

Easily implemented iterative solution methods for a class of variational inequalities

Alexander Lapin (Kazan Federal University, Kazan, Russia)

1. Variational inequality with a nonlinear diffusion-convection differential operator

$$\int_{\Omega} a(x)g_1(\nabla u) \cdot \nabla(v - u) dx + \int_{\Omega} b(x)g_2(u, \nabla u)(v - u) dx + \phi(|\nabla v|) - \phi(|\nabla u|) \geq \int_{\Omega} f(v - u) dx \quad \forall v \in V, u \in V. \quad (1)$$

Here $\Omega \subset \mathbb{R}^2$ is a bounded domain with a piecewise smooth boundary $\partial\Omega$, $H_0^1(\Omega) \subseteq V \subseteq H^1(\Omega)$, $H^1(\Omega)$ and $H_0^1(\Omega)$ are Sobolev spaces; $f \in L_2(\Omega)$, $a, b \in L_{\infty}(\Omega)$ and $a(x) \geq a_0 > 0 \forall x \in \Omega$.

Let the operator $P : V \rightarrow V^*$ be defined by the left-hand side of (1):

$$\langle Pu, v \rangle = \int_{\Omega} a(x)g_1(\nabla u) \cdot \nabla v dx + \int_{\Omega} b(x)g_2(u, \nabla u)v dx.$$

The main assumptions:

P is uniformly monotone.

ϕ is a proper, convex and lower semicontinuous function.

The sufficient conditions to ensure continuity and uniform monotonicity of P are:

$$\begin{cases} g_1(\bar{t}) \text{ and } g_2(s, \bar{t}) \text{ are continuous and } |g_1(\bar{t})| \leq c|\bar{t}|, \\ (g_1(\bar{t}_1) - g_1(\bar{t}_2), \bar{t}_1 - \bar{t}_2) \geq \sigma_0|\bar{t}_1 - \bar{t}_2|^2, \quad \sigma_0 > 0, \\ |g_2(s_1, \bar{t}_1) - g_2(s_2, \bar{t}_2)| \leq \beta_1|s_1 - s_2| + \beta_2|\bar{t}_1 - \bar{t}_2|, \end{cases} \quad (2)$$

$$a_0\sigma_0 - b\beta_1c_f^2 - b\beta_2c_f \equiv \sigma > 0,$$

where $b = \sup_{x \in \Omega} |b(x)|$, c_f is the constant from Friedrichs inequality.

Classical examples of function ϕ :

$$1) \phi(|\nabla u|) = \int_{\Omega} |\nabla u| dx,$$

2) $\phi(|\nabla u|)$ is the indicator function of the convex and closed set $K = \{u \in H_0^1(\Omega) : |\nabla u(x)| \leq 1 \forall x \in \Omega\}$.

Under aforementioned assumptions for the operator P and the functional ϕ variational inequality (1) has a unique solution.

Further for the definiteness we consider variational inequality (1) in the case $V = H_0^1(\Omega)$ and $\phi(|\nabla u|)$ is the indicator function of K .

Remark 1 *The result on the existence of a unique solution for variational inequality (1) remains valid if the operator P is continuous and uniformly monotone only on the set $K = \{u \in H_0^1(\Omega) : |\nabla u(x)| \leq 1 \forall x \in \Omega\}$. Because of this, the assumptions (2) can be satisfied only for $\bar{t} \in \mathbb{R}^2 : |\bar{t}| \leq 1$.*

2. Approximation of the variational inequality

1-st order f.e.m. on the triangle grids in the case of a polygonal Ω .

$T_h = \{e\}$, $\cup e = \bar{\Omega}$, is a conforming triangulation of $\bar{\Omega}$ into triangle finite elements e ,

$V_h \subset H_0^1(\Omega)$ is the space of the continuous and piecewise linear functions,

$U_h \in L_2(\Omega)$ is the space of the piecewise constant functions,

$K_h = \{u_h \in V_h : |\nabla u_h| \leq 1 \ \forall x \in \Omega\}$ is the convex and closed set in V_h ,

$f_h = (\text{meas } e)^{-1} \int_{t \in e} f(t) dt$ and similar for a_h and b_h .

Discrete variational inequality, approximating (1):

$$\begin{aligned} \int_{\Omega} a_h g_1(\nabla u_h) \cdot \nabla(v_h - u_h) dx + \int_{\Omega} b_h g_2(u_h, \nabla u_h)(v_h - u_h) dx &\geq \\ &\geq \int_{\Omega} f_h(v_h - u_h) dx \quad \forall v_h \in K_h, u_h \in K_h. \quad (3) \end{aligned}$$

The operator defined by the left-hand side of this variational inequality inherits the properties of P , so, (3) has a unique solution.

Matrix (operator)-vector form of the discrete variational inequality.

$w \in \mathbb{R}^{N_e}$ and $u \in \mathbb{R}^{N_u}$ are the vectors of nodal values of the functions $w_h \in U_h$ and $u_h \in V_h$, respectively:

$$U_h \ni u_h \Leftrightarrow w \in \mathbb{R}^{N_e} \text{ and } V_h \ni u_h \Leftrightarrow u \in \mathbb{R}^{N_u}.$$

If $\bar{q}_h = (q_{1h}, q_{2h}) \in U_h \times U_h$ then $q_h \Leftrightarrow q \in \mathbb{R}^{N_y}$, $N_y = 2N_e$, $q = (q_{11}, q_{21}, \dots, q_{1i}, q_{2i}, \dots, q_{1N_e}, q_{2N_e})$.

(Further by (\cdot, \cdot) and $\|\cdot\|$ we mean the Euclidian scalar products and norms in the corresponding spaces.)

Define the matrices $L \in \mathbb{R}^{N_u \times N_y}$, $M_p \in \mathbb{R}^{N_y \times N_y}$, $M_u \in \mathbb{R}^{N_u \times N_u}$ and the operators:

$$(Lu, q) = \int_{\Omega} \nabla u_h(x) \cdot \bar{q}_h(x) dx, \quad (M_p p, q) = \int_{\Omega} \bar{p}_h(x) \cdot \bar{q}_h(x) dx,$$

$$(M_u u, v) = \int_{\Omega} u_h(x) v_h(x) dx,$$

$$k_1 : \mathbb{R}^{N_y} \rightarrow \mathbb{R}^{N_y}, \quad (k_1(p), q) = \int_{\Omega} a_h(x) g_1(\bar{p}_h(x)) \cdot \bar{q}_h(x) dx,$$

$$k_2 : \mathbb{R}^{N_u} \times \mathbb{R}^{N_y} \rightarrow \mathbb{R}^{N_u}, \quad (k_2(u, p), v) = \int_{\Omega} b_h(x) g_2(u_h, \bar{p}_h(x)) v_h(x) dx.$$

Finally, denote by $\theta : \mathbb{R}^{N_y} \rightarrow \overline{\mathbb{R}}$ the indicator function of the set $\mathcal{K} = \{p : p_{2j}^2 + p_{2j-1}^2 \leq 1 \ \forall j = 1, \dots, N_e\}$. Using the notations variational inequality (3) can be written as follows:

$$(L^T k_1(Lu) + k_2(u, Lu), (v - u)) + \theta(Lv) - \theta(Lu) \geq (f, v - u). \quad (4)$$

The equivalent form of writing for discrete variational inequality is the inclusion

$$L^T k_1(Lu) + k_2(u, Lu) + L^T \partial\theta(Lu) \ni f, \quad (5)$$

which we will solve.

Properties of the matrices and operators:

Matrices $L^T L$, M_u and M_p are symmetric and positive definite, M_p has **block diagonal form with 2×2 blocks** (N_e blocks corresponding to the finite elements).

The operator k_1 is continuous and uniformly monotone, while k_2 is Lipschitz-continuous:

$$\begin{aligned} (k_1(p) - k_1(q), p - q) &\geq a_0 \sigma_0 \|p - q\|_{M_p}^2, \\ \|k_2(u, p) - k_2(v, q)\|_{M_u^{-1}} &\leq b\beta_1 \|u - v\|_{M_u} + b\beta_2 \|p - q\|_{M_p}. \end{aligned}$$

k_1 has the **block diagonal form** with 2×2 blocks.

$\partial\theta$ is a maximal monotone operator – subdifferential of the proper, convex and lower semicontinuous function θ – and it has the **block diagonal form** with 2×2 blocks as k_1 and M_p .

Construction of the saddle point problem

Consider inclusion (5):

$$L^T k_1(Lu) + k_2(u, Lu) + L^T \partial\theta(Lu) \ni f.$$

Define the auxiliary vectors $p = M_p^{-1/2}Lu$ and $\lambda \in k_1(M_p^{1/2}p) + \partial\theta(M_p^{1/2}p)$. Then the triple (u, p, λ) satisfies the following system:

$$\begin{aligned} k_2(u, M_p^{1/2}p) &+ L^T \lambda \ni 0, \\ k_1(M_p^{1/2}p) + \partial\theta(M_p^{1/2}p) &- M_p^{1/2} \lambda \ni 0, \\ Lu - M_p^{1/2}p &= 0. \end{aligned} \quad (6)$$

The operator $A_0 \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} k_2(u, p) \\ k_1(p) \end{pmatrix}$ is not monotone and this impedes the application of the iterative solution methods.

We make the equivalent transformations of the system by using the equation $M_p^{1/2}p - Lu = 0$ to get a monotone operator:

$$\begin{aligned} rL^T Lu - rL^T M_p^{1/2}p + k_2(u, M_p^{1/2}p) + L^T \lambda &= 0, \\ k_1(M_p^{1/2}p) + \partial\theta(M_p^{1/2}p) - M_p^{1/2} \lambda &\ni 0, \\ Lu - M_p^{1/2}p &= 0. \end{aligned} \quad (7)$$

Lemma 1 *If*

$$0 \leq r_1 < r < r_2, \quad r_{1,2} = 2a_0\sigma_0 - b\beta_2c_f \mp 2\sqrt{a_0\sigma_0\sigma}$$

then

the operator $A \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} rL^T Lu - rL^T p + k_2(u, p) \\ k_1(p) \end{pmatrix}$ is uniformly monotone, problem (7) has a solution (u, p, λ) with the unique component (u, p) .

4. Iterative solution method for saddle point problem

(7):

$$\begin{aligned} k_1(M_p^{1/2}p^{k+1}) + \partial\theta(M_p^{1/2}p^{k+1}) - M_p^{1/2}\lambda^k &\ni 0, \\ rL^T Lu^{k+1} - rL^T p^{k+1} + k_2(u^k, M_p^{1/2}p^{k+1}) + L^T\lambda^k &= 0, \\ \lambda^{k+1} &= \lambda^k + \tau(M_p^{1/2}p^{k+1} - Lu^{k+1}) \end{aligned} \quad (8)$$

with an initial guess (λ^0, u^0) .

Further for the definiteness we take $r = 2a_0\sigma_0 - b\beta_2c_f$ – the midpoint of the admissible interval for r .

Implementation of the method:

1) solve the inclusion

$$k_1(M_p^{1/2}p^{k+1}) + \partial\theta(M_p^{1/2}p^{k+1}) \ni F^k = M_p^{1/2}\lambda^k;$$

2) solve the system of **linear** equations

$$rL^T Lu^{k+1} = rL^T p^{k+1} - k_2(u^k, M_p^{1/2}p^{k+1}) - L^T\lambda^k;$$

3) update λ : $\lambda^{k+1} = \lambda^k + \tau(M_p^{1/2}p^{k+1} - Lu^{k+1})$.

Owing to block diagonal form of the operators k_1 and $\partial\theta$ the inclusion is splitted into N_e two-dimensional problems for the coordinates of vector p^{k+1} corresponding to the finite elements.

Thus, the method is very easy to implement.

Some other (well-known) iterative methods.

Uzawa-type method for solving saddle point problem (7):

$$\begin{aligned} k_1(M_p^{1/2}p^{k+1}) + \partial\theta(M_p^{1/2}p^{k+1}) - M_p^{1/2}\lambda^k &\ni 0, \\ rL^T Lu^{k+1} - rL^T p^{k+1} + k_2(u^{k+1}, M_p^{1/2}p^{k+1}) + L^T\lambda^k &= 0, \\ \lambda^{k+1} &= \lambda^k + \tau(M_p^{1/2}p^{k+1} - Lu^{k+1}) \end{aligned} \quad (9)$$

([A. Lapin, 2010] and [E. Laitinen, A. Lapin and S. Lapin, 2012])

We can also use another transformation of the system (6) (similar to augmented Lagrangian technique; see [M. Fortin and R. Glowinski *Augmented Lagrangian methods* – 1983] and [R. Glowinski and P. LeTallec *Augmented Lagrangian and operator-splitting methods in nonlinear mechanics* – 1989]) and obtain the following saddle point problem:

$$\begin{aligned} rL^T Lu - rL^T M_p^{1/2}p + k_2(u, M_p^{1/2}p) + L^T\lambda &= 0, \\ -rLu + rM_p^{1/2}p + k_1(M_p^{1/2}p) + \partial\theta(M_p^{1/2}p) - M_p^{1/2}\lambda &\ni 0, \\ Lu - M_p^{1/2}p &= 0. \end{aligned}$$

For any $r > 0$ this problem has a solution (u, p, λ) with the unique component (u, p) . Iterative method for its solving (Algorithm 2 due to the terminology of R. Glowinski):

$$\begin{aligned} -rLu^k + rM_p^{1/2}p^{k+1} + k_1(M_p^{1/2}p^{k+1}) + \partial\theta(M_p^{1/2}p^{k+1}) - M_p^{1/2}\lambda^k &\ni 0, \\ rL^T Lu^{k+1} - rL^T p^{k+1} + k_2(u^{k+1}, M_p^{1/2}p^{k+1}) + L^T\lambda^k &= 0, \\ \lambda^{k+1} &= \lambda^k + \tau(M_p^{1/2}p^{k+1} - Lu^{k+1}) \end{aligned} \quad (10)$$

with an initial guess (λ^0, u^0) ;

Implementation: on every iteration of methods (9) and (10) we solve the inclusion as in method (8) and the system of **nonlinear** equations

$$rL^T Lu^{k+1} + k_2(u^{k+1}, M_p^{1/2}p^{k+1}) = -L^T\lambda^k + rL^T p^{k+1}.$$

This is the most time consuming step in the implementation of these methods.

Convergence of the iterative methods follows from the following general result on the convergence of the iterative method for constrained saddle point problem.

5. Iterative solution methods for the constrained saddle point problem

$$\begin{pmatrix} A & -B^T \\ -B & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} + \begin{pmatrix} \partial\psi(x) \\ 0 \end{pmatrix} \ni \begin{pmatrix} f \\ -g \end{pmatrix}. \quad (11)$$

Assumptions:

$$\begin{aligned} & \text{operator } A : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_x} \text{ is continuous, strictly monotone and coercive,} \\ & B \in \mathbb{R}^{N_\lambda \times N_x} \text{ is a full column rank matrix: } \text{rank } B = N_\lambda \leq N_x, \\ & \psi : \mathbb{R}^{N_x} \rightarrow \bar{\mathbb{R}} \text{ is a proper, convex and lower semi-continuous function,} \\ & \text{int dom } \psi \cap \{x \in \mathbb{R}^{N_x} : Bx = g\} \neq \emptyset. \end{aligned} \quad (12)$$

Further we suppose that the following representation takes place:

$Ax = A(x, x)$, where $A(x, y) : \mathbb{R}^{N_x} \times \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_x}$ is a continuous operator.

A particular case of this representation is $A = A_1 + A_2$ with the continuous operators $A_i : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_x}$.

Iterative method for solving system (11):

$$\begin{aligned} & A(x^{k+1}, x^k) + \partial\psi(x^{k+1}) - B^T \lambda^k \ni f, \\ & \frac{1}{\tau} D(\lambda^{k+1} - \lambda^k) + Bx^{k+1} = g, \quad D = D^T > 0. \end{aligned} \quad (13)$$

Theorem 1 *Let assumptions (12) be fulfilled, then saddle point problem (11) has a solution (x, λ) with a unique component x .*

If, in addition, there exist a number $\alpha > 1$ and a non-negative and continuous function $\rho(t) : \mathbb{R} \rightarrow \mathbb{R}$, $\rho(0) = 0$, such that

$$\begin{aligned} (A(x_1, y_1) - A(x_2, y_2), x_1 - x_2) & \geq \frac{\alpha\tau}{2} (D^{-1}B(x_1 - x_2), B(x_1 - x_2)) + \\ & + \rho(x_1 - x_2) - \rho(y_1 - y_2) \quad \forall x_i, y_i \in \mathbb{R}^n, \end{aligned} \quad (14)$$

then iterative method (13) converges starting from any initial guess (x^0, λ^0) .

Theorem 2 *Iterative method (8) converges if $\tau < r = 2a_0\sigma_0 - b\beta_2c_f$.*

For the proof we use theorem 1 with

$$B = \begin{pmatrix} L & -M_p^{1/2} \end{pmatrix}, D = E, \text{ and } \psi(x) = \theta(p)$$

and

$$A(x, y) = \begin{pmatrix} rL^T Lu - rL^T p + k_2(v, p) \\ k_1(p) \end{pmatrix} \text{ for } x = \begin{pmatrix} u \\ p \end{pmatrix}, y = \begin{pmatrix} v \\ q \end{pmatrix},$$

$$B = \begin{pmatrix} L & -M_p^{1/2} \end{pmatrix}, D = E, \text{ and } \psi(x) = \theta(p).$$

Inequality (14):

$$\begin{aligned} (A(x_1, y_1) - A(x_2, y_2), x_1 - x_2) &\geq \frac{\alpha\tau}{2} (D^{-1}B(x_1 - x_2), B(x_1 - x_2)) + \\ &\quad + \rho(x_1 - x_2) - \rho(y_1 - y_2) \quad \forall x_i, y_i \in \mathbb{R}^n, \end{aligned}$$

is satisfied with $\rho(x) = \beta_1 c_f^2 / 2 \|Lu\|^2$.

Remark 2 *The convergence of methods (9) and (10) can be proved by using theorem 1 as well.*

Two-stage iterative method

Let now, when implementing method (8), we solve equation

$$rL^T Lu^{k+1} = rL^T M_p^{1/2} p^{k+1} - k_2(u^k, M_p^{1/2} p^{k+1}) - L^T \lambda^k \equiv F$$

for u^{k+1} by an "inner" iterative method with initial guess u^k . Denote by u_m the m -th iteration of this method, then

$$u_m - u^{k+1} = T_m(u^k - u^{k+1}) \Rightarrow u^{k+1} = (E - T_m)^{-1}(u_m - T_m u^k),$$

where T_m is the corresponding matrix of this method. Whence, u_m satisfies the equation

$$rL^T L(E - T_m)^{-1}(u_m - T_m u^k) = F.$$

If we take u_m as a new, $k + 1$ -th, iteration of method (8), then it becomes

$$\begin{aligned}
& rL^T L(E - T_m)^{-1}u^{k+1} - rL^T L(E - T_m)^{-1}T_m u^k - rL^T M_p^{1/2}p^{k+1} + \\
& \quad + k_2(u^k, M_p^{1/2}p^{k+1}) + L^T \lambda^k = f, \\
& \quad k_1(M_p^{1/2}p^{k+1}) + \partial\theta(M_p^{1/2}p^{k+1}) - M_p^{1/2}\lambda^k \ni 0, \\
& \quad \lambda^{k+1} = \lambda^k + \tau(M_p^{1/2}p^{k+1} - Lu^{k+1}).
\end{aligned} \tag{15}$$

with initial guess (λ^0, u^0) .

Theorem 3 *Iterative method (15) converges if*

$$\|L(E - T_m)^{-1}T_m u\|^2 \leq \gamma \|Lu\|^2, \tag{16}$$

$$\tau < (1 - 4\gamma^{1/2})r, \tag{17}$$

where $\gamma > 0$ is small enough.

We use theorem 1 with $A(x, y) = A_1(x, y) + A_2(x, y)$, where for

$$\begin{pmatrix} u \\ M_p^{1/2}p \end{pmatrix}, y = \begin{pmatrix} v \\ M_p^{1/2}q \end{pmatrix}$$

$$A_1(x, y) = \begin{pmatrix} -rL^T M_p^{1/2}p + k_2(v, M_p^{1/2}p) \\ k_1(M_p^{1/2}p) \end{pmatrix}$$

$$\text{and } A_2(x, y) = \begin{pmatrix} rL^T L(E - T_m)^{-1}u - rL^T L(E - T_m)^{-1}T_m v \\ 0 \end{pmatrix}.$$

The inequality (14) of theorem 1 is valid with

$$\rho(x) = b\beta_1 c_f^2 / 2 \|Lu\|^2 + \frac{r}{2\gamma^{1/2}} \|L(E - T_m)^{-1}T_m u\|^2.$$

Remark 3 *If T_m commutes with the matrix $A_0 = L^T L$, then assumption (16) is fulfilled when*

$$\|T_m\| \leq \frac{\gamma}{1 + \gamma}. \tag{18}$$

This is the situation e.g. of the conjugate gradient method.

Another variational inequality

Let $V = H_0^1(\Omega)$, $K = \{u : |\nabla u| \leq 1\}$ and differential operator $P : V \rightarrow V^*$ be defined by :

$$\langle Pu, v \rangle = \int_{\Omega} a(u)g_1(\nabla u) \cdot \nabla v \, dx.$$

We suppose that $g(s, \bar{t})$ satisfies the monotonicity property:

$$(g(s, \bar{t}_1) - g(s, \bar{t}_2), \bar{t}_1 - \bar{t}_2) \geq \sigma_0 |\bar{t}_1 - \bar{t}_2|^2, \quad \sigma_0 > 0.$$

After approximation of the corresponding variational inequality we get the inclusion

$$L^T k(u, Lu) + L^T \partial\theta(Lu) \ni f, \quad k(u, Lu) = a(u) g_1(Lu)$$

Saddle point problem

$$\begin{aligned} rL^T Lu - rL^T M_p^{1/2}p + L^T \lambda &= 0, \\ k(u, M_p^{1/2}p) + \partial\theta(M_p^{1/2}p) - M_p^{1/2}\lambda &\ni 0, \\ Lu - M_p^{1/2}p &= 0 \end{aligned}$$

has a solution if $0 < r < 4\sigma_0$.

Iterative method (with an initial guess (p^0, λ^0))

$$\begin{aligned} rL^T Lu^{k+1} - rL^T p^k + L^T \lambda^k &= 0, \\ a(u^{k+1}) g_1(M_p^{1/2}p^{k+1}) + \partial\theta(M_p^{1/2}p^{k+1}) - M_p^{1/2}\lambda^k &\ni 0, \\ \lambda^{k+1} &= \lambda^k + \tau(M_p^{1/2}p^{k+1} - Lu^{k+1}) \end{aligned}$$

is easily implementable and converges for $\tau < \tau(r)$.

Numerical experiments

We solved variational inequality (1) with $a(x) = b(x) = 1$, $f(x) = 10$ and

$$g_1(\nabla u) = \begin{cases} \nabla u & \text{if } |\nabla u| < 1/2 \\ \frac{\nabla u}{\sqrt{2|\nabla u|}} & \text{if } |\nabla u| \geq 1/2 \end{cases}, \quad g_w(u, \nabla u) = k \sin u \frac{\partial u}{\partial x_1}$$

In $\Omega = (0, 1) \times (0, 1)$ the finite difference approximation on the uniform grid with steps from 0.01 to 0.002 was used.

We controlled the L_2 -norm of the residual $M_p^{1/2} p^{k+1} - Lu^{k+1}$.

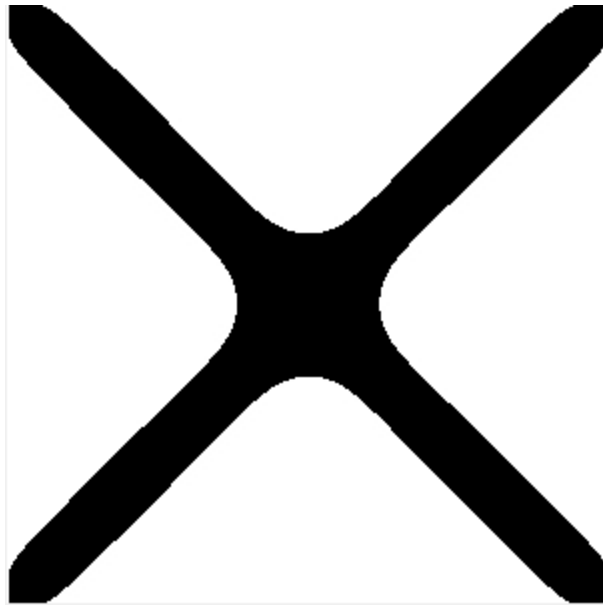


Рис. 1: Black:k=1

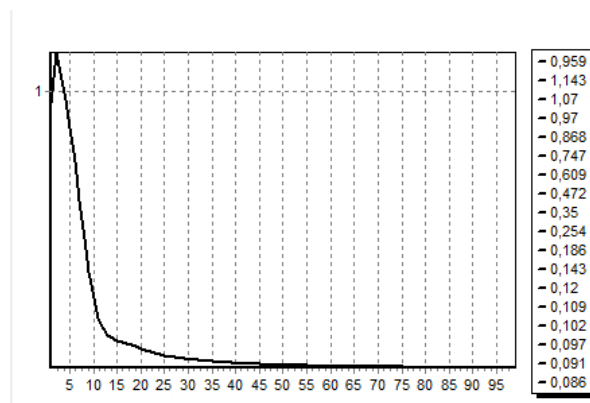


Рис. 2: residual for convection coefficient $k=1$

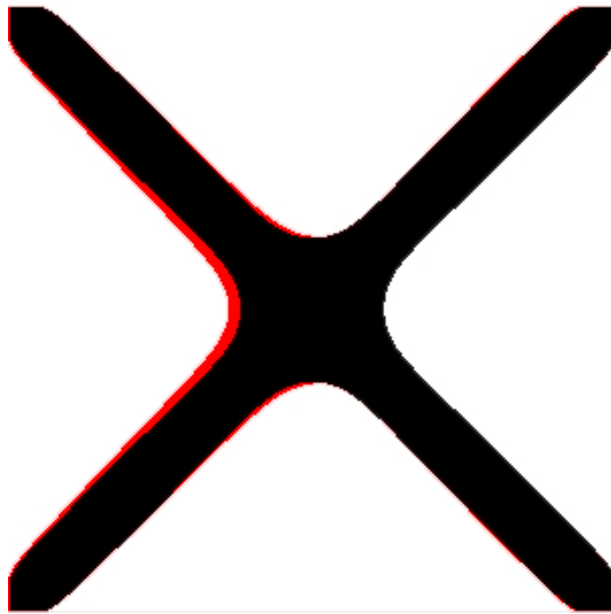


Рис. 3: Black:k=1, Red:k=10

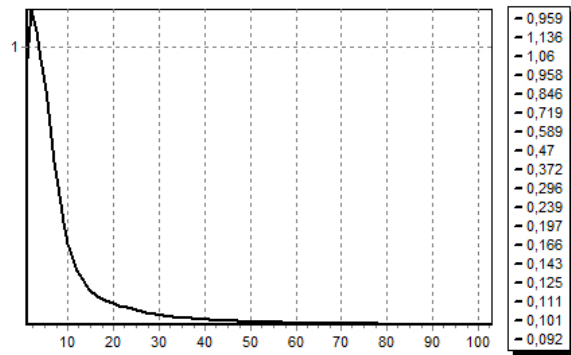


Рис. 4: residual for coefficient k=10