An Analysis of Convergence of the Multigrid Method for Stiff Problems

V. Kucherenko*

Abstract. If on tetrahedrons $\omega_i$ with vertices at 'coarse' nodes the variations of values of coefficients of the elliptic operator are uniformly bounded by quantity $S$, the convergence rate of the two-grid method suggested in the paper is dependent only on the quantity $S$, i.e. independent of the value of a coefficient jump at the boundary of 'coarse' tetrahedrons $\omega_i$.

1. Introduction. Let $\Omega$ be a bounded domain in $\mathbb{R}^3$ with the piecewise-differentiable boundary.

Let then functions $a_{ij}$ be piecewise-continuous and for the quadratic form

$$a(x) = \sum_{i,j=1}^{3} a_{ij}(x) \xi_i \xi_j$$

a piecewise-continuous function $\lambda(x) > 0$ exists such that

$$\inf_{x \in \Omega, |\xi| = 1} \frac{a(\xi x)}{\lambda(x)} > 0; \quad \sup_{x \in \Omega, |\xi| = 1} \frac{a(\xi x)}{\lambda(x)} < \infty. \quad (1.1)$$

Let then $b(x) \geq 0$, $b \in L_1(\partial \Omega)$ on a set of positive measure. Introduce on $W^1_2(\Omega)$ a scalar product

$$[u, v] = \left\{ \sum_{i,j=1}^{3} a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} \right\}_{\Omega} + \int_{\partial \Omega} b u v \, d\Sigma. \quad (1.2)$$

For $f \in L_2(\Omega)$, $\psi \in L_2(\partial \Omega)$ consider the functional minimum problem

$$\Phi(u) = [u, u] - 2 \int_{\Omega} uf \, dx - 2 \int_{\partial \Omega} u \psi \, d\Sigma. \quad (1.3)$$

* Moscow Civil Engineering Institute, Moscow, USSR.
In this paper we will analyse the multigrid iterative method of computation of the minimum of the functional \( \Phi(u) \) on finite element spaces of functions.

Let \( \Omega_0 \) be a partition of the domain \( \Omega \) into tetrahedrons \( \Delta^0 \). Denote by \( \omega_0 \) a set of vertices of tetrahedrons of the partition \( \Omega_0 \). Assume that the nodes of \( \omega_0 \) which are not internal points of the domain \( \Omega \) belong to its boundary \( \partial \Omega \). Then again partition the tetrahedrons \( \Delta^0 \) into finer tetrahedrons \( \Delta^1 \). Denote by \( \omega_1 \) a set of tetrahedrons of fine partition and by \( \omega_1 \) a set of their vertices. By \(|\omega_k|, k=0,1\), denote the number of nodes in \( \omega_k \).

Let \( R^*_k, k=0,1 \), be linear spaces of grid functions on the nodes \( \omega_k \), and let \( x_k^m \) be a node from \( \omega_k \). Denote by \( st(x_k^m) \) the star of the node \( x_k^m \), i.e., the set of tetrahedrons of the partition \( \Omega_k \) which have the node \( x_k^m \) as their vertex. Then introduce the basis function \( \varphi_k^m(x) \) equal to unity at the node \( x_k^m \), to zero at the other nodes of the star \( st(x_k^m) \) and linear on each tetrahedron of the star. Outside \( st(x_k^m) \) the function \( \varphi_k^m \) is assumed to be equal to zero. Then functions \( \varphi_0^m \) are said to be 'coarse-basis', and \( \varphi_1^m \) are said to be 'fine-basis'. Now construct finite element spaces \( H_k, k=0,1 \):

\[
H_k : v = \sum_{x_k^m \in \omega_k} v(x_k^m) \varphi_k^m(x). \tag{1.4}
\]

This formula defines isomorphism \( J_k : R^*_k \rightarrow H_k \). It is obvious that \( H_0 \subset H_1 \). Denote this inclusion by \( J \) and define the inclusion \( i_1 \) of \( R_0 \) into \( R_1 \):

\[
i_1 = J^{-1}H_0. \tag{1.5}
\]

It is not difficult to verify that \( (i_1\mu)(x_k^m) = (J\mu)(x_k^m) \). Obviously, \( i_1 \) extends definition of grid functions from \( R_0 \) given at coarse nodes of \( \omega_0 \) to fine nodes of \( \omega_1 \setminus \omega_0 \) by interpolation defined by the mapping \( J_0 \).

Let us introduce on \( R_k \) the scalar products

\[
(\varphi, \psi) = \int_\Omega (J_k \varphi) \cdot (J_k \psi) \, dx, \quad (\varphi, \psi) \overset{\text{def}}{=} [J_k \varphi, J_k \psi] \tag{1.6}
\]

\[
<\varphi, \psi>_k = \sum_{x_k^m \in \omega_k} \varphi(x_k^m) \psi(x_k^m).
\]

2. Projector \( P \) of space \( R_1 \) onto \( i_1 R_0 \). The finite element approximation of the elliptic operator \( L \) corresponding to problem (1.4) is the linear operator \( L_k \) on the space \( R_k \) satisfying the relation \( <L_k u, \psi>_k = [u, \psi] \forall u, \psi \in R_k \). As well known, minimization problem (1.3) on the space \( H_k \) can be reduced to the solution of the linear equation

\[
L_k u_k = f_k, \quad u_k \in R_k \tag{2.1}
\]

where the right-hand side \( f_k \) can be computed by standard procedures using \( f \) and \( \psi \) [1].
In the multigrid iterative process contained in [2, 3] for solving the equation
\( L_1 u_1 = f_1 \), the fundamental part is played by the projector \( P \) of orthogonal projection with respect to the scalar product \( \langle \cdot, \cdot \rangle \) of the space \( R_1 \) onto \( i_1 R_0 \).

The function \((1 - P)u\) is obviously orthogonal to \( i_1 R_0 \). To compute this function, it is convenient to pass to the space \( H_1 \). Then,

\[ J_1(1 - P)u = J_1 u + \sum_{m=1}^{\omega_0} v_m \phi_0^m \]

and the condition \( \langle J_1(1 - P)u, \phi_0^n \rangle = 0, \ n = 1, \ldots, \omega_0 \rangle \), leads to the system of equations

\[ \sum_{m=1}^{\omega_0} v_m [\phi_0^m, \phi_0^n] = -\langle J_1 u, \phi_0^n \rangle \].

(2.2)

As known, the matrix \( \| [\phi_0^m, \phi_0^n] \| \) coincides with the matrix \( L_{10} \). To compute the function \((1 - P)u\), it is therefore necessary to solve equation (2.2) on the coarse grid \( \omega_0 \).

3. Iterative process. For the case where the estimating function \( \lambda(\chi) \) from condition (1.1) is continuous on tetrahedrons of the coarse partition \( \omega_{1p} \) in [3] we suggested an iterative process whose convergence rate is independent of the value of \( \lambda(\chi) \).

Thus, denote by \( D_1 \) a diagonal matrix equal to the diagonal of the operator \( L_1 \), and let \( u_0 \) be the solution to equation (2.1) on the coarse grid, i.e., for \( k = 0 \). Let us consider the iterative process

\[ Y^n = (1 - P)X^n + i_1 u_0, \quad X^0 = 0, \quad n \geq 0 \]  

\[ X^{n+1} = Y^n - \tau_n D_1^{-1}(L_1 Y^n - f_1), \quad n \geq 0. \]  

(3.1)  

(3.2)

Here, the step size \( \tau_n \) is determined by the gradient descent method for the operator \( D_1^{-1}L_1(1 - P) \):

\[ \tau_n = \langle D_1 \xi_n, \xi_n \rangle^{-1} \langle L_1(1 - P)\xi_n, \xi_n \rangle \]

(3.3)

Thus, to start iterative process (3.1), (3.2), it is first necessary to solve equation (2.1) on the coarse grid and then to successively perform computations by formulae (3.1), (3.2). It was proved in [2] that

\[ Pu_1 = i_1 u_0. \]  

(3.4)

Hence, iterative process (3.1), (3.2) fixes the projection of \( Y^n \) onto \( i_1 R_0 \).

The modifications of iterative process (3.1), (3.2) without justification of its
convergence for stiff problems can be found in publications by various authors (see, for example, [4]). Another approach to the multigrid method in case of the coefficients being continuous on the tetrahedrons \(D_0\) was suggested in [5].

Below, we impose constraints on the function \(\lambda(x)\) under which in the estimate

\[
<L_{L_n+1}^r r_{n+1}>_1 \leq (1 - q) <L_{L_n}^r r_n>_1, \quad r_n = Y^n - u_1
\]

(3.5)

the contraction coefficient \(q\) is independent of the step size of grid nodes and the spread in values of \(\lambda(x)\), i.e. of the 'stiffness' of problem.

To prove it, we will suggest another technique making it possible to avoid the use of the second main inequality for elliptic equations.

Let us analyze the algebraic structure of the formulae of process (3.1), (3.2). Denote by \(P^*\) an operator adjoint to \(P\) with respect to the scalar product \(<,>_1\) and let \(D_1\) be an arbitrary symmetric, positive definite operator in \(R^*_1\). Introduce the linear space \(R_D \overset{def}{=} D_1^{-1}(1 - P^*)R_1\) and set

\[
m(L) = \inf_{\nu \in R_D^*} <L_1(1 - P)\nu,\nu>_1 / <D_1^*\nu,\nu>_1
\]

\[
M(L) = \sup_{\nu \in R_D^*} <L_1(1 - P)\nu,\nu>_1 / <D_1^*\nu,\nu>_1
\]

(3.6)

\[
q(L) = [M(L) - m(L)]/[M(L) + m(L)].
\]

**Lemma 3.1.** Iterative process (3.1), (3.2) for \(n \geq 0\) satisfies estimate (3.5) with \(q = q(L)\).

**Proof.** The space \(R_1\) can be decomposed into the direct sum \(R_1 = R_D \oplus PR_1\) orthogonal to \(<D_1 \cdot,\cdot>_1\). Indeed, assume that \(<D_1 \cdot, (1 - P^*)w, w>_1 = 0\) \(\forall w \in R_1\). It means that \((1 - P)w = 0\), i.e. \(w \in PR_1\). Denote by \(Q\) the projector \(R_1\) onto \(R_D^*\) orthogonal to \(<D_1 \cdot,\cdot>_1\).

Then the equality \(L_1 P = P^*L_1\) proved in [2] implies the equality

\[
D_1^{-1}L_1(1 - P) = D_1^{-1}(1 - P^*)L_1.
\]

(3.7)

Hence, we have \(D_1^{-1}L_1(1 - P)R_1 \subset R_D\) and the operator \(D_1^{-1}L_1(1 - P)\) is self-adjoint to \(<D_1 \cdot,\cdot>_1\).

Formula (3.1) implies the equality

\[
Y^n - u_1 = (1 - P)(X^n - u_1).
\]

(3.8)

Since for the introduced projector \(Q\) we have \((1 - P) = (1 - P)Q\), then by virtue of (3.7) we obtain

\[
<L_1(Y^{n+1} - u_1), (Y^{n+1} - u_1)>_1 = <L_1(1 - P)(X^{n+1} - u_1), (1 - P)(X^{n+1} - u_1)>_1
\]

\[
= <L_1(1 - P)Q(X^{n+1} - u_1), (1 - P)Q(X^{n+1} - u_1)>_1.
\]

(3.9)
Then determine \( Q(X^{n+1} - u_1) \) from (3.2) and substitute it into (3.9). Equality (3.8) implies that \( P \gamma^n = P \nu_1 \). Therefore,

\[
L_1 Y^n - f_1 = L_1 (Y^n - u_1) = L_1 (1 - P)(Y^n - u_1).
\]

Hence, by virtue of (3.2) we have

\[
X^{n+1} - u_1 = Y^n - u_1 - \tau_\eta D_1^{-1} L_1 (Y^n - u_1) = Y^n - u_1 - \tau_\eta D_1^{-1} L_1 (1 - P)(Y^n - u_1).
\]  

(3.10)

Formula (3.7) implies that the second term in the right-hand side of (3.10) belongs to \( R_D \). Make the operator \( Q \) act on equality (3.10). Then taking into account that \( 1 - P = (1 - P)Q \) we obtain

\[
Q(X^{n+1} - u_1) = Q(Y^n - u_1) - \tau_\eta D_1^{-1} L_1 (1 - P)Q(Y^n - u_1).
\]  

(3.11)

The right-hand side in (3.11) defines the iterative process on \( R_D \) for the self-adjoint operator \( D_1^{-1} L_1 (1 - P) \) whose step \( \tau_\eta \) in the gradient descent method is computed by formula (3.3). Substitute (3.11) into (3.9). Taking into account the convergence rate estimates for the steepest descent method [6] and the estimate

\[
<L_1 (1 - P)\nu, (1 - P)\nu> = [(1 - P)\nu, \nu] \leq \nu, \nu = <L_1 \nu, \nu>.
\]

we obtain

\[
<L_1 (1 - P)Q(X^{n+1} - u_1), Q(X^{n+1} - u_1)> \leq (1 - q(L))<L_1 (1 - P)Q(Y^n - u_1), (1 - P)Q(Y^n - u_1)> \leq (1 - q(L))<L_1 (1 - P)(Y^n - u_1), (1 - P)(Y^n - u_1)>.
\]  

(3.12)

Equality (3.8) implies that \( Y^n - u_1 = (1 - P)(Y^n - u_1) \). Hence, (3.9), (3.12) imply estimate (3.5) for \( q = q(L) \). This completes the proof of the lemma.

4. Estimates for \( m(L), M(L) \). Let us introduce the value \( \delta_\lambda^i \):

\[
\delta_\lambda^i = \max\left( \frac{\text{Vrai max } \lambda}{\text{Vrai min } \lambda} \right)
\]

(4.1)

and similarly define \( \delta_\beta^i \) [the function \( b(x) \) is contained in (1.2)]. Let it be known that in the partitions \( D_\kappa \) on each of the tetrahedrons \( \Delta_m^k \) the ratio of the length of the maximal edge to that of the minimal one does not exceed \( \gamma \), and each of the tetrahedrons \( \Delta_m^0 \) contains, at most, \( \nu \) fine tetrahedrons \( \Delta_n^1 \). The following theorem is valid.
Theorem 4.1. A constant $C$ dependent on $\gamma$, $\nu$ and the boundary $\partial \Omega$ exists such that

$$m(L) \geq \frac{C}{\max(\delta_A, \delta_B)^2}. \quad (4.2)$$

For arbitrary integrated values $a_j$ and $b$ we have

$$M(L) \leq 4. \quad (4.3)$$

Auxiliary manipulations. Introduce the notation

$$[u, \nu](\Delta) \overset{\text{def}}{=} \frac{\partial u}{\partial x_j} \frac{\partial \nu}{\partial \delta_j} \int_{\Omega \cap A} \sum_{i, j=1}^3 a_{i j} \frac{\partial u}{\partial x_i} \frac{\partial \nu}{\partial x_j} \, dx + \iota\nu \, d\Sigma. \quad (4.4)$$

Let $\omega(A_m^i)$ be a set of numbers of vertices of the tetrahedron $A_m^i$. Set

$$<D_1 u, \nu>_1(A_m^1) = \sum_{n \in \omega(A_m^1)} [\varphi_n^r, \varphi_n^s](\Delta_m^1) u_n v_n \quad (4.5)$$

$$<D_1 u, \nu>_1(A_m^0) = \sum_{A_m^0 \in A_k^0} <D_1 u, \nu>_1(A_k^1).$$

Then for each function $\nu \in R_1$ construct a coarse-basis function $\varphi$ coinciding with $\nu$ at coarse nodes of $\omega_0$:

$$\varphi = \Pi_0 \overset{\text{def}}{=} \sum_{m=1}^{\omega_0} \nu(x_0^m) \varphi_0^m(x). \quad (4.6)$$

It is obvious that $\Pi^2 = \Pi$. Denote the linear space $(1 - \Pi)R_1$ by $R_1/R_0$. To estimate $m(L)$ from below, we will make use of the values

$$C_{10} \overset{\text{def}}{=} \sup_{v \in R_1} \frac{[\Pi v, \Pi v]/[v, v]}{v} \quad (4.7)$$

$$m_{10} \overset{\text{def}}{=} \inf_{A_m^0} \inf_{\psi \in R_1/R_0} \frac{[\psi, \psi](A_m^0)/<D_1 \psi, \psi>_1(A_m^0).} {<D_1 \nu, \nu>_1(A_m^0)}. \quad (4.8)$$

The proof of the estimate from below.

Lemma 4.1. The following estimate is valid:

$$m(L) \geq m_{10}/(1 + C_{10}). \quad (4.9)$$

Proof. Set $\varphi = \Pi(u - Pu)$. By virtue of orthogonality of the subspaces $R_1$ and $PR_1$ to $<D_1 \cdot, \cdot>$ we have

$$<D_1 u, u>_1 \leq <D_1(u - Pu - \varphi), (u - Pu - \varphi)>_1 \leq$$

$$= <D_1 u, u>_1 + <D_1(Pu + \varphi), (Pu + \varphi)>_1. \quad (4.10)$$
Then (4.7) implies the inequality
\[ <L_1(u - Pu - \phi), (u - Pu - \phi)>_1 \leq 2(1 + C_{10}) <L_1(1 - P)u, u>_1. \] (4.11)

Inequalities (4.10) and (4.11) on \( R_D \) imply the estimate
\[ \frac{<L_1(u - Pu - \phi), (u - Pu - \phi)>_1}{<D_1(u - Pu - \phi), (u - Pu - \phi)>_1} \leq 2(1 + C_{10}) \frac{<L_1(1 - P)u, u>_1}{<D_1u, u>_1}. \] (4.12)

Set \( \psi = u - Pu - \phi \). It is obvious that \( \psi \) belongs to \( R_D \).

Below we will show that
\[ <D_1 \psi, \psi>_1 = \sum_{k \in \Omega_0} [\psi, \psi](A^0_k). \] (4.13)

Therefore, the left-hand side of (4.12) can be presented in the form of fraction
\[ \left[ \sum_{A^0_m \in \Omega_0} [\psi, \psi](A^0_m) \right] / \left[ \sum_{A^0_m \in \Omega_0} <D_1 \psi, \psi>_1(A^0_m) \right]. \] (4.14)

Now make use of the elementary inequality for positive values \( a_i \) and \( b_i \):
\[ \min_i (a_i/b_i) \leq (\sum a_i)/(\sum b_i) \leq \max_i (a_i/b_i). \] (4.15)

Hence, by virtue of (4.15) and the definition of \( m_{10} \) from (4.12), (4.14) we obtain estimate (4.9).

The proof of relation (4.13). Let \( st(x^n_1) \) be stars of nodes \( x^n_1, n = 1, \ldots, |\omega_1| \). Obviously, we have
\[ <D_1 \psi, \psi>_1 = \sum_{n=1}^{\omega_1} [\phi^n_1, \phi^n_1](x^n_1)^2 = \sum_{n=1}^{\omega_1} v_n^2 \sum_{A^1_m \in \omega(\phi^n_1)} [\phi^n_1, \phi^n_1](A^1_m). \] (4.16)

Rearrange the terms in sum (4.16). Note that the tetrahedron \( A^1_m \) is contained in those stars \( st(x^n_1) \) whose nodes \( x^n_1 \) are its vertices: \( x^n_1 \in \omega(A^1_m) \). Hence, if we group all summands relating to simplexes \( A^1_m, m = 1, \ldots, |\omega_1| \), we obtain formula (4.13). This completes the proof of the lemma.

The new technique suggested for estimating \( m(L) \) consists in exploiting values (4.7) and (4.8) and in estimating them via \( \delta_n, \delta_p, \nu \) and \( \gamma \). In this way we manage to avoid using the second main inequality for the elliptic equation.

The proof of Theorem 4.1. Estimate from above (4.3). The following string of equalities is valid
\[ <L_1 \nu, \nu>_1 = [v^1_1, ]^1_1 v^1_1 = \sum_{n=1}^{\omega_1} [v^1_1, ]^1_1 (A^1_m) = \sum_{A^1_m \in \Omega_1} \sum_{l,n \in \omega(\phi^n_1)} [\phi^n_1, \phi^n_1](A^1_m) v_l v_n. \] (4.17)
Then exploiting the estimate
\[ |[\varphi^1_1, \varphi^m_1](A^1_m)\nu'_n| \leq 2^{-1}[\varphi^1_1, \varphi^m_1](A^1_m)\nu'_n^2 + 2^{-1}[\varphi^m_1, \varphi^m_1](A^1_m)\nu'_n^2 \]
by virtue of (4.17) we find
\[ <L_1 v, v>_1 \leq 4 \sum_{A^1_m \in \Omega} <D_1 v, v>_1(A^1_m). \]  
(4.18)

Inequality (4.3) is implied by relations (4.13), (4.5) and inequality (4.18).

*Estimate from below.* Estimate first $C_{10}$. Show that if
\[ [v, v](A^0_m) = 0, \text{ then } [\Pi v, \Pi v](A^0_m) = 0. \]  
(4.19)

Formula (1.2) implies that if $[v, v](A^0_m) = 0$, then almost everywhere on $\partial \Omega \cap A^0_m$ we have $b(x) = 0$. Then the expression $[v, v](A^0_m)$ is a quadratic form of values of the function $v$ at nodes of $\bigcup_{A^1_n \in A^0_m} \omega(A^1_n)$, and the expression $[\Pi v, \Pi v](A^0_m)$ is a quadratic form of values of the function $v$ at nodes of $\omega(A^0_n)$. The form $[\Pi v, \Pi v](A^0_n)$ vanishes in the linear subspace
\[ v(x^0_0) = \ldots = v(x^0_k), \quad \{x^0_0, \ldots, x^0_k\} = \omega(A^0_m) \]  
(4.20)

and the form $[v, v](A^0_m)$ vanishes on the linear subspace
\[ v(x^1_0) = \ldots = v(x^1_k), \quad \{x^1_0, \ldots, x^1_k\} = \bigcup_{A^1_n \in A^0_m} \omega(A^1_n). \]  
(4.21)

Since $\omega(A^0_m) \subset \bigcup_{A^1_n \in A^0_m} \omega(A^1_n)$, null space (4.21) is contained in null space (4.20) and, hence, relation (4.19) is valid. Thus,
\[ \sup_{v \in H} [\Pi v, \Pi v]/[v, v] \leq \max' \sup_m [\Pi v, \Pi v]/[v, v](A^0_m). \]  
(4.22)

Here, max' is taken for those $m$ for which we have $[v, v](A^0_m) \neq 0$. For such $m$ we have
\[ [\Pi v, \Pi v]/[v, v] \leq \max_{m} (\delta_{A^0_m} \delta_{\partial \Omega}) \max_{m} \frac{\int \nabla(J_0\Pi v)^2 \, dx + \int_{A^0_m \cap \partial \Omega} (J_0\Pi v)^2 \, d\Sigma}{\int_{A^0_m \cap \partial \Omega} \nabla(J_0\Pi v)^2 \, d\Sigma}. \]  
(4.23)

The quadratic forms
\[ \int_{A^0_m} |\nabla(J_0\Pi v)|^2 \, dx, \quad \int_{A^0_m} |\nabla(J_0 v)|^2 \, dx \]
have null spaces (4.20) and (4.21), respectively. Consider their relation on the
subspace orthogonal to null space (4.21) and using standard arguments we obtain

\[ \int_{\Delta_m^0} |\nabla (I_0 I^0 v)|^2 \, dx \leq C_1 (\gamma, v, \partial \Omega) \int_{\Delta_m^0} |\nabla (I_0 v)|^2 \, dx. \]  
(4.24)

Likewise, we find

\[ \int_{\Delta_m^0 \cap \partial \Omega} (I_0 I^0 v)^2 \, d\Sigma \leq C_1 (\gamma, v, \partial \Omega) \int_{\Delta_m^0 \cap \partial \Omega} (I_0 v)^2 \, d\Sigma. \]  
(4.25)

Estimates (4.24) and (4.25) imply the estimate for \( C_{10}. \)

By virtue of (4.20) the estimate for \( m_{10} \) is obtained similarly to that for \( C_{10}. \) This completes the proof of the theorem.

**Application to computing real problems.** If on the tetrahedrons of the coarse partition \( \Omega_0 \) the variations of values of the functions \( \lambda \) and \( b \) are great, it is necessary to modify the coarse partition \( \Omega_0 \) by adding to it, instead of some coarse tetrahedrons with large values \( \delta_\lambda \) and \( \delta_b \), tetrahedrons of fine partition which make part of coarse tetrahedrons. For new ‘coarse nodes’ \( x_0^{nf} \) which are vertices of the newly included (formerly fine) tetrahedrons we take old fine-basis functions for new ‘coarse-basis’ functions. The partition \( \Omega_1 \) and the basis functions corresponding to it are taken without any modifications.

**REFERENCES**


