*Journal of Engineering Physics and Thermophysics, Vol. 82, No. 5, 2009*

# NUMERICAL ALGORITHM WITHOUT SATURATION FOR SOLVING NONSTATIONARY PROBLEMS

**S. D. Algazin**

UDC 519.632.4

*Time discretization without saturation, i.e., the discretization automatically accounting for the smoothness of the solution of the problem studied, is considered. As an example, a heat conduction equation is used, but the method is applicable to any nonstationary problem, such as where the discrete operator operating on spatial variables has a full system of eigenvectors and the eigenvalues are real.*

***Keywords:*** *numerical algorithms without saturation, nonstationary problems, heat conduction equation.*

**Introduction.** In [1], numerical algorithms without saturation for solving the stationary problems of mathematical physics are considered. In the present investigation these results are extended to nonstationary problems. Numerical algorithms without saturation were suggested by K. I. Babenko at the beginning of 70s of the last century [2] (the second augmented edition was published in 2002). At the present time the most widespread method for solving problems of the mechanics of a deformable solid body is the method of finite elements. Its drawbacks are well known: in approximating the displacement by a piecewise linear function we obtain rupture stresses. At the same time, it should be noted that the majority of the problems of the mechanics of a deformable solid body is described by elliptic-type equations that have smooth solutions. It seems of current interest to develop algorithms that could allow for this smoothness. Many-years use of this technique by the present author for elliptical eigenvalue problems has proved their high efficiency.

For example, an eigenvalue problem for the zero Bessel equation was considered: on a grid consisting of 23 nodes the first eigenvalue of this problem was found with 28 decimal places. In contrast to the classical difference methods and finite-element method, where the dependence of the speed of convergence on the number of nodes of the grid is exponential, here there is an exponential decrease of an error.

However, up to now only stationary problems have been analyzed. Below, this gap is being filled. In [3], a one-dimensional heat conduction equation is investigated; in the present work a two-dimensional heat conduction equation is considered. First, we will describe the problem in which a nonlinear heat conduction equation with variable coefficients appears. This is the problem of gas percolation in a porous medium. The sought-for equation has the form

$$\frac{\partial (\varepsilon\rho)}{\partial t} + \operatorname{div} (\rho \mathbf{v}) = 0 , \tag{1}$$

where $\varepsilon$ for real beds has a value within the range 0.15–0.22; $\varepsilon\rho$ is the concentration. Equation (1) is derived from the ordinary mass conservation law:

$$\frac{d}{dt} \int\limits_{V_{\text{por}}} \rho d\tau = \frac{d}{dt} \int\limits_{V} \rho\varepsilon d\tau = 0 , \tag{2}$$

where both volumes are immobile. From Eq. (2), with the aid of differentiation with respect to a mobile volume [4], we obtain

$$\frac{\partial (\varepsilon\rho)}{\partial t} = \operatorname{div} (\varepsilon\rho \mathbf{w}) , \quad \mathbf{v} = \varepsilon \mathbf{w} ,$$

Institute of Problems of Mechanics, Russian Academy of Sciences, 101-1 Vernadskii Ave., Moscow, 119526, Russia; email: algazinsd@mail.ru. Translated from Inzhenerno-Fizicheskii Zhurnal, Vol. 82, No. 5, pp. 950–960, September–October, 2009. Original article submitted March 6, 2008; revision submitted December 23, 2008.

 

as a result of which we arrive at Eq. (1).

The Darcy law (1856) is valid for slow motions of a liquid in an isotropic porous medium, i.e., for small values of Re ($\mathrm{Re} < \mathrm{Re}_{cr}$):

$$\mathbf{v} = -\frac{\kappa}{\mu_g} \operatorname{grad} p \,. \tag{3}$$

For real porous media $\kappa$ = 100–1000 mD. The permeability is a geometric characteristic of a porous medium, i.e., it is determined by the dimensions of particles, their shape, and packing. The equation of state has the form $\rho = \frac{M_g}{RT}\frac{p}{z(p)}$, and $z(p)$ is determined experimentally ($z(p) = 1$ for a perfect gas, i.e., a barotropic one). Equation (1) relates to the case where there are no gas sources in a bed (wells). In the general case the continuity equation has the form

$$\frac{\partial (\varepsilon\rho)}{\partial t} + \operatorname{div} (\rho\mathbf{v}) = f(z, t) \,, \quad z \in G \,, \tag{4}$$

where $f(z, t)$ is the given function; $G$ is the two-dimensional region with a smooth boundary $\partial G \in C^{\infty}$. Let $z = \varphi(\zeta) = r \exp(i\theta)$ be the conformal mapping of a single circle onto the region $G$. We will write out Eq. (4) in new variables [4]:

$$ds^2 = (dr^2 + r^2 d\theta^2) \left|\varphi'(\zeta)\right|^2 \Rightarrow g_{11} = \left|\varphi'(\zeta)\right|^2, \quad g_{22} = r^2 \left|\varphi'(\zeta)\right|^2, \quad \sqrt{g} = \left|\varphi'(\zeta)\right|^2 r \,,$$

$$\operatorname{grad} p\big|_r = \frac{1}{|\varphi'(\zeta)|}\frac{\partial p}{\partial r}, \quad \operatorname{grad} p\big|_\theta = \frac{1}{|\varphi'(\zeta)|\, r}\frac{\partial p}{\partial \theta} \,.$$

Substituting the components of the gradient into Eqs. (3) and (4), we obtain

$$\frac{\partial (\varepsilon\rho)}{\partial t} = \left|\varphi'(\zeta)\right|^{-2} L(w) + f(\zeta, t) \,, \quad \zeta = r\exp(i\theta) \,, \quad 0 \le r \le 1 \,, \quad 0 \le \theta < 2\pi \,, \quad |\zeta| \le 1 \,; \tag{5}$$

$$L(w) = \frac{1}{r}\frac{\partial}{\partial r}\left(r\kappa(r, \theta)\frac{\partial w}{\partial r}\right) + \frac{1}{r^2}\frac{\partial}{\partial \theta}\left(\kappa(r, \theta)\frac{\partial w}{\partial \theta}\right). \tag{6}$$

Here $\varepsilon = \varepsilon(r, \theta)$; $p = p(r, \theta, t)$; $\kappa = \kappa(r, \theta, p) = \kappa(r, \theta)\psi(p)$; $\mu_g = \mu_g(p)$; $w(p) = \int \frac{\rho(p)\psi(p)}{\mu_g(p)}\, dp$. Thus, the expressions (5) and (6) represent the sought-for formulation of the percolation problem. These equations should be supplemented with the boundary condition

$$\left.\frac{\partial p}{\partial n}\right|_{\partial G} = 0 \,, \tag{7}$$

which means the absence of a gas flow through the boundary of the region $\partial G$ (this follows from Eq. (3)). Note that the function $w$ also satisfies this boundary condition. Thus, the problem of gas percolation in a porous medium is reduced to a nonlinear heat conduction equation with variable coefficients, but in the present work we will consider only a linear heat conduction equation with variable coefficients.

**Discretization of a Two-Dimensional Problem over Spatial Variables.** For the discretization of problem (5)–(7) we will first perform discretization of the operator $L(w)$. We will consider the spectral problem

$$L(w) + \lambda w = 0 \,, \quad \left.\frac{\partial w}{\partial r}\right|_{r=1} = 0 \,. \tag{8}$$

Note that $-\int\limits_{|\zeta|\le 1} L(w)w\,d\zeta = \int\limits_{|\zeta|\le 1}\left[\kappa\left(\frac{\partial w}{\partial r}\right)^2 + \frac{\kappa}{r^2}\left(\frac{\partial w}{\partial \theta}\right)^2\right]d\zeta$. Thus, the boundary-value problem (8) is equivalent to the following extreme problem:

$$J(w) = \int\limits_{|\zeta|\le 1}\left[\kappa\left(\frac{\partial w}{\partial r}\right)^2 + \frac{\kappa}{r^2}\left(\frac{\partial w}{\partial \theta}\right)^2 - \lambda w^2\right]d\zeta \to \min . \tag{9}$$

In fact, $\delta J$ (variation of the functional $J$) is the principal linear part of the increment $J(w+h) - J(w)$, whence we obtain

$$\delta J = 2\int\limits_{|\zeta|\le 1}\left[\kappa\, w_r h_r + \frac{\kappa}{r^2} w_\theta h_\theta - \lambda w h\right]d\zeta$$

$$= 2\left\{\kappa r w_r h\,\big|_{r=1} - \int\limits_{|\zeta|\le 1}\left[\frac{1}{r}\frac{\partial}{\partial r}(r\kappa w_r) + \frac{1}{r^2}\frac{\partial}{\partial \theta}(\kappa w_\theta) + \lambda w\right]h\,d\zeta\right\} = 0 .$$

Since $h$ is an arbitrary function, we obtain relations (8). Thus, in searching for the minimum of functional (9) there is no need to satisfy the Neumann boundary-value condition beforehand, i.e., this boundary-value condition is natural. For the discretization of functional (9) we will apply the quadrature formula:

$$\int\limits_{|\zeta|\le 1} F(\zeta)\,d\sigma = \sum_{\nu,l} c_{\nu l} F_{\nu l}, \quad F_{\nu l} = F\left(r_\nu \exp(i\theta_l)\right), \tag{10}$$

$$r_\nu = 0.5 + 0.5\cos\frac{(2\nu-1)\,\pi}{2m}, \quad \nu = 1, 2, ..., m; \quad \theta_l = \frac{2\pi l}{N}, \quad l = 0, 1, ..., 2n_\theta; \quad N = 2n_\theta + 1 .$$

It is obtained after replacement of the function under integral by an interpolation formula for the function of two variables in the circle:

$$\left(P_{M_p} F\right)(r, \theta) = \sum_{l=0}^{2n_\theta}\sum_{\nu=1}^{m} F_{\nu l} L_{\nu l}(r, \theta), \quad F_{\nu l} = F(r_\nu, \theta_l), \tag{11}$$

$$L_{\nu l}(r, \theta) = \frac{T_m(2r-1)}{NT_m'(2r_\nu - 1)(r - r_\nu)} D_{n_\theta}(\theta - \theta_l); \quad D_{n_\theta}(\theta) = 0.5 + \sum_{k=1}^{n_\theta}\cos k\theta; \quad T_m(x) = \cos(m \arccos x) .$$

The interpolation equation (11) possesses the needed properties. Indeed, it is exact on polynomials of two variables of degree $\omega = \min(n, m-1)$. We will designate the set of these polynomials by $P_\omega$, and $E_\omega$ will designate the best approximation of the function $F \in C[D]$ ($D$ is the single circle) by the polynomial of $P_\omega$. This will determine the projector

$$P_{M_p} C[D] \to L^{M_p}, \quad L^{M_p} = L\left(L_1, ..., L_{M_p}\right),$$

where $L_1$, ..., $L_{M_p}$ are the fundamental functions of the interpolation formula (11) numbered by the same index. The following classical inequality is valid:

$$\left|F(r, \theta) - \left(P_{M_p} F\right)(r, \theta)\right| \le \left(1 + \left|P_{M_p}\right|_\infty\right) E_\infty(F), \tag{12}$$

in which $|P_{M_p}|_\infty$ is the norm of the projector $P_{M_p}$. Just as in a one-dimensional case, inequality (12) shows that the corresponding interpolation formula has no saturation. The norm of the projector $P_{M_p}$ satisfies the relation $|P_{M_p}|_\infty = O(\ln^2 M_p)$; moreover this estimate can be easily refined. Making some assumptions on the smoothness of the class of interpolated functions, one can estimate the rate of decrease of the best approximation of $E_\omega$ for $M_p \to \infty$ and obtain specific estimates of the error of the interpolation formula (11). Let $F(r, \theta) = (P_{M_p}F)(r, \theta) + \rho_{M_p}(r, \theta; F)$, where $\rho_{M_p}(r, \theta; F)$ is the error of the interpolation formula (11) (residual). Then the following theory advanced by K. I. Babenko [2, pp. 238–239] is valid.

T h e o r e m 1. *Let the class of functions* $H_\infty^{M_p}(K; D) \subset C(D)$ *in the circle* $D$ *satisfy the conditions* $\left|\dfrac{\partial^{k+1}F}{\partial x^k \partial y^1}\right| \le K$, $k+1 \le \mu$. *Then, if* $F \in H_\infty^{M_p}(K; D)$, *then*

$$|\rho_M(.; F)|_\infty \le c_\mu K M_p^{-\mu/2} \log^2 M_p, \tag{13}$$

*where* $c_\mu$ *is a constant depending on* $\mu$.

Thus, from consideration of Eq. (13) it is seen that at the same number of nodes of the interpolation of $M_p$ the rate of a decrease of the error of interpolation formula (11) increases with $\mu$, i.e., with increase in the smoothness of the interpolated function $F$. This means that the interpolation formula obtained has no saturation.

Based on the interpolation formula (11), one can easily construct a quadrature formula for calculating certain integrals, when a circle is the region of integration. Actually, replacing the integrand by expression (11), we obtain the quadrature formula (10), where $d\sigma$ is an element of the area; $c_{\nu l}$ are the weight coefficients, and $\delta(F)$ is the error; $c_{\nu l} = \int\limits_D L_{\nu l}(r, \theta) d\sigma$. In this case, $c_{\nu l}$ is independent of $l$. We will introduce into consideration the block-diagonal matrix $C$ = diag $(c_1, c_2, ..., c_m)$, where $c_\nu$ ($\nu$ = 1, 2, ..., $m$) are the diagonal matrices of size $N \times N$ with identical numbers on the diagonal. For the error of the quadrature formula we have the following estimate:

$$|\delta(F)| \le 2\pi E_\omega(F).$$

Note that all $c_{\nu l}$ are positive at a large enough number of interpolation nodes. For the coefficients of the quadrature formula (10) we have the expression

$$c_\nu = \frac{2\pi}{N}\left\{\frac{(-1)^{m+1}-1}{(m^2-1)\, m\, (-1)^{\nu-1}} \sin\theta_\nu + \frac{r_\nu}{m}\left(1 + 2\sum_{l=2(2)}^{m-1} \frac{\cos l\theta_\nu}{1-l^2}\right)\right\}, \quad r_\nu = \frac{\cos + \theta_\nu}{2}, \quad \theta_\nu = \frac{(2\nu-1)\,\pi}{2m}.$$

Next, we introduce formulas for numerical differentiation with respect to $r$ and $\theta$:

$$\left(\frac{\partial w}{\partial r}\right)_{\zeta=\zeta_{\nu l}} = \sum_{\mu=1}^{m} D_{\nu\mu}^{(r)} w_{\mu l}, \quad \left(\frac{\partial w}{\partial \theta}\right)_{\zeta=\zeta_{\nu l}} = \sum_{p=1}^{N} \tilde{B}_{lp} w_{\nu p}.$$

The matrices $\tilde{B}$ and $D^{(r)}$ have been obtained by differentiation of the altered interpolation formula (11). Over $r$, the interpolation formula that at $r$ = 1 satisfies the Neumann boundary-value problem has been applied

$$P_m(x; F) = \sum_{j=1}^{m} \left[\frac{T_m(x)}{m\dfrac{(-1)^{j-1}}{\sin\theta_j}(x - x_j)} - A_j T_m(x)\right] F_j, \quad x_j = \cos\theta_j;$$

$$\theta_j = \frac{(2j-1)\,\pi}{2m}, \quad j = 1, 2, ..., m; \quad x = 2r - 1;$$

$A_j$ will be selected so as to satisfy the boundary-value condition $F'(1) = 0$.

With the aid of the quadrature formula (10), functional (9) will be transformed into a quadratic form:

$$J(w) = \sum_{\nu,l} c_{\nu l}\left[k_{\nu l}\left(\frac{\partial w}{\partial r}\right)^2_{\zeta=\zeta_{\nu l}} + \frac{k_{\nu l}}{r_\nu^2}\left(\frac{\partial w}{\partial \theta}\right)^2_{\zeta=\zeta_{\nu l}} + \lambda w^2_{\nu l}\right], \tag{14}$$

where $k_{\nu l} = \kappa(r_\nu, \theta_l)$ is the value of the function $\kappa$ at the node of the grid. Differentiating (14) with respect to $w_{\tilde{\mu}\tilde{l}}$, we obtain

$$\sum_{p=1}^{N} B^*_{\tilde{\nu}\tilde{l},p} w_{\tilde{\nu}p} + \sum_{\mu=1}^{m} A^*_{\tilde{\nu}\tilde{l},\mu} w_{\mu\tilde{l}} = \lambda c_{\tilde{\nu}} w_{\tilde{\nu}\tilde{l}},$$

where $B^*_{\tilde{\nu}\tilde{l},p} = \dfrac{\tilde{c}_\nu}{\tilde{r}^2_\nu}\sum_{l=1}^{N} k_{\tilde{\nu}l}\tilde{B}_{lp}\tilde{B}_{l\tilde{l}}$; $A^*_{\tilde{\nu}\tilde{l},\mu} = \sum_{\nu=1}^{m} c_\nu k_{\nu\tilde{l}} D^{(r)}_{\nu\mu} D^{(r)}_{\nu\nu}$ is the discrete analog of the eigenvalue problem

$$\operatorname{div}(\kappa \operatorname{grad} w) + \lambda w = 0\,, \quad r < 1\,, \quad \left.\frac{\partial w}{\partial r}\right|_{r=1} = 0\,.$$

Estimation of the error of the discretization described can be made using the scheme given in [1, 5].

**Statement of the Problem.** In a cylinder $D = \{|\zeta| \le 1,\ 0 \le t \le 1\}$, we will consider the heat conduction equation

$$\frac{\partial u(\zeta, t)}{\partial t} = |\varphi'(\zeta)|^{-2} L(u) + f(\zeta, t) = r\exp(i\theta)\,, \quad 0 \le r \le 1\,, \quad 0 \le \theta < 2\pi\,, \quad |\zeta| \le 1\,; \tag{15}$$

$$u|_{t=0} = u_0(\zeta)\,, \tag{16}$$

$$\left.\frac{\partial u}{\partial n}\right|_{r=1} = 0\,. \tag{17}$$

Without loss of generality, we may assume that $u_0(\zeta) \equiv 0$. Otherwise, we will introduce a new unknown function $v(\zeta, t) = u(\zeta, t) - u_0(\zeta)$ which is the solution of the same boundary-value problem (15)–(17) but with a different right-hand side. The boundary condition (17) is satisfied, since it is fulfilled for the function $u_0(\zeta)$.

**Discretization in Time.** For $t$ we select a grid consisting of $k$ nodes: $t_\nu = \dfrac{1}{2}(z_\nu + 1)$, $z_\nu = \cos\chi_\nu$, $\chi_\nu = \dfrac{(2\nu - 1)\pi}{2k}$, $\nu = 1, 2, \ldots, k$, and apply interpolation by a polynomial:

$$q(t) = \sum_{\nu=1}^{k} \frac{T_k(t)\, tq_\nu}{k\dfrac{(-1)^{\nu-1}}{\sin\chi_\nu}\, t_\nu (z - z_\nu)}\,. \tag{18}$$

The quantities entering into Eq. (18) have been determined above. The values of the first derivative of $u(\zeta, t)$ over $t$ that enter into the left-hand side of relations (15) will be obtained by differentiation of the interpolation formula (18).

Let $A$ be the matrix of the discrete operator $-|\varphi'(\zeta)|^2 L(u)$; then, by having designated $u_{\mu\nu} = u(\zeta_\mu, t_\nu)$, $\mu = 1, 2, ..., N_t$; $\nu = 1, 2, ..., k$, we obtain $\dfrac{\partial u(\zeta_\mu, t)}{\partial t} + \sum\limits_{p=1}^{N_i} A_{\mu p} u(\zeta_p, t) = f(\zeta_\mu, t)$.

Let $B$ be the matrix of numerical differentiation with respect to $t$ over [0, 1]. As a result we obtain $\sum\limits_{q=1}^{k} B_{\nu q} u_{\mu q} + \sum\limits_{p=1}^{N_i} A_{\mu p} u_{p\nu} = f_{\mu\nu}$. We will number the nodes of the grid by one index along the lines, i.e., the first index $I \to (\mu, \nu) = (\nu - 1)N_t + \mu$ changes most rapidly. Then we obtain a discrete problem

$$(B \otimes I_{N_t} + I_k \otimes A)\, u = f, \tag{19}$$

where $B$ is a matrix of size $k \times k$ — differentiation with respect to $t$; $A$ is a matrix of size $N_t \times N_t$ — a discrete operator $-|\varphi'(\zeta)|^{-2} L(u)$; $I_{N_t}$ and $I_k$ are unit matrices. We will represent $A$ in the form (see [1])

$$A = \sum_p \lambda_p h_p, \quad h_p^2 = h_p, \quad h_p h_l = 0, \quad p \neq l \Rightarrow \sum_p h_p = I_m$$

$$\Rightarrow B \otimes \sum_p h_p + I_k \otimes \left( \sum_p \lambda_p h_p \right) = \sum_p (B + \lambda_p I_k) \otimes h_p$$

$$\Rightarrow (B \otimes I_{N_t} + I_k \otimes A)^{-1} = \sum_p (B + \lambda_p I_k)^{-1} \otimes h_p . \tag{20}$$

Note that the operator $-|\varphi'(\zeta)|^{-2} L(u)$ is degenerate, i.e., it has a zero eigenvalue. In this case, one has to reverse the matrix of numerical differentiation, i.e., to approximately reverse the operator of differentiation. Even though the matrix of numerical differentiation is degenerate, the reverse one exists; this is integration. For practical realization of such an approach, we will introduce formulas for numerical integration of the function that would satisfy the boundary-value condition $u|_{t=0} = 0$ over the segment [0, 1]. For $t$ we will select a grid consisting of $m$ nodes: $r_\nu = \dfrac{1}{2}(x_\nu + 1)$, $x_\nu = \cos \theta_\nu$, $\theta_\nu = \dfrac{(2\nu - 1)\pi}{2m}$, $\nu = 1, 2, ..., m$, $x = 2r - 1$, and apply the interpolation by a polynomial:

$$u(r) = \sum_{\nu=1}^{m} \frac{T_m(x)\, r u_\nu}{m \dfrac{(-1)^{\nu-1}}{\sin \theta_\nu} r_\nu (x - x_\nu)} .$$

Further we have

$$\sum_{\nu=1}^{m} \frac{T_m(x)\, u_\nu}{m \dfrac{(-1)^{\nu-1}}{\sin \theta_\nu} (x - x_\nu)} = \frac{2}{m} \sum_{l=0}^{m-1}{}' \cos l\theta_\nu T_l(x) \Rightarrow u(r) = \frac{2}{m} \sum_{\nu=1}^{m} \left( \sum_{l=0}^{m-1}{}' \cos l\theta_\nu T_l(x) \frac{r}{r_\nu} \right) u_\nu .$$

To find $\int\limits_0^r u(r)dx$, it is necessary to calculate the integral $I_l(r) = \int\limits_0^r T_l(x) r dr$, $x = 2r - 1$:

$$I_0(r) = \frac{r^2}{2}; \quad I_1(r) = \frac{2}{3} r^3 - \frac{r^2}{2}; \quad I_2(r) = 2r^4 - \frac{8}{3} r^3 + \frac{1}{2} r^2 ,$$

$$l \geq 3\,, \quad 4I_l(r) = \frac{x^2 T_l(x)}{l+2} + \frac{xT_l(x)}{l+1} - \frac{xT_{l-1}(x)}{l+2} - \frac{lT_{l-1}(x)}{l^2-1} - \frac{T_{l-2}(x)}{l^2-4} + \frac{(-1)^l}{l^2-4} - \frac{(-1)^l}{l^2-1}\,.$$

The matrix of numerical integration has the form

$$I^{(r)}_{\mu\nu} = \frac{2}{m}\sum_{l=0}^{m-1} \frac{\cos l\theta_\nu I_l(r_\mu)}{r_\mu} \Rightarrow \int_0^{r_\mu} u(r)\,dr = \sum_{\nu=1}^{m} I^{(r)}_{\mu\nu} u_\nu\,.$$

Thus, the solution of discrete problem (19) will be obtained by multiplying the matrix (20) by the vector of the right-hand side. Note that to construct a matrix reciprocal of matrix (19) it is sufficient to reverse $N_t$ matrices of size $k \times k$. We will also note that above we have nowhere used the specific feature of the matrix $A$, i.e., $A$ can be the matrix of a two-dimensional, three-dimensional, and of any other problem. It is only necessary that the matrix can have the full system of eigenvectors and that the eigenvalues be real.

**Numerical Example.** As a numerical example we will consider problem (15)–(17) with the right-hand side: $f(t, r, \varphi) = (r^3 - 3r)^3 \cos\varphi + r(r^2-3)t(r\cos 2\varphi + \cos\varphi) - \{2(r\cos\varphi + 1)[9r(r^2-3)^2(r^2-1)t\cos\varphi](r\cos\varphi + 2)18[r(r^3-3r)^2 + (r^2-1)(r^3-3r)^3(r^2-1)]t\cos\varphi\}$ and with the function $k(r, \varphi) = r\cos(\varphi) + 2$; then the solution $u(t, r, \varphi) = (r^3 - 3r)^3 t\cos\varphi$. Let $M$ be the number of points over the radius; $N$ the number of points over $\theta$ (over the circles of the grid); $K$ the number of points in time; BNORM the norm of the matrix reciprocal of the matrix of a discreet problem; BNORM the norm of the difference between an exact and approximate solutions. The results of calculations are presented below.

$$M = 5\,, \quad N = 5\,, \quad K = 5\,; \quad \text{BNORM} = 2.23\,; \quad \text{RNORM} = 0.16\,;$$

$$M = 10\,, \quad N = 5\,, \quad K = 5\,; \quad \text{BNORM} = 2.26\,; \quad \text{RNORM} = 2.86\cdot 10^{-2}\,;$$

$$M = 20\,, \quad N = 5\,, \quad K = 5\,; \quad \text{BNORM} = 2.29\,; \quad \text{RNORM} = 4.67\cdot 10^{-3}\,;$$

$$M = 100\,, \quad N = 5\,, \quad K = 5\,; \quad \text{BNORM} = 2.40\,; \quad \text{RNORM} = 3.54\cdot 10^{-4}\,;$$

$$M = 300\,, \quad N = 5\,, \quad K = 5\,; \quad \text{BNORM} = 2.61\,; \quad \text{RNORM} = 5.20\cdot 10^{-5}\,.$$

A further increase in the number of grid nodes is senseless, since the function $\kappa = \kappa(r, \theta)$ is assigned in the program with a unary accuracy. Thus, to obtain a solution with one decimal place (an acceptable accuracy for investigation of the development of gas deposits), it is sufficient to take 75 nodes of a grid in a cylinder.

**Theoretical Investigation of an Error.** We will carry out this investigation on the example of a one-dimensional problem. In a rectangle $D = \{0 \leq x \leq 1,\ 0 \leq t \leq 1\}$ we will consider the heat conduction equation:

$$\frac{\partial u(x,t)}{\partial t} = \frac{\partial^2 u(x,t)}{\partial x^2} + f(x,t)\,, \quad (x,t) \in D\,; \quad u\big|_{t=0} = u_0(x)\,; \quad u\big|_{x=0} = u\big|_{x=1} = 0\,. \tag{21}$$

As has already been noted above, without loss of generality we may assume that $u_0(x) \equiv 0$.

**Discretization of a One-Dimensional Problem.** Over $x$ we approximate the sought-for function $u(x, t)$ by a polynomial; for this purpose, over $x$ we take a grid consisting of $m_1$ nodes:

$$x_\mu = \frac{1}{2}(z_\mu + 1)\,, \quad z_\mu = \cos\chi_\mu\,, \quad \chi_\mu = \frac{(2\mu - 1)\pi}{2m_1}\,, \quad \mu = 1, 2, ..., m_1\,,$$

and apply the interpolation formula:

$$q(x) = \sum_{\mu=1}^{m_1} \frac{T_{m_1}(x)\,(x-1)\,xq_k}{m_1 \dfrac{(-1)^{\mu-1}}{\sin\chi_\mu}\,(x_\mu - 1)\,x_\mu\,(z - z_\mu)}, \quad q_\mu = q(x_\mu), \quad z = 2x - 1. \tag{22}$$

The second derivative with respect to $x$ entering into Eq. (21) will be found by differentiation of the interpolation formula (22). Over $t$ we will select a grid consisting of $k$ nodes:

$$t_\nu = \frac{1}{2}(z_\nu + 1), \quad z_\nu = \cos\chi_\nu, \quad \chi_\nu = \frac{(2\nu - 1)\,\pi}{2k}, \quad \nu = 1, 2, ..., k,$$

as well as apply interpolation by a polynomial:

$$q(t) = \sum_{\nu=1}^{k} \frac{T_k(t)\,tq_\nu}{k \dfrac{(-1)^{\nu-1}}{\sin\chi_\nu}\,t_\nu\,(z - z_\nu)}. \tag{23}$$

The quantities entering into Eq. (23) have been determined above. The values of the first derivative at the nodes of the grid from $u(x, t)$ over $t$ that enter into the left-hand side of relation (22) will be obtained by differentiation of the interpolation formula (23).

The traditional methods of solving this problem are varied (see, e.g., [6]). The main drawback of the difference methods is that they are with saturation. Irrespective of the smoothness of a solution, the error of discretization of the difference method in spatial variables is $O(h^p)$, where $p$ is the order of the difference scheme. An analogous statement is valid for discretization in time, e.g., for an explicit scheme of first order in time the error of discretization in time is $O(\tau)$. The total error cannot be smaller than the maximum of these values.

In contrast to the difference methods, in the present work an approximation of the solution by polynomials is applied. Let $f(x) \in C[a, b]$ be a continuous function and $(P_n f)(x)$ be its interpolation polynomial. Thereby the projector $P_n$: $C \to L^n$ has been determined, where $L^n \subset C$ is the corresponding $n$-dimensional sub-space of polynomials. Then

$$\left| f(x) - (P_n f)(x) \right| \le \left(1 + \left| P_n \right|_\infty\right) E_{n-1}(f).$$

Here $\left| P_n \right|_\infty$ is the norm of the projector and $E_{n-1}(f)$ is the best approximation of the function $f$ by polynomials of degree not higher than $(n - 1)$ in the norm $C$. Moreover, the nodes are selected so that $\left| P_n \right|_\infty = O(\ln(n))$. According to the Weierstrass theorem for each continuous function we have $\lim\limits_{n\to\infty} E_n(y) = 0$; the rate of a decrease in $E_n(y)$ for $n \to \infty$ depends on the smoothness of the function $y$.

Thus, let $y(x)$ be a continuous function given in the interval $[-1, +1]$, $P_n(x)$ be the polynomial of degree $n$ that deviates least from $y(x)$ in the interval considered, and $E_n(y)$ be the best approximation of $y(x)$ by means of a polynomial of degree $n$ so that

$$E_n(y) = \max_{|x|\le 1} \left| y(x) - P_n(x) \right|,$$

then the Jackson theorem holds [7, p. 296]:

T h e o r e m 2. *If the function $y(x)$ in the interval $-1 \le x \le +1$ has a continuous derivative $y^{(p)}(x)$ that satisfies the Lipschitz condition*

$$\left| y^{(p)}(x') - y^{(p)}(x'') \right| < K \left| x' - x'' \right|, \quad \left(x' \ne x''; \left| x' \right|, \left| x'' \right| \le 1\right),$$

*then for its best approximation by means of ordinary polynomials the inequality $E_n(y) < \dfrac{c_p K}{n^{p+1}}$ ($c_p = \dfrac{c^{p+1}(p+1)^{p+1}}{(p+1)!}$, $n \ge p + 1$) is valid, where $c$ is an absolute constant.*

Theorem 2 shows the rate of a decrease of the best approximation of $E_n(y)$ depending on the smoothness of the function $y$.

Thus, on the same grid ($n$ is fixed) the considered approximation of the function is improved with increase in the smoothness of the solution. Moreover, we may not know a priori the smoothness, but the method will adjust itself to it. This is the crux of the methods without saturation suggested by K. I. Babenko.

A brief account of the principles of the theory of unsaturable numerical methods is contained in the first edition of the book by K. I. Babenko [2]. Note that investigations in the computational mathematics along these lines have not been adequately propagandized and up to now are practically not known abroad. This is confirmed by the fact that today there has begun factual "rediscovery" (evidently independent) of these very computational methods in the West under the name of "spectral" methods (S. Orszag, D. Gotlieb, E. Tadmor, USA), as well as in the form of the present-day ($h$–$p$) specializations of the method of finite elements (O. Widlund, USA and S. Schwab, Switzerland) in which, on making meshes of a grid finer (i.e., when $h \to 0$), the degree $p$ of polynomials used in approximation of functions inside one finite element increases simultaneously. We regret that by now the works of K. I. Babenko and of his pupils have been practically forgotten.

Let $A$ be the matrix of the discrete operator $-\dfrac{d^2}{dx^2}$; then, designating $u_{\mu\nu} = u(x_\mu, t_\nu)$, $\mu = 1, 2, ..., m_1$, $\nu = 1, 2, ..., k$, we obtain $\dfrac{\partial u(x_\mu, t)}{\partial t} + \sum\limits_{p=1}^{m_1} A_{\mu p} u(x_p, t) = f(x_\mu, t)$. Let $B$ be the matrix of numerical differentiation with respect to $t$ over [0, 1]. As a result we find

$$\sum_{q=1}^{k} B_{\nu q} u_{\mu q} + \sum_{p=1}^{m_1} A_{\mu p} u_{p\nu} = f_{\mu\nu} . \tag{24}$$

We will number the grid nodes by one index along the lines (i.e., the first index $I \to (\mu, \nu) = (\nu - 1)m_1 + \mu$ changes most rapidly). Then we obtain the discrete problem:

$$(B \otimes I_{m_1} + I_k \otimes A)\, u = f , \tag{25}$$

where $B$ is the matrix of size $k \times k$ — differentiation with respect to $t$; $A$ is the matrix of size $m_1 \times m_1$ — the second differentiation with respect to $x$; $I_{m_1}$ and $I_k$ are single matrices. Next, proceeding as in derivation of Eq. (20), we obtain

$$(B \otimes I_{m_1} + I_k \otimes A)^{-1} = \sum_{p} (B + \lambda_p I_k)^{-1} \otimes h_p . \tag{26}$$

Thus, the solution of discrete problem (24) will be obtained by the multiplication of matrix (26) by the vector of the right-hand side of $f$. Note that to construct a matrix reciprocal of (26) it is sufficient to reverse $m_1$ matrices of size $k \times k$, where $k$ is the number of the nodes of interpolation in time. We note also that nowhere was the specificity of the matrix $A$ used, i.e., $A$ can be the matrix of a two-dimensional, three dimensional, and of any other problem. It is only necessary that the matrix can have a full system of eigenvectors and that the eigenvalues be real.

To construct the discretization of the above-described problem, we had to differentiate the interpolation formulas. To estimate the error of this operation, there exists the following theorem [2].

T h e o r e m 3. *Let $f \in W_\infty^n(M, I)$, $0 \le s < n$, $s$ be the whole number. Then*

$$\left| f^{(s)}(x) - p^{(s)}(x; f) \right| \le M \frac{(b-a)^{n-s}}{(n-s)!} , \quad x \in [a, b] .$$

The stability of numerical methods in the classical case is considered for the difference schemes [2, p. 763] in the following way. We will assume that there exist two linear normalized spaces $F$ and $U$ such that the equation

$$\mathfrak{L}u = f \tag{27}$$

is solvable, and the solution is the single one for any element $f \in F$ and $\mathfrak{L}^{-1}F \in U$. The norms in these spaces are designated by $\|\cdot\|_U$ and $\|\cdot\|_F$.

In the region $\Omega$ we consider the grid of $\Omega_h$ depending on the parameter $h$. We introduce the mappings of $J_h$: $U \to U_h$, $J_h$: $u \to u_h$, $I_h$: $F \to F_h$, $I_h$: $f \to f_h$, where $U_h$ and $F_h$ are the finite-dimensional spaces, and we assume that $F_h$ contains all the grid functions on $\Omega_h$. On the grid functions we introduce two norms: $\|\cdot\|_U$ and $\|\cdot\|_F$ that obey the conditions

$$\lim_{h\to 0} \|J_h u\|_{U_h} = \|u\|_U, \quad \lim_{h\to 0} \|I_h f\|_{F_h} = \|f\|_F$$

for arbitrary $u \in U$, $f \in F$. Let

$$\mathfrak{L}_h v_h = f_h \tag{28}$$

be the discretization of Eq. (27).

The operator $\mathfrak{L}_h$ is well conditioned with the order $\rho$, if for any grid function $\nu_h$

$$\|v_h\|_{U_h} \le M h^{-p} \|\mathfrak{L}_h v_h\|_{F_h} .$$

We recall that the operator $\mathfrak{L}_h$ approximates the operator $\mathfrak{L}$ with the order $\omega$, if

$$\|I_h \mathfrak{L}u - \mathfrak{L}_h J_h u\|_{F_h} \le C(u)\, h^{\omega} .$$

T h e o r e m (of Ryaben'kii–Filippov). *If the operator $\mathfrak{L}_h$ approximates the operator $\mathfrak{L}$ with the order $\omega$ and is well conditioned with the order $\rho$, then for the solution of problem (27), (28) the following estimate is valid:*

$$\|J_h u - v_h\| U_h \le C_1(u)\, h^{\omega-\rho} .$$

P r o p o s a l 1. $|u|_\infty < |f|_\infty$ (see Eqs. (24)–(26)).

P r o o f of proposal 1. We easily obtain that this norm of matrix (26) does not exceed the maximum norm of the matrices $(B + \lambda_p I_k)^{-1}$ (these matrices are the analogs of eigenvalues). To construct these matrices, one does not need to supply the procedure of numerical reversal. Note that in this case the differential operator on the left-hand side of the relation $y' + ay = g(x) \Rightarrow y = \exp(-ax)(c + \int g(x) \exp(ax)dx)$ is reversed in our case: $y(t) = \int_0^t g(x) \exp(a(x-t))dx$. Consequently, $|y|_\infty \le |g|_\infty$, since $a > 0$, which was to be proved.

Proposal 1 points to the stability of the considered algorithm of the solution of the one-dimensional heat condition equation on its right-hand side. To investigate the stability over the initial data (in the case where they are nonzero), we introduce the function

$$v(x, t) = u(x, t) - u_0(x) \to \frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2} + u_0''(x) + f(x, t), \quad v(0, t) = v(1, t) = 0, \quad v(x, 0) = 0 .$$

In a discrete form $\tilde{G}\nu = -e \otimes Au_0 + f$, $e = (1, 1, ..., 1)'$ is the vector-column of dimensionality $k$, where the matrix $\tilde{G}$ has been defined on the right-hand side of relation (28).

P r o p o s a l 2. $|v|_\infty < |u_0|_\infty + |f|_\infty$.

P r o o f of proposal 2. We avail ourselves of the property of the Kronecker product [8, p. 20]: $(\tilde{A} \otimes \tilde{B})(\tilde{C} \otimes \tilde{D}) = \tilde{A}\tilde{C} \otimes \tilde{B}\tilde{D}$, as well as of the evident statement that $h_p A = \lambda_p h_p$, from which the relation $v = -\sum_p [(B + \lambda_p I_p)^{-1} e] \otimes \lambda_p (h_p u_0) + \tilde{G}^{-1} f$ follows: the calculation of the product $[(B + \lambda_p I_p)^{-1} e]$ is equivalent to the com-

putation of the integral $\int_0^t \exp(a(x-t))dx = \frac{1}{a}(1 - \exp(-at)) < \frac{1}{a}$ at $a > 0$, $a = \lambda_p$ (see also the proof of proposal 1). From this and proposal 1, the proof of proposal 2 follows.

We will consider an example. Let $u_0(x) = x(x-1)$ and the right-hand side $f(x, t) = (\cos t - \pi^2 \sin t) \times \sin \pi x + 2$, then $u(x, t) = \sin(\pi t) \sin(\pi x) + x(x-1)$. We introduce the new function $v(x, t) = u(x, t) - x(x-1)$; then $v(x, 0) = 0$, and the right-hand side of heat conduction equation (22) is $F(x, t) = (\cos t - \pi^2 \sin t) \sin(\pi x)$. Consequently, the problem can be solved by the technique described above for the heat conduction equation with zero initial conditions. This example was considered in [3]. On a $5 \times 5$ grid five decimal places have been obtained.

**Conclusions.** From the arguments given it follows that a multilayer, implicit, unconditionally stable method without saturation has been obtained for solving nonstationary problems of mathematical physics. The discrete operator over spatial variables must have a full system of eigenfunctions, and the corresponding eigenvalues must be real. The stability can be guaranteed for negatively determined operators.

This work was carried out with financial support from the Russian Foundation of Basic Research, project No. 09-08-00011-a.

## NOTATION

$A$, matrix of a discrete operator over spatial variables, of size $m_1 \times m_1$ for a one-dimensional problem and of size $N_t \times N_t$ for a two-dimensional problem; $B$, matrix of numerical differentiation with respect to time, of size $k \times k$; $\tilde{B}$, matrix of numerical differentiation with respect to $\theta$ of a two-dimensional interpolation formula, of size $N \times N$, $N = 2n_\theta + 1$; $D$, square $\{0 \le x \le 1,\ 0 \le t \le 1\}$; $D^{(r)}$, matrix of numerical differentiation with respect to $r$ of a two-dimensional interpolation formula, of size $m \times m$; $f(z, t)$, density of gas sources, kg/(m$^2$·sec); $G$, region in which gas percolation is considered; $\partial G$, boundary of region $G$; $k$, number of nodes of interpolation in time; $L$, two-dimensional differential operator standing on the right-hand side of percolation equation; $m$, number of nodes over the radius in two-dimensional interpolation; $m_1$, number of nodes over spatial variable in a one-dimensional problem; $M_p$, number of nodes of interpolation in the region considered; $N$, number of nodes over $\theta$ in two-dimensional interpolation, $N = 2n_\theta + 1$; $N_t m \times N$, number of nodes of interpolation in a two-dimensional region; $n_\theta$, $N = 2n_\theta + 1$; $p$, pressure, Pa; $r$, polar coordinate in a circle of single radius, m; $R$, universal gas constant, J/(K·mole); Re, Reynolds number; $T$, absolute temperature, K; $t$, time, sec; $v$, percolation velocity, m/sec; $V_{por}$, volume of pores, m$^3$; $V$, total volume, m$^3$; $w$, gas velocity, m/sec; $\varepsilon$, porosity; $\kappa$, permeability, D (1D = $10^{-8}/0.981$ cm$^2$; $\mu_g$, dynamic viscosity of a gas, Pa·sec; $\rho$, density, kg/m$^3$; $\varphi(\zeta)$, function assigning conformal mapping of a circle of single radius $|\zeta| \le 1$ onto the region $G$ considered. Subscripts: cr, critical; g, gas; p, point; por, porous.

## REFERENCES

1. S. D. Algazin, *Numerical Algorithms without Saturation in the Classical Problems of Mathematical Physics* [in Russian], Nauchnyi Mir, Moscow (2002).
2. K. I. Babenko, *Principles of Numerical Analysis* [in Russian], Nauka, Moscow (1986).
3. S. D. Algazin, *Numerical Algorithms of the Classical Mathematical Physics. XIV. Numerical Algorithm without Saturation for Solving the Heat Conduction Equation*, Preprint No. 816 of the Institute of Applied Mechanics, Russian Academy of Sciences, Moscow (2008).
4. L. I. Sedov, *Mechanics of a Continuous Medium* [in Russian], Vol. 1, Nauka, Moscow (1970).
5. S. D. Algazin, Concerning the localization of the eigenvalues of closed linear operators, *Sib. Mat. Zh.*, **24**, No. 2, 3–8 (1983).
6. V. I. Lebedev, *Explicit Difference Schemes with Variable Time Steps for Solving Strict Systems of Equations*, Preprint No. 177 of the Dept. of Higher Mathematics, Academy of Sciences of the USSR, Moscow (1987).
7. V. L. Goncharov, *Theory of Interpolation and Approximation of Functions* [in Russian], Gostekhizdat, Moscow (1934).
8. M. Markus and H. Mink, *Review on the Theory of Matrices and Matrix Inequalities* [Russian translation], Nauka, Moscow (1972).