (c_{34} + c_{39}) h_i \sum_{y \in \Omega_i} v^2(y).$$

Together with (2.134) and (2.135) this leads to

$$v^T L_i v \le c_{31}(c_{34} + c_{39}) h_i \sum_{y \in \Omega_i} v^2(y).$$

Since the matrix L_i is symmetric, due to the Rayleigh ratio the above inequality implies the estimate

$$\lambda \le c_{30} h_i = c_{31}(c_{34} + c_{39}) h_i.$$

The lower estimate follows immediately from positive definiteness of the bilinear form $\mathcal{L}$. □

2.3.6 Convergence of the cascadic algorithm. On the sequence of grids Ω_i, $i = 0, 1, \ldots, l$ we obtain the sequence of problems:

for given $f_i \in M_i$ *find* $v_i \in M_i$ *such that*

$$L_i v_i = f_i.$$

For their solving we use the cascadic iterative algorithm with the conjugate-gradient method or the Jacobi type one (see Section 2.1).

To prove the convergence of the cascadic algorithm besides the results obtained above we need also some auxiliary inequalities.

Lemma 28. *For an isomorphic pair $v_i \in M_i$ and $\tilde{v}_i \in \mathcal{H}^i$ the inequalities*

$$c_{41} h_i^{3/2} \|v\|_i \le \|\tilde{v}\|_{0,\Omega} \le c_{42} h_i^{3/2} \|v\|_i \tag{2.143}$$

hold.

Proof. Consider an arbitrary tetrahedron Δ from $\mathcal{F}_i^1$ with vertices $x_j,\ j = 1, \ldots, 4$. We have

$$(\tilde{v}, \tilde{v})_{0,\Delta} = \sum_{j,l=1}^{4} v(x_j) v(x_l) (\varphi_{x_j}^i, \varphi_{x_l}^i)_{0,\Delta}. \tag{2.144}$$

The mapping which transforms the *reference* tetrahedron Δ' (see Fig. 9) into Δ is linear. Therefore its Jacobian J given by (2.112) is constant. Because of (2.54) there exists a constant c_{43} such that

$$|J| = c_{43} h_i^3.$$

Then

$$(\varphi_{x_j}^i, \varphi_{x_l}^i)_{0,\Delta} = \int_{\Delta} \varphi_{x_j}^i \varphi_{x_l}^i dx = \int_{\Delta'} \Phi_{x'_j}^i \Phi_{x'_l}^i |J| d\eta = \\ c_{43} h_i^3 (\Phi_{x'_j}^i, \Phi_{x'_l}^i)_{0,\Delta'}, \tag{2.145}$$

where $x'_j,\ j = 1, \ldots, 4$ are the vertices of Δ', and $\Phi_{x'_j}^i$ are the usual basis functions corresponding to x'_j. Let us introduce the matrix D with the elements $(\Phi_{x'_j}^i, \Phi_{x'_l}^i)_{0,\Delta'},\quad j, l = 1, \ldots, 4$. Since $\Phi_{x'_j}^i$ are linear independent, the matrix D is nonsingular, symmetric, and positive definite. Hence its eigenvalues are positive, i.e.,

$$0 < c_{44} \le \lambda_{\min}(D) \le \lambda_{\max}(D) \le c_{45}. \tag{2.146}$$

With (2.144) and (2.145) this yields

$$(\tilde{v}, \tilde{v})_{0,\Delta} = c_{43} h_i^3 \sum_{j,l=1}^{4} v(x_j) v(x_l) (\Phi_{x'_j}^i, \Phi_{x'_l}^i)_{0,\Delta'} \le c_{43} c_{45} h_i^3 \sum_{j=1}^{4} v^2(x_j), \tag{2.147}$$

$$(\tilde{v}, \tilde{v})_{0,\Delta} \ge c_{43} c_{44} h_i^3 \sum_{j=1}^{4} v^2(x_j). \tag{2.148}$$

Sum up over all Δ of $\mathcal{F}_i^1$:

$$(\tilde{v}, \tilde{v})_{0,\Omega'} = \sum_{\Delta \in \mathcal{F}_i^1} (\tilde{v}, \tilde{v})_{0,\Delta}.$$

Denote by c_{46} and c_{47} respectively the minimal and the maximal number of tetrahedra being joined at the same node. Since all nodes of Ω_i belong to Ω', (2.147) and (2.148) imply

$$(\tilde{v}, \tilde{v})_{0,\Omega'} \leq \sum_{\Delta \in \mathcal{F}_i^1} c_{43} c_{45} h_i^3 \sum_{j=1}^{4} v^2(x_j) \leq c_{43} c_{45} c_{47} h_i^3 \sum_{x \in \Omega_i} v^2(x),$$

$$(\tilde{v}, \tilde{v})_{0,\Omega'} \geq c_{43} c_{45} c_{46} h_i^3 \sum_{x \in \Omega_i} v^2(x),$$

i.e.,

$$c_{48} h_i^3 \|v\|_i^2 \leq \|\tilde{v}\|_{0,\Omega'}^2 \leq c_{49} h_i^3 \|v\|_i^2. \tag{2.149}$$

Next we consider a tetrahedra of $\mathcal{F}_i^2$, for example, the tetrahedron $ABCD$ shown in Fig. 5. The tetrahedron $AEFG$ obtained by dividing $ABCD$ belongs to $\mathcal{F}_{i+1}^1$. Let us introduce the vector w_j, $j = 1, \dots, 4$ whose components are equal to the values of $\tilde{v}$ at the vertices of $AEFG$, i.e.,

$$w_1 = \tilde{v}(A) = v(A), \qquad w_2 = \tilde{v}(E) = v(A)/2,$$
$$w_3 = \tilde{v}(F) = v(A)/2, \quad w_4 = \tilde{v}(G) = v(A)/2.$$

On $AEFG$ the estimate (2.147) holds. Hence

$$\begin{aligned}(\tilde{v}, \tilde{v})_{0,AEFG} &\leq c_{43} c_{45} h_{i+1}^3 \sum_{j=1}^{4} w_j^2 \\ &= \frac{7}{4} c_{43} c_{45} h_{i+1}^3 v^2(A) \leq c_{50} h_{i+1}^3 v^2(A).\end{aligned} \tag{2.150}$$

The same estimate is valid also on the tetrahedron $EFGM$ of $\mathcal{F}_{i+1}^1$.

As in the proof of Lemma 25, we divide the tetrahedra of $\mathcal{F}_{i+1}^2$ obtained by dividing $ABCD$ into smaller ones. Then we select tetrahedra of $\mathcal{F}_{i+2}^1$ and obtain on them the estimates similar to (2.150). We repeat this procedure up to level l.

Denote by Δ_{i+k} an arbitrary tetrahedron of $\mathcal{F}_{i+k}^1$ obtained by dividing a tetrahedron of $\mathcal{F}_{i+k-1}^2$. The values of the function $\tilde{v}$ at its vertices are equal to $v(A)/2^k$ or $v(A)/2^{k-1}$. Hence

$$(\tilde{v}, \tilde{v})_{0,\Delta_{i+k}} \leq c_{50} h_{i+k}^3 v^2(A) \frac{1}{2^{2(k-1)}}.$$

The number of such tetrahedra does not exceed the quantity $c_{17} m_{i+k-1}$ where m_{i+k-1} satisfies the estimate (2.121) (see the proof of Lemma 25).

Then

$$(\tilde{v}, \tilde{v})_{0,\widetilde{ABCD}} \le c_{51} v^2(A) \sum_{k=1}^{l-i} h_{i+k}^3 \frac{1}{2^{2(k-1)}} m_{i+k-1}, \tag{2.151}$$

where $\widetilde{ABCD}$ is the polyhedron that is formed by uniting the elementary tetrahedra obtained by dividing $ABCD$. The inequality (2.123) implies

$$h_{i+k} \le 2\tilde{h}_{i+k} = \frac{\tilde{h}_i}{2^{k-1}} \le \frac{h_i}{2^{k-1}}.$$

Applying to (2.151) this inequality together with (2.121) and evaluating the sum of the geometric progression, we get

$$(\tilde{v}, \tilde{v})_{0,\widetilde{ABCD}} \le 3c_{51} v^2(A) h_i^3 \sum_{k=1}^{l-i} \frac{1}{2^{3k-5}} \le \frac{96}{7} c_{51} v^2(A) h_i^3.$$

Consider an arbitrary tetrahedron $KLMN$ of $\mathcal{F}_i^2$ which has only two vertices belonging to Γ. Then

$$(\tilde{v}, \tilde{v})_{0,KLMN} \le c_{52}(v^2(K) + v^2(L)) h_i^3,$$

where K and L are the vertices lying inside Ω.

As a result we obtain on Ω'' the estimate

$$(\tilde{v}, \tilde{v})_{0,\Omega''} \le c_{53} h_i^3 \sum_{x \in \Omega_i} v^2(x) = c_{53} h_i^3 \|v\|_i^2.$$

Combining this inequality with the latter inequality (2.149) and taking into account that $\tilde{v}$ vanishes on ω^δ, we obtain the latter inequality (2.143). The former inequality (2.149) yields

$$\|\tilde{v}\|_{0,\Omega}^2 \ge \|\tilde{v}\|_{0,\Omega'}^2 \ge c_{48} h_i^3 \|v\|_i^2. \qquad \square$$

Thus we have proved all the conditions being used in the proof of the convergence of the cascadic algorithm in the Section 2.1. Therefore we formulate the main result without proof.

Theorem 29. *Let the conditions* (2.79) *be fulfilled for the problem* (2.77)–(2.78) *on a convex bounded domain $\Omega \subset R^3$ with a boundary Γ of the class C^2. Then for the solution u_l of the cascadic algorithm with the conjugate-gradient method or the Jacobi-type iterations on each level $j = 1, \ldots, l$ the estimate*

$$|\!|\!| v_l - u_l |\!|\!|_l \le d_1 \sum_{j=1}^{l} \frac{h_{j-1}}{2m_j + 1} \|f\|_{0,\Omega} \tag{2.152}$$

holds. $\square$

A number of iterations in the cascadic algorithm is chosen in the same way as in the Section 2.3.

Theorem 30. *Under the hypotheses of Theorem 29 the error of the cascadic algorithm with the conjugate-gradient method or the Jacobi-type iterations is estimated as*

$$\|u_l - v_l\|_l \leq \frac{d_5 h_l}{2m+1} \|f\|_{0,\Omega}. \tag{2.153}$$

The prolongation $\tilde{u}_l \in \mathcal{H}^l$ of the vector u_l obeys the estimate

$$\|\tilde{u}_l - u\|_\Omega \leq h_l \left(c_{29} + \frac{d_5}{2m+1} \right) \|f\|_{0,\Omega}. \tag{2.154}$$

The number of arithmetic operations is evaluated from above by

$$S_l \leq (d_6 m + d_7) n_l \tag{2.155}$$

with constants d_5, c_{29}, d_6, d_7 which are independent of m and n_l. □

3 Numerical results

3.1 Dependence of the convergence rate of the V-cycle upon smoothing

3.1.1 Preliminary remarks. In recent years many variants of multigrid methods have been developed. In [3] additive and multiplicative multigrid algorithms were compared. Theoretical as well as numerical results show the superiority of the multiplicative multigrid algorithm over the additive one on sequential computers. For parallel computers corresponding results are obtained in [2].

The usual V-cycle is the typical multiplicative algorithm. In the numerical tests we investigated the dependence of the convergence rate of this algorithm upon a kind of iterative process and upon the application of pre-smoothing and post-smoothing iterations.

It has been often argued in literature that the convergence rate does not depend on whether only pre-smoothing, only post-smoothing, or a combination of pre- and post-smoothing iterations are applied. Indeed, theoretically for all these variants the same upper bound is valid [25]. But in practice, as our tests show, these strategies yield different results.

Concerning the choice of iterative process, presented numerical results demonstrate that using the Jacobi-type iterations with the special Chebyshev parameters improves the convergence of the algorithm in comparison with the Jacobi-type iterations with constant iterative parameters.

3.1.2 Formulation of the multigrid algorithm. Assume that we have a sequence of finite-dimensional vector spaces

$$M_0 \subset M_1 \subset \ldots \subset M_l.$$

Assume also that interpolation operators

$$I_i : M_i \to M_{i+1}, \quad i = 0, \ldots, l-1,$$

restriction operators

$$R_i : M_{i+1} \to M_i, \quad i = 0, \ldots, l-1,$$

and linear invertible operators

$$L_i : M_i \to M_i, \quad i = 0, \ldots, l$$

are given.

The main objective of the multigrid algorithm is to solve the problem:

$$\begin{aligned} &for\ given\ \ f_l \in M_l\ \ find\ \ v_l \in M_l\ \ satisfying \\ &L_l v_l = f_l. \end{aligned}$$

As a smoothing operator

$$S_i : M_i \to M_i, \quad i = 1, \ldots, l$$

we consider one of two iterative processes.
Jacobi-type iterations damped by 1/2 (ν iterations);
Procedure $S_i^{(\nu)}(L_i, w_i, f_i)$;

$$\begin{aligned} &y_0 = w_i; \\ &if\ \nu \neq 0\ then \\ &\quad \{\ for\ k = 1, \ldots, \nu\ do \\ &\quad\quad y_k = y_{k-1} - N_i(L_i y_{k-1} - f_i);\}; \\ &set\ S_i^{(\nu)} = y_\nu. \end{aligned} \tag{3.1}$$

The matrix N_i is defined by

$$N_i = \frac{1}{2} D_i^{-1}, \quad D_i : \text{diagonal of } L_i.$$

Jacobi-type iterations with the special Chebyshev parameters (ν iterations);
Procedure $S_i^{(\nu)}(L_i, w_i, f_i)$;

$$
\begin{aligned}
& y_0 = w_i; \\
& if\ \nu \neq 0\ then \\
& \quad for\ k = 1, \dots, \nu\ do \\
& \qquad \left\{ \tau_{k-1} = \frac{1}{\lambda_i^*} \cos^{-2} \frac{\pi(2k-1)}{2(2\nu+1)}; \right. \\
& \qquad\qquad \left. y_k = y_{k-1} - \tau_{k-1}(L_i y_{k-1} - f_i); \right\}; \\
& set\ S_i^{(\nu)} = y_\nu .
\end{aligned}
\tag{3.2}
$$

Here λ_i^* is the upper bound of eigenvalues of L_i.

Now we can formulate the multigrid algorithm.

Multigrid algorithm (V-cycle);
Procedure $MG(i, u_i, L_i, f_i)$;

$$
\begin{aligned}
& if\ i = 0\ then\ u_i = L_i^{-1} f_i \\
& \qquad else\ \{\ w_i = 0; \\
& \qquad\qquad u_i = S_i^{(\nu_1)}\ (L_i, w_i, f_i); \\
& \qquad\qquad g_{i-1} = R_{i-1}(L_i u_i - f_i); \\
& \qquad\qquad MG(i-1, q_{i-1}, L_{i-1}, g_{i-1}); \\
& \qquad\qquad w_i = u_i - I_{i-1} q_{i-1}; \\
& \qquad\qquad u_i = S_i^{(\nu_2)}(L_i, w_i, f_i); \}.
\end{aligned}
\tag{3.3}
$$

3.1.3 Numerical tests.

Consider the Poisson equation on the unit square:

$$
\begin{aligned}
-\Delta u &= f \ \text{ in } \Omega = (0,1) \times (0,1), \\
u &= 0 \ \text{ on } \Gamma.
\end{aligned}
$$

The right-hand side is of the form

$$f = (12x^2 - 2) y(1-y) + 2x^2(1-x^2).$$

This problem has the exact solution

$$u = x^2(1-x^2) y(1-y).$$

We set up a sequence of uniform triangulations on Ω in the usual way. The coarsest grid consists of 3×3 nodes and the finest one of 255×255 nodes, i.e., $l = 6$. On each grid the discrete system

$$L_i v_i = f_i$$

is constructed. The matrix L_i is the discrete Laplacian symbolized by the stencil

$$\begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix}.$$

Denote by $\varepsilon_0 = v_l - w_l$ the error of the initial guess $w_l = 0$ and by $\varepsilon = v_l - u_l$ the error of the approximate solution u_l obtained by the multigrid algorithm (3.3). The error reduction is characterized by the quantity

$$\rho = \frac{|\!|\!|\varepsilon|\!|\!|_l}{|\!|\!|\varepsilon_0|\!|\!|_l}$$

where $|\!|\!|\cdot|\!|\!|_l$ is the energy norm defined by

$$|\!|\!|v|\!|\!|_l = (v^T L_l v)^{1/2} \qquad \forall\, v \in M_l.$$

In the tests we investigated three versions of the multigrid algorithm (3.3):

a) applying only pre-smoothing iterations, i.e., $\nu_1 = \nu$ and $\nu_2 = 0$;

b) applying only post-smoothing iterations, i.e., $\nu_1 = 0$ and $\nu_2 = \nu$;

c) applying the combination of $\nu_1 = \nu/2$ pre- and $\nu_2 = \nu/2$ post-smoothing steps, supposing ν to be even (the symmetric V-cycle).

To calculate the value of ρ, we obtained the exact solution v_l of the discrete system by the Gauss elimination.

Tables 1 and 2 contain the results of tests for the iterative processes (3.1) and (3.2) respectively.

Table 1 shows that for the iterative process (3.1) the symmetric V-cycle and the application of only post-smoothing iterations yield the values of ρ which differ only slightly for the same values of ν. Besides, we see that these two versions of the algorithm (3.3) do better than the application of only pre-smoothing iterations.

From Table 2 we see that three versions of the algorithm (3.3) using the iterative process (3.2) give different results. As in the case of the iterative process (3.1), the application of only pre-smoothing iterations is the least efficient. But the application of only post-smoothing iterations has a considerable advantage over the symmetric V-cycle.

Table 1.
Dependence of ρ upon the number of pre- and post-smoothing iteration steps for the iterative process (3.1)

ν_1	ν_2	ρ
2	0	0.401
0	2	0.232
1	1	0.254
4	0	0.346
0	4	0.178
2	2	0.184
8	0	0.281
0	8	0.128
4	4	0.128

Table 2.

Dependence of ρ upon the number of pre- and post-smoothing iteration steps for the iterative process (3.2)

ν_1	ν_2	ρ
2	0	0.340
0	2	0.169
1	1	0.217
4	0	0.220
0	4	0.081
2	2	0.121
8	0	0.118
0	8	0.019
4	4	0.050

Comparing the results listed in Tables 1 and 2, we see the superiority of the iterative process (3.2) over (3.1).

It is well known that on the single grid the conjugate gradient method converges faster than the iterative process (3.2). We used the conjugate gradient method as a smoother in the multigrid algorithm (3.3) applying only post-smoothing iterations. The results listed in Table 3 show that in the multigrid algorithm the conjugate gradient method has no advantage over the iterative process (3.2) in convergence. In addition, the computational expence of the iterative process (3.2) is lower.

Basing on the obtained numerical results we can make the following conclusion. The application of only post-smoothing iterations using the iterative process (3.2) enables to improve the convergence of the multigrid algorithm (3.3).

Table 3.
Dependence of ρ upon the number of post-smoothing iteration steps.

ν_1	ν_2	ρ(conjugate gradient method)	ρ (iterative process (3.2))
0	2	0.168	0.169
0	4	0.076	0.081
0	8	0.025	0.019

3.2 Cascadic algorithm for the Poisson equation

3.2.1 Two-step semi-iterative process. The theoretical estimates of the convergence of the cascadic algorithm were obtained for the conjugate gradient method and the Jacobi-type iterative process with the special parameters used as a smoother. Unfortunately, in practice the latter iterative process is unstable for a large number of iteration steps due to round-off errors.

One way to overcome this problem is to re-order the iterative parameters [30]. However, the procedure of re-ordering is complicated and leads to a large programming overhead.

Another way of ensuring the numerical stability of iterative process is to apply the two-step semi-iterative process [38]. We formulate this algorithm for a self-adjoint and positive definite matrix L_i.

Two-step semi-iterative process (m_i iteration steps on level i);
Procedure $S_i(L_i, w_i, f_i)$;

$$\begin{aligned}
&y_0 = w_i;\\
&\beta_0 = \frac{4}{3\lambda_i^*};\ y_1 = y_0 - \beta_0(L_i y_0 - f_i);\\
&for\ k = 2, \ldots, m_i\ do\\
&\qquad \Big\{\ \alpha_{k-1} = \frac{2(2k-1)}{2k+1};\ \beta_{k-1} = \frac{2}{\lambda_i^*};\\
&\qquad\quad y_k = \alpha_{k-1}(y_{k-1} - \beta_{k-1}(L_i y_{k-1} - f_i)) + (1-\alpha_{k-1})y_{k-2};\ \Big\}\\
&set\ S_i = y_{m_i}.
\end{aligned} \tag{3.4}$$

Here λ_i^* is the maximal eigenvalue of L_i.

In the Jacobi-type iterative process the value of τ_k and thus the approximate solution y_k depend on the number m_i of iteration steps. Denote by $y_k^{(m_i)}$ the approximate solution obtained on the k-th step of the Jacobi-type iterative process. The parameters α_k and β_k in (3.4) are chosen in such a way that the sequence of the approximate solutions $\{y_k\}_{k=1}^{m_i}$ obtained by (3.4) coincides with $\{y_k^{(k)}\}_{k=i}^{m_i}$. Consequently, the results on the convergence of the cascadic algorithm obtained for the Jacobi-type iterative process hold for the two-step semi-iterative process (3.4). In addition, the algorithm (3.4) is numerically stable.

3.2.2 Dependence of the convergence on the number of iteration steps. Consider the Poisson equation on the unit square:

$$-\Delta u = f \quad \text{in } \Omega = (0,1)\times(0,1), \tag{3.5}$$

$$u = 0 \quad \text{on } \Gamma. \tag{3.6}$$

For $H_2^{1+\alpha}$-regular problems ($0 < \alpha \le 1$) the error of the cascadic algorithm satisfies the estimate [5]

$$|||u_l - v_l|||_l \le c\frac{h_l^\alpha}{m_l^\alpha}\|f\|_{\Omega,\alpha-1} \tag{3.7}$$

where u_l is the approximate solution obtained by the cascadic algorithm and v_l is the exact solution of the discrete problem. In our numerical experiments we wanted to investigate the dependence of the cascadic error on the number of iteration steps m_l for different classes of regularity, i.e., for different values of α.

The conjugate gradient method as well as the two-step semi-iterative process were used in the cascadic algorithm as a smoother. The number of iteration steps on the i-th level was chosen as the smallest integer satisfying the inequality

$$2m_i + 1 \ge (2m_l + 1)\left(\frac{n_l h_i}{n_i h_l}\right)^{1/2}.$$

To obtain the exact solution of the discrete problem the Gauss elimination was applied.

First we consider the problem (3.5)–(3.6) with the right-hand side

$$f = -\frac{25}{256}x^{-7/16}(9-41x)y^2(1-y) - 2x^{25/16}(1-x)(1-3y). \tag{3.8}$$

This problem has the exact solution

$$u = x^{25/16}(1-x)y^2(1-y) \tag{3.9}$$

which belongs to $H_2^2(\Omega)$.

Table 4.
Results of the cascadic algorithm
for the problem (3.5), (3.6), (3.8) ($m_l = 30$).

i	Grid	Two-step semi-iterative process		Conjugate gradient method		$\|v_i - u_i^*\|_i$
		$\|u_i^* - u_i\|_i$	$\|u_i - v_i\|_i$	$\|u_i^* - u_i\|_i$	$\|u_i - v_i\|_i$	
1	31×31	$2.227 \cdot 10^{-4}$	$7.375 \cdot 10^{-6}$	$2.228 \cdot 10^{-4}$	$5.452 \cdot 10^{-16}$	$2.228 \cdot 10^{-4}$
2	63×63	$1.094 \cdot 10^{-4}$	$9.497 \cdot 10^{-6}$	$1.089 \cdot 10^{-4}$	$4.756 \cdot 10^{-15}$	$1.089 \cdot 10^{-4}$
3	127×127	$5.352 \cdot 10^{-5}$	$1.196 \cdot 10^{-5}$	$5.302 \cdot 10^{-5}$	$6.171 \cdot 10^{-6}$	$5.274 \cdot 10^{-5}$
4	255×255	$3.651 \cdot 10^{-5}$	$2.198 \cdot 10^{-5}$	$3.559 \cdot 10^{-5}$	$1.996 \cdot 10^{-5}$	$2.540 \cdot 10^{-5}$

Table 4 shows the results of the cascadic algorithm for $l = 4, h_l = \sqrt{2}/256$, and $m_l = 30$. The second column contains the number of nodes of the i-th grid. In the columns 3 through 7 the grid norms of errors are listed. Here u_i^* denotes the vector of values of the exact solution (3.9) at the nodes of the i-th grid. The columns 3 and 5 contain the total error and the columns 4 and 6 contain the cascadic one for the two-step semi-iterative process (3.4) and the conjugate gradient method respectively. The last column contains the discretization error.

In Tables 5 and 6 the total and cascadic errors are listed for the two-step semi-iterative process and the conjugate gradient method respectively.

Table 5.
Dependence of the total and cascadic errors on m_l
for the problem (3.5), (3.6), (3.8) (two-step semi-iterative process).

i	Grid	$m_l = 20$		$m_l = 40$	
		$\|u_i^* - u_i\|_i$	$\mu_i - v_i\|_i$	$\|u_i^* - u_i\|_i$	$\|u_i - v_i\|_i$
1	31×31	$2.227 \cdot 10^{-4}$	$1.053 \cdot 10^{-5}$	$2.225 \cdot 10^{-4}$	$5.687 \cdot 10^{-6}$
2	63×63	$1.085 \cdot 10^{-4}$	$1.278 \cdot 10^{-5}$	$1.090 \cdot 10^{-4}$	$7.711 \cdot 10^{-6}$
3	127×127	$5.726 \cdot 10^{-5}$	$2.198 \cdot 10^{-5}$	$5.358 \cdot 10^{-5}$	$1.201 \cdot 10^{-5}$
4	255×255	$4.811 \cdot 10^{-5}$	$3.610 \cdot 10^{-5}$	$3.272 \cdot 10^{-5}$	$1.727 \cdot 10^{-5}$

Table 6.
Dependence of the total and cascadic errors on m_l
for the problem (3.5), (3.6), (3.8) (conjugate gradient method).

i	Grid	$m_l = 20$		$m_l = 40$	
		$\|u_i^* - u_i\|_i$	$\|u_i - v_i\|_i$	$\|u_i^* - u_i\|_i$	$\|u_i - v_i\|_i$
1	31×31	$2.228 \cdot 10^{-4}$	$5.452 \cdot 10^{-16}$	$2.228 \cdot 10^{-4}$	$5.452 \cdot 10^{-16}$
2	63×63	$1.089 \cdot 10^{-4}$	$1.725 \cdot 10^{-9}$	$1.089 \cdot 10^{-4}$	$1.597 \cdot 10^{-15}$
3	127×127	$5.666 \cdot 10^{-5}$	$1.853 \cdot 10^{-5}$	$5.267 \cdot 10^{-5}$	$1.493 \cdot 10^{-6}$
4	255×255	$4.588 \cdot 10^{-5}$	$3.262 \cdot 10^{-5}$	$3.224 \cdot 10^{-5}$	$1.568 \cdot 10^{-5}$

Comparing the values of the cascadic error for different values of m_l, we see that the cascadic error decreases proportionally with increasing m_l according to (3.7) with $\alpha = 1$.

The presented results show that on the coarser grids the conjugate gradient method is considerably more efficient than the two-step semi-iterative process. However, on the finest grid the conjugate gradient method gives only slightly better results.

Comparing the cascadic error with the discretization one, we see that the value of m_l should be taken about 30. Further increasing m_l does not significantly reduce the value of the total error.

Next we investigate the convergence of the cascadic algorithm for the $H_2^{3/2+\varepsilon}$ – regular problem with a small ε, i.e., $\varepsilon << 3/2$. Consider the problem (3.5)–(3.6) with the right-hand side

$$f = -\frac{17}{256}x^{-15/16}(1-33x)y^2(1-y) - 2x^{17/16}(1-x)(1-3y). \qquad (3.10)$$

It has the exact solution

$$u = x^{17/16}(1-x)y^2(1-y).$$

The results of the cascadic algorithm for this problem are listed in Tables 7–9.

For the cascadic error we have the estimate (3.7) with $\alpha \approx 1/2$, i.e., the cascadic error has to be inversely proportional to $\sqrt{m_l}$ for fixed h_l. However, our experimental results show that in fact the cascadic algorithm converges significantly faster. From Tables 8 and 9 we see that doubling the number m_l of iteration steps results in the reduction of the cascadic error by a factor close to 1/2. Thus, the cascadic error is almoust inversely proportional to m_l rather than to $\sqrt{m_l}$.

Comparing the results for different smoothers, we see that the conjugate gradient method is slightly more efficient.

The last test is concerned with the $H_2^{1+\varepsilon}$ – regular problem with a small ε. We consider the problem (3.5)–(3.6) with the right-hand side

$$f = \frac{9}{256}x^{-7/16}(7x^{-1}+25)y^2(1-y) - 2x^{9/16}(1-x)(1-3y). \qquad (3.11)$$

Its exact solution is of the form

$$u = x^{9/16}(1-x)y^2(1-y).$$

Tables 10–12 contain numerical results for this problem.

Table 7.

Results of the cascadic algorithm
for the problem (3.5), (3.6), (3.10) ($m_l = 30$).

i	Grid	Two-step semi-iterative process		Conjugate gradient method		$\|\|v_i - u_i^*\|\|_i$
		$\|\|u_i - v_i\|\|_i$	$\|\|u_i^* - u_i\|\|_i$	$\|\|u_i - v_i\|\|_i$	$\|\|v_i - u_i^*\|\|_i$	
1	31×31	$3.251 \cdot 10^{-4}$	$9.246 \cdot 10^{-6}$	$3.257 \cdot 10^{-4}$	$7.559 \cdot 10^{-16}$	$3.257 \cdot 10^{-4}$
2	63×63	$2.268 \cdot 10^{-4}$	$1.166 \cdot 10^{-5}$	$2.258 \cdot 10^{-4}$	$3.557 \cdot 10^{-15}$	$2.258 \cdot 10^{-4}$
3	127×127	$1.550 \cdot 10^{-4}$	$1.447 \cdot 10^{-5}$	$1.546 \cdot 10^{-4}$	$7.051 \cdot 10^{-6}$	$1.547 \cdot 10^{-4}$
4	255×255	$1.139 \cdot 10^{-4}$	$2.989 \cdot 10^{-5}$	$1.134 \cdot 10^{-4}$	$2.704 \cdot 10^{-5}$	$1.054 \cdot 10^{-4}$

Table 8.

Dependence of the total and cascadic errors on m_l
for the problem (3.5), (3.6), (3.10) (two-step semi-iterative process).

i	Grid	$m_l = 20$		$m_l = 40$	
		$\|\|u_i^* - u_i\|\|_i$	$\|\|u_i - v_i\|\|_i$	$\|\|u_i^* - u_i\|\|_i$	$\|\|u_i - v_i\|\|_i$
1	31×31	$3.269 \cdot 10^{-4}$	$1.294 \cdot 10^{-5}$	$3.262 \cdot 10^{-4}$	$7.060 \cdot 10^{-6}$
2	63×63	$2.252 \cdot 10^{-4}$	$1.598 \cdot 10^{-5}$	$2.259 \cdot 10^{-4}$	$8.871 \cdot 10^{-6}$
3	127×127	$1.563 \cdot 10^{-4}$	$2.627 \cdot 10^{-5}$	$1.550 \cdot 10^{-4}$	$1.364 \cdot 10^{-5}$
4	255×255	$1.220 \cdot 10^{-4}$	$4.625 \cdot 10^{-5}$	$1.107 \cdot 10^{-4}$	$2.347 \cdot 10^{-5}$

Table 9.

Dependence of the total and cascadic errors on m_l
for the problem (3.5), (3.6), (3.10) (conjugate gradient method).

i	Grid	$m_l = 20$		$m_l = 40$	
		$\|\|u_i^* - u_i\|\|_i$	$\|\|u_i - v_i\|\|_i$	$\|\|u_i^* - u_i\|\|_i$	$\|\|u_i - v_i\|\|_i$
1	31×31	$3.257 \cdot 10^{-4}$	$7.559 \cdot 10^{-16}$	$3.257 \cdot 10^{-4}$	$7.559 \cdot 10^{-16}$
2	63×63	$2.258 \cdot 10^{-4}$	$2.398 \cdot 10^{-9}$	$2.258 \cdot 10^{-4}$	$2.210 \cdot 10^{-15}$
3	127×127	$1.562 \cdot 10^{-4}$	$1.982 \cdot 10^{-5}$	$1.545 \cdot 10^{-4}$	$2.322 \cdot 10^{-6}$
4	255×255	$1.210 \cdot 10^{-4}$	$4.144 \cdot 10^{-5}$	$1.106 \cdot 10^{-4}$	$2.124 \cdot 10^{-5}$

Table 10.

Results of the cascadic algorithm
for the problem (3.5), (3.6), (3.11) ($m_l = 30$)

i	Grid	Two-step semi-iterative process		Conjugate gradient method		$\|v_i - u_i^*\|_i$
		$\|u_i^* - u_i\|_i$	$\|u_i - v_i\|_i$	$\|u_i^* - u_i\|_i$	$\|u_i - v_i\|_i$	
1	31×31	$2.150 \cdot 10^{-2}$	$5.172 \cdot 10^{-5}$	$2.149 \cdot 10^{-2}$	$1.124 \cdot 10^{-15}$	$2.149 \cdot 10^{-2}$
2	63×63	$2.111 \cdot 10^{-2}$	$1.620 \cdot 10^{-4}$	$2.112 \cdot 10^{-2}$	$4.133 \cdot 10^{-14}$	$2.112 \cdot 10^{-2}$
3	127×127	$2.049 \cdot 10^{-2}$	$5.427 \cdot 10^{-4}$	$2.052 \cdot 10^{-2}$	$5.326 \cdot 10^{-4}$	$2.048 \cdot 10^{-2}$
4	255×255	$2.020 \cdot 10^{-2}$	$1.935 \cdot 10^{-3}$	$2.022 \cdot 10^{-2}$	$2.004 \cdot 10^{-3}$	$1.974 \cdot 10^{-2}$

Table 11.

Dependence of the total and cascadic errors on m_l
for the problem (3.5), (3.6), (3.11) (two-step semi-iterative process)

i	Grid	$m_l = 20$		$m_l = 40$	
		$\|u_i^* - u_i\|_i$	$\|u_i - v_i\|_i$	$\|u_i^* - u_i\|_i$	$\|u_i - v_i\|_i$
1	31×31	$2.148 \cdot 10^{-2}$	$7.265 \cdot 10^{-5}$	$2.148 \cdot 10^{-2}$	$4.156 \cdot 10^{-5}$
2	63×63	$2.112 \cdot 10^{-2}$	$2.525 \cdot 10^{-4}$	$2.111 \cdot 10^{-2}$	$9.503 \cdot 10^{-5}$
3	127×127	$2.058 \cdot 10^{-2}$	$1.018 \cdot 10^{-3}$	$2.048 \cdot 10^{-2}$	$3.829 \cdot 10^{-4}$
4	255×255	$2.057 \cdot 10^{-2}$	$2.767 \cdot 10^{-3}$	$2.004 \cdot 10^{-2}$	$1.538 \cdot 10^{-3}$

Table 12.

Dependence of the total and cascadic errors on m_l
for the problem (3.5), (3.6), (3.11) (conjugate gradient method).

i	Grid	$m_l = 20$		$m_l = 40$	
		$\|u_i^* - u_i\|_i$	$\|u_i - v_i\|_i$	$\|u_i^* - u_i\|_i$	$\|u_i - v_i\|_i$
1	31×31	$2.149 \cdot 10^{-2}$	$1.124 \cdot 10^{-15}$	$2.149 \cdot 10^{-2}$	$1.124 \cdot 10^{-15}$
2	63×63	$2.112 \cdot 10^{-2}$	$3.207 \cdot 10^{-8}$	$2.112 \cdot 10^{-2}$	$3.232 \cdot 10^{-15}$
3	127×127	$2.059 \cdot 10^{-2}$	$9.172 \cdot 10^{-4}$	$2.050 \cdot 10^{-2}$	$3.127 \cdot 10^{-4}$
4	255×255	$2.056 \cdot 10^{-2}$	$2.722 \cdot 10^{-3}$	$2.006 \cdot 10^{-2}$	$1.590 \cdot 10^{-3}$

As in the case of the $H_2^{3/2+\varepsilon}$-regular problem, in practice the cascadic algorithm converges faster than the theoretical estimate (3.7) suggests.

Unlike the above two tests, the conjugate gradient method has no advantage over the two-step semi-iterative process. For the $H_2^{1+\varepsilon}$-regular problem both smoothers yield almoust equal values of the cascadic error.

The numerical experiments allow us to make the following conclusions. The estimate (3.7) is valid for H_2^2-regular problems. For $H_2^{1+\alpha}$-regular problems with $\alpha < 1$ the cascadic algorithm converges faster than might be expected from (3.7). Besides, the experiments demonstrated high efficiency of the two-step semi-iterative process (3.4).

References

1. Bakhvalov N.S. On the convergence of a relaxation method with natural constraints of the elliptic operator. – USSR Comput. Math. and Math. Phys., 1966, V. 6, № 5, P. 101–135.
2. Bastian P. Adaptive parallele Mehrgitterverfahren. – Teubner Skripten zur Numerik. Stuttgart: Teubner, 1996.
3. Bastian P., Hackbush W., Wittum G. Additive and multiplicative multi-grid – a comparison. – Computing, 1998, № 60, P. 345–364.
4. Bey J. Tetrahedral grid refinement. – Computing, 1995, № 55, P. 355-378.
5. Bornemann F.A. On the convergence of cascadic iterations for elliptic problems. – Preprint SC 94–8, Berlin: Konrad–Zuse–Zentrum, 1994.
6. Bornemann F.A., Deuflhard P. The cascadic multigrid method for elliptic problems. – Numer. Math., 1996, 75, № 2, P. 135–152.
7. Ciarlet P. The Finite Element Method for Elliptic Problems. – Amsterdam: North-Holland Publishing Company, 1978.
8. Deuflhard P. Cascadic conjugate gradient methods for elliptic partial differential equations. I. Algorithm and numerical results. – Technical Report SC 93–23, Berlin: Konrad–Zuse–Zentrum, 1993.
9. Deuflhard P. Cascadic conjugate gradient methods for elliptic partial differential equations. Algorithm and numerical results. – In: Proceedings of 7th International Conference on Domain Decomposition Methods 1993. – Providence: AMS, 1994.
10. Faux I.D., Pratt M.J. Computational Geometry for Design and Manufacture. – N.Y.: John Wiley, 1979.
11. Fedorenko R.P. A relaxation method for solving elliptic difference equations. – USSR Comput. Math. and Math. Phys., 1961, V. 1, № 5, P. 1092–1096.
12. Fedorenko R.P. The speed of convergence of one iterative process. – USSR Comput. Math. and Math. Phys., 1964, V. 4, № 3, P. 227–235.
13. Fichera G. Existence Theorems in Elasticity. – Berlin: Springer, 1972.
14. Gajewski H., Gröger K., Zacharias K. Nichtlineare Operatorgleichungen und Operatordifferentzialgleichungen. – Berlin: Academie-Verlag, 1974.

15. Gilyova L.V. A cascadic multigrid algorithm in the finite element method for the three-dimensional Dirichlet problem (in Russian). – Siberian J. of Numer. Mathematics, 1998, V. 1, № 3, P. 217–226.
16. Gilyova L.V., Shaidurov V.V. A cascadic multigrid algorithm in the finite element method for the plane elasticity problem (in Russian). – Preprint № 15, Krasnoyarsk: Computing Center of Sib. Branch of Russ. Acad. of Sci., 1996, 15 P.
17. Gilyova L.V., Shaidurov V.V. A cascadic multigrid algorithm in the finite element method for the plane elasticity problem. – Preprint№ 296, Berlin: Weierstrass-Institut für Angewandte Analysis und Stochastik, 1996, 13 P.
18. Gilyova L.V., Shaidurov V.V. A cascadic multigrid algorithm in the finite element method for the plane elasticity problem. – East-West J. Numer. Math., 1997, V. 5,№ 1, P. 143–156.
19. Gilyova L.V., Shaidurov V.V. A cascadic multigrid algorithm in the finite element method for the elasticity problem (in Russian). – In: Theses of the International Conference on Mathematical Models and Methods of Their Investigation. – Krasnoyarsk: Krasnoyarsk State University, 1997, P. 61–62.
20. Gilyova L.V., Shaidurov V.V. A cascadic multigrid algorithm in the finite element method for an indefinite-sign problem (in Russian). – In: Theses of the Third Siberian Congress on Applied and Industrial Mathematics, Part 2. – Novosibirsk: Institute of Mathematics of Sib. Branch of Russ. Acad. of Sci., 1998, P. 10.
21. Gilyova L.V., Shaidurov V.V. A cascadic multigrid algorithm in the finite element method for an indefinite-sign elliptic problem. – 5-th European Multigrid Conference. Special Topics and Applications. – Stuttgart: Institut für Computeranwendungen der Universität Stuttgart, 1998, P. 74-89.
22. Gilyova L.V., Shaidurov V.V. A cascade algorithm for solving a discrete analogue of weakly nonlinear elliptic equation. – Technical Report IOKOMO-01-98, Dresden: Technische Universität Dresden, Fakultät für Mathematik und Naturwissenschaften, 1998, 12 P.
23. Gilyova L.V., Shaidurov V.V. A cascade algorithm for solving a discrete analogue of weakly nonlinear elliptic equation. – Russ. J. Numer. Anal. Math. Modelling, 1999, V. 14, № 1, P. 59-69.
24. Gilyova L.V., Shaidurov V.V. Justification of asymptotic stability of the triangulation algorithm for the three-dimensional domain (in Russian). – Siberian J. of Numer. Mathematics, 2000, V. 3,№ 2, P. 123–136.
25. Hackbush W. Multi-Grid Methods and Applications. – Berlin, N.Y.: Springer Verlag, 1985.
26. Korn G.A., Korn T.M. Mathematical Handbook for Scientists and Engeneers. – N.Y., London: McGraw-Hill, 1961.
27. Ladyzhenskaya O.A., Uraltseva N.N., Linear and Quasilinear Equations of Elliptic Type (in Russian). – Moscow: Nauka, 1973.
28. Oganesjan L.A., Rivkind V.Ya., Rukhovets L.A. Variation-Difference Methods for the Solution of Elliptic Equations, Part 2 (in Russian). – Proceedings

of the seminar on differential equations and their applications. – Vilnius: Institute of Physics and Mathematics of Acad. of Sci. of the Lithuanian SSR, 1974.
29. Oganesjan L.A., Rukhovets L.A. Variation-Difference Methods for the Solution of Elliptic Equations (in Russian). – Erevan: Acad. of Sci. of the Armenian SSR, 1979.
30. Samarskii A.A., Nikolaev E.S. Methods of Solving Grid Equations (in Russian). – Moscow: Nauka, 1978.
31. Shaidurov V.V. Multigrid Methods for Finite Elements (in Russian). – Moscow: Nauka, 1989.
32. Shaidurov V.V. Multigrid Methods for Finite Elements. – Dordrecht: Kluwer Academic Publishers, 1995.
33. Shaidurov V.V. Some estimates of the rate of convergence for the cascadic conjugate–gradient method. – Preprint№ 4, Magdeburg: Otto–von–Guericke–Universität, 1994.
34. Shaidurov V.V. The convergence of the cascadic conjugate–gradient method under a deficient regularity. – In: Problems and Methods in Mathematical Physics. – Stuttgart: Teubner, 1994, P. 185-194.
35. Shaidurov V.V. Cascadic algorithm with nested subspaces in domains with curvilinear boundary. – In: Advanced Mathematics: Computations and Applications. Eds.: A.S. Alekseev and N.S. Bakhvalov. – Novosibirsk: NCC Publisher, 1995, P. 588-595.
36. Shaidurov V.V. Some estimates of the rate of convergence for the cascadic conjugate-gradient method. – Computers Math. Applic., 1996, V. 31, № 4/5, P. 161–171.
37. Shaidurov V.V., Timmermann G. A cascadic algorithm for the solution of nonlinear signindefinite problems. – Technical Report IOKOMO–09–98, Dresden: Technische Universität Dresden, Fakultät für Mathematik und Naturwissenschaften, 1998.
38. Shaidurov V.V., Timmermann G. Stable semi-iterative smoother for cascadic and multigrid algorithms. – In: Proceedings of the sixth European Multigrid Conference: Multigrid methods VI. – Berlin, Heidelberg, N.Y.: Springer Verlag, 2000, P. 221–227.
39. Shaidurov V.V., Tobiska L. The convergence of the cascadic conjugate-gradient method applied to elliptic problems in domains with re-entrant corners. – Preprint№ 37, Magdeburg: Otto-von-Guericke-Universität, 1997.
40. Timmermann G. A cascadic algorithm for the solution of a weakly nonlinear problem. – Technical Report IOKOMO-10-97, Dresden: Technische Universität Dresden, Fakultät für Mathematik und Naturwissenschaften, 1997.
41. Voevodin V.V., Kuznetsov Yu.A. Matrices and Computations (in Russian). – Moscow: Nauka, 1984.