ALTERNATING DIRECTION ITERATIONAL SCHEMES FOR THE NUMERICAL SOLUTION OF THE DIRICHLET PROBLEM*

A. A. SAMARSKII and V. B. ANDREEV

(Moscow)

(*Received 8 June 1964*)

The fourth order difference scheme of approximation on the rectangular mesh ($h_\alpha \neq h_\beta$ for $\alpha \neq \beta$) will be considered for Poisson's equation. We shall prove the convergence in the mean of the schemes in question at the rate $O(|h|^4)$, for the Dirichlet problem in the p-dimensional rectangular parallelepiped ($p = 2, 3$), where

$$|h|^2 = \sum_{\alpha=1}^{p} h_\alpha^2,$$

whatever the ratio h_α between the intervals. The conditions under which the maximum principle holds for the proposed schemes on a rectangular mesh will be discussed, and they will be shown to be uniformly convergent at a rate $O(|h|^4)$ for $p \leqslant 4$.

An alternating directions iterational process will be considered, and the choice of sequence of iterational parameters $\{\tau_n\}$ "reasonably high" speed of convergence of the process will be discussed. The choice of optimum ratios between the terms of the sequence $\{\tau_n\}$, minimizing the number of iterations, will also be examined.

1. Given the p-dimensional parallelepiped

$$\bar{D}_p = \{x = (x_1, \ldots, x_p) : 0 \leqslant x_\alpha \leqslant l_\alpha, \quad \alpha = 1, \ldots, p\}$$

with boundary Γ, we seek the solution in it of the problem

---

* *Zh. Vych. Mat.* 4, No. 6, 1025-1036, 1964.

$$Lu \equiv \sum_{\alpha=1}^{p} L_\alpha u = -f(x), \qquad L_\alpha u = \frac{\partial^2 u}{\partial x_\alpha^2}, \tag{1}$$

$$u|_\Gamma = g(x). \tag{2}$$

Let $\bar{\omega}_h = \{x_i = (i_1 h_1, \ldots, i_p h_p) \in \bar{D}_p: \quad i_\alpha = 0, \ldots, N_\alpha; \quad h_\alpha = \frac{l_\alpha}{N_\alpha}, \alpha = 1, \ldots, p\}$ be a difference mesh, uniform with respect to each of directions x_α, and $\gamma = \{x_i \in \Gamma\}$ the boundary of the mesh $\bar{\omega}_h$.

Problem (1) - (2) was considered in [1] - [3] with $p = 2, 3$ on a difference mesh $\bar{\omega}_h$ with $h_\alpha = h$, $\alpha = 1, \ldots, p$, and difference schemes of the fourth order of approximation to the sufficiently smooth solution of equation (1) were proposed; provided the sides of the parallelepiped l_α, $\alpha = 1, \ldots, p$ were comparable in size, the schemes were shown to be uniformly convergent at the rate $O(h^4)$.

An alternating directions iterational process (see (13) with $\sigma = 1$) was proposed for these schemes in [4]* with the number of iterations

$$\nu \asymp \nu_0 \lg \frac{1}{h} \lg \frac{1}{\varepsilon}, \tag{3}$$

where $\varepsilon > 0$ is the required accuracy. The choice of optimum iterational parameters $\{\tau_n\}$ minimizing ν_0 was likewise discussed. The expression for the rate of convergence was only proved in [4] for $p = 2$.

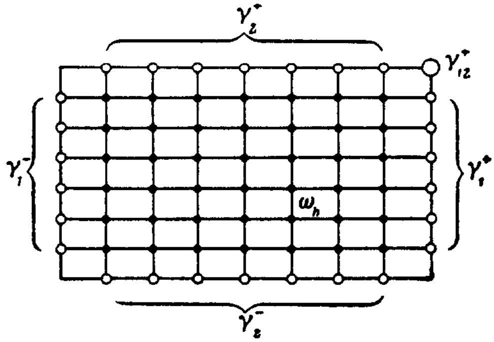

FIG. 1.

* In [4], note 1, there is an error in the evaluation of d and φ. The correct expressions are

$$d = q + \frac{h^2}{12} q^2, \qquad \varphi = f + \frac{h^2}{12}(qf + Lf) \text{ and } q = \text{const}.$$

In the present paper we consider the more general family of iterational schemes with

$$\sigma \geqslant \frac{1}{2}, \tag{4}$$

where σ is a parameter characterizing the iterational process. A "one-dimensional" procedure is proposed for choosing the iterational parameters $\{\tau_n\}$, minimizing ν_0 for difference schemes of the fourth order of accuracy on the rectangular mesh ($h_\alpha \neq h_\beta$ for $\alpha \neq \beta$). (Expression (3) and the optimum value of ν_0 are also obtained in passing for schemes of the 2nd order of accuracy.)

2. We consider the following difference scheme * for the approximation of problem (1) - (2):

$$\Lambda' y + \varphi = 0, \qquad y|_\gamma = g(x), \tag{5}$$

where

$$\Lambda' = \Lambda + \frac{\theta}{12}\sum_{\alpha=1}^{p} h_\alpha^2 \sum_{\beta\neq\alpha}^{1-p} \Lambda_\alpha\Lambda_\beta, \qquad \Lambda = \sum_{\alpha=1}^{p}\Lambda_\alpha, \qquad \Lambda_\alpha y = y_{\bar{x}_\alpha x_\alpha},$$
$$\varphi = f + \frac{\theta}{12}\sum_{\alpha=1}^{p} h_\alpha^2 \Lambda_\alpha f, \qquad \theta = 0,\ 1. \tag{6_θ}$$

Here (see [6])

$$x = x_i = (h_1 i_1, \ldots, h_p i_p),$$
$$x^{(\pm 1_\alpha)} = (h_1 i_1, \ldots, h_{\alpha-1} i_{\alpha-1},\ h_\alpha (i_\alpha \pm 1),\ h_{\alpha+1} i_{\alpha+1}, \ldots, h_p i_p),$$
$$y = y_i = y(x), \qquad y^{(\pm 1_\alpha)} = y(x^{(\pm 1_\alpha)}),$$
$$y_{\bar{x}_\alpha} = (y - y^{(-1_\alpha)})/h_\alpha, \qquad y_{x_\alpha} = (y^{(+1_\alpha)} - y)/h_\alpha.$$

It is easily shown (see also [5]) that, for $\theta = 1$, scheme (5)-(6_θ) (in future scheme (5)-(6_1)) has the 4th order of approximation in $|h|$ on the class of sufficiently smooth solutions of (1), so that

$$\psi = \Lambda' u + \varphi = O(|h|^4).$$

For $\theta = 0$ scheme (5)-(6_θ) (scheme (5)-(6_0)) becomes the familiar scheme

---

* Scheme (5) - (6^1) for $f = 0$ and $p = 2$ was proposed without a proof of convergence in [5].

of the second order of accuracy.

Let us show that (5)-(6_1) has the 4th order of accuracy in $|h|$. Let u be a solution of problem (1)-(2) and y a solution of problem (5)-(6_1). We now obtain for the function $z = y - u$:

$$\Lambda' z + \psi = 0, \qquad z|_\gamma = 0. \tag{7}$$

We require the scalar products (see [6])

$$(y, v) = \sum_{\omega_h} yvH, \quad (yv]_\alpha = \sum_{\omega_h^{+\alpha}} yvH, \quad (y, v]_{\alpha,\beta} = \sum_{\omega_h^{+\alpha+\beta}} yvH$$

and the corresponding norms

$$\|v\| = \sqrt{(v, v)}, \quad \|v_{\bar{x}_\alpha}\| = \sqrt{(v_{\bar{x}_\alpha}, v_{\bar{x}_\alpha}]_\alpha}, \quad \|v_{\bar{x}_\alpha \bar{x}_\beta}\| = \sqrt{(v_{\bar{x}_\alpha \bar{x}_\beta}, v_{\bar{x}_\alpha \bar{x}_\beta}]_{\alpha,\beta}},$$

where

$$H = \prod_{\alpha=1}^{p} h_\alpha,$$

$$\gamma_\alpha^+ = \{x_i \in \Gamma : i_\alpha = N_\alpha;\ i_\beta \neq 0,\ N_\beta \text{ for } \beta \neq \alpha\},$$

$$\gamma_\alpha^- = \{x_i \in \Gamma : i_\alpha = 0;\ i_\beta \neq 0,\ N_\beta \text{ for } \beta \neq \alpha\},$$

$$\gamma_{\alpha\beta}^+ = \{x_i \in \Gamma : i_\alpha = N_\alpha,\ i_\beta = N_\beta;\ i_\delta \neq 0,\ N_\delta \text{ for } \delta \neq \alpha, \beta\},$$

$$\omega_h = \bar{\omega}_h - \gamma, \qquad \omega_h^{+\alpha} = \omega_h + \gamma_\alpha^+, \qquad \omega_h^{+\alpha+\beta} = \omega_h^{+\alpha} + \gamma_\beta^+ + \gamma_{\alpha\beta}^+$$

$$\gamma_\alpha = \gamma_\alpha^+ + \gamma_\alpha^-.$$

Multiplying (7) scalarly by z and applying Green's difference formula (see [6]), we obtain the energy identity

$$I = \frac{1}{12} \sum_{\alpha=1}^{p} h_\alpha^2 \sum_{\beta \neq \alpha}^{1-p} \|z_{\bar{x}_\alpha \bar{x}_\beta}\|^2 + (\psi, z), \qquad I \equiv \sum_{\alpha=1}^{p} \|z_{\bar{x}_\alpha}\|^2. \tag{8}$$

Using Lemmas 2 and 3 of [7], we have

$$\|z\|^2 \leqslant \frac{l_\alpha^2}{4} \|z_{\bar{x}_\alpha}\|^2, \qquad \|z\|^2 \leqslant M_0 I, \qquad M_0 = \frac{1}{4} \left(\sum_{\alpha=1}^{p} \frac{1}{l_\alpha^2} \right)^{-1}$$

$$h_\alpha^2 \|z_{\bar{x}_\alpha \bar{x}_\beta}\|^2 \leqslant 4 \|z_{\bar{x}_\beta}\|^2, \qquad \frac{1}{12} \sum_{\alpha=1}^{p} h_\alpha^2 \sum_{\beta \neq \alpha}^{1\ p} \|z_{\bar{x}_\alpha \bar{x}_\beta}\|^2 \leqslant \frac{p-1}{3} I. \tag{9}$$

We consider (ψ, z):

$$(\psi, z) \leqslant \|\psi\| \|z\| \leqslant (M_0 I)^{1/2} \|\psi\| \leqslant c_0 I + \frac{M_0}{4c_0} \|\psi\|^2. \tag{10}$$

Substituting (9) and (10) in (8) and suitably fixing c_0, we find that

$$I \leqslant \frac{9M_0}{(4-p)^2} \|\psi\|^2,$$

or, from (9),

$$\|z\| \leqslant \frac{3M_0}{4-p} \|\psi\|.$$

We have now proved:

Theorem 1. If the condition

$$\|\psi\| \leqslant M|h|^4,$$

is satisfied, the difference scheme (5)-(6_1) with $p \leqslant 3$ is convergent in the mean at a rate $O(|h|^4)$ so that

$$\|y - u\| \leqslant M' |h|^4, \qquad M' = M \, \frac{3M_0}{4-p},$$

where M is a positive constant independent of $|h|$.

Theorem 1 proves the convergence of scheme (5)-(6_1) in the mean on any sequence of rectangles uniform with respect to each mesh direction, provided only that $|h| \to 0$. If we impose certain restrictions on the ratio between the intervals h_α of the mesh $\bar{\omega}_h$, we can prove uniform convergence for scheme (5)-(6_1). We expand (5)-(6_1) in points

$$\frac{7-p}{3} \sum_{\alpha=1}^{p} \frac{1}{h_\alpha^2} y = \frac{1}{6} \sum_{\alpha=1}^{p} \left[(7-p) \frac{1}{h_\alpha^2} - \sum_{\beta \neq \alpha}^{1-p} \frac{1}{h_\beta^2} \right] (y^{(+1_\alpha)} + y^{(-1_\alpha)}) + \\ + \frac{1}{12} \sum_{\alpha=1}^{p-1} \sum_{\beta=\alpha+1}^{p} \left(\frac{1}{h_\alpha^2} + \frac{1}{h_\beta^2} \right) (y^{(+1_\alpha, +1_\beta)} + y^{(+1_\alpha, -1_\beta)} + y^{(-1_\alpha, +1_\beta)} + y^{(-1_\alpha, -1_\beta)}) + \varphi. \tag{11}$$

It is clear from (11) that the coefficient of y on the left-hand side is equal to the sum of all the coefficients on the right-hand side. Let h_α, $\alpha = 1, \ldots, p$, be such that all the coefficients of (11) are non-negative, i.e.

$$(7-p) \frac{1}{h_\alpha^2} - \sum_{\beta \neq \alpha}^{1-p} \frac{1}{h_\beta^2} \geqslant 0. \tag{12}$$

The maximum principle (see [2]) will then hold for equation (11), and we can prove, by the same arguments as in [2]*, that

Theorem 2. If

$$\| \psi \|_0 = \max_{\omega_h} |\psi| \leqslant M|h|^4$$

and conditions (12) are satisfied, the difference scheme (5) - (6_1) with $p \leqslant 4$ is uniformly convergent at a rate $O(|h|^4)$ so that

$$\| y - u \|_0 \leqslant M'|h|^4,$$

where M' is a positive constant independent of $|n|$.

For $p = 2$, conditions (12) become $1/\sqrt{5} \leqslant h_1/h_2 \leqslant \sqrt{5}$. For $p = 3$, the ratios h_1^2/h_2^2 and h_1^2/h_3^2, which satisfy conditions (12) are given by the coordinates of the part of the plane inside the triangle with vertices $A(\frac{1}{3}\ \frac{1}{3})$, $B(1,\ 3)$, $C(3,\ 1)$. If $p = 4$, it follows from (12) that Theorem 2 only holds with $h_\alpha = h$ ($\alpha = 1, \ldots, 4$), which is only possible if the sides l_α, $\alpha = 1, \ldots, 4$ of the region D_p are commensurable.

Note 1. If D_p has commensurable sides l_α, we can introduce the difference scheme $\bar{\omega}_h$ with $h_\alpha = h$, $\alpha = 1, \ldots, p$, into it. On this mesh, problem (1) - (2) can be associated with the difference scheme

$$\Lambda^* y + \varphi^* = 0, \qquad y|_\gamma = g(x), \tag{5*}$$

where

$$\begin{aligned} \Lambda^* &= \Lambda + \frac{h^2}{6} \sum_{\alpha=1}^{p-1} \sum_{\beta=\alpha+1}^{p} \Lambda_\alpha \Lambda_\beta + \frac{h^4}{30} \sum_{\alpha=1}^{p-2} \sum_{\beta=\alpha+1}^{p-1} \sum_{\delta=\beta+1}^{p} \Lambda_\alpha \Lambda_\beta \Lambda_\delta, \\ \varphi^* &= f + \frac{h^2}{12} Lf + \frac{h^4}{360} \left(L^2 f + 2 \sum_{\alpha=1}^{p-1} \sum_{\beta=\alpha+1}^{p} L_\alpha L_\beta f \right). \end{aligned} \tag{6*}$$

Scheme (5*) - (6*) with $p = 2,\ 3$ was proposed in [3], [9], where its uniform convergence at a rate $O(h^6)$ was proved.

3. We consider the alternating directions method for the approximate

---

* See [8] for the case $p = 2$.

solution of problem (5) - (6_θ). Let $v = v^{(n+1)}$ be the $(n+1)$-th iteration, $\check{v} = v^{(n)}$, $\tau = \tau_n$ the iterational parameter, which will be chosen later, and $v_{\bar{t}} = (v - \check{v})/\tau$. For the derivative scheme (see [4]), connecting v and $\check{v}$, we take

$$A v_{\bar{t}} = \Lambda' \check{v} + \varphi, \qquad v|_\gamma = g(x), \qquad v^{(0)}(x) = v_0(x), \tag{13}$$

where

$$A = \prod_{\alpha=1}^{p} A_\alpha, \qquad A_\alpha = E - \sigma\tau\Lambda_\alpha, \qquad Ev = v, \qquad \sigma > \frac{1}{2},$$

and $v_0(x)$ is the zero approximation.

For $\theta = 0$, scheme (13) was proposed and investigated in [10] (see also [11] - [13]). One of the alternating directions algorithms (see [4], [7], [10] - [13]) may be used for determining v from (13); some of these only operate for $\theta = 0$. Let us prove, for instance, the algorithm proposed in [7]

$$\begin{gathered} A_1 w_{(1)} = \Lambda' \check{v} + \varphi, \qquad \Lambda_\alpha w_{(\alpha)} = w_{(\alpha-1)}, \qquad \alpha = 2, \ldots, p, \\ w_{(\alpha)} = 0 \text{ for } x \in \gamma_\alpha, \qquad \alpha = 1, \ldots, p, \qquad v = \check{v} + \tau w_{(p)}. \end{gathered} \tag{14}$$

Notice that the algorithm proposed in [12] follows from (14) on carrying out the substitution $w_{(\alpha)} = (v_{(\alpha)} - \check{v})/\tau$ in (14); however, (14) is more economic, and in addition, the $w_{(\alpha)}$, $\alpha = 1, \ldots, p$, always satisfy the zero boundary conditions.

4. Let us consider the convergence of the iterational process (13).

We obtain for $w = v - y$, where y is a solution of problem (5) - (6_θ), and v a solution of problem (13)

$$A w_{\bar{t}} = \Lambda' \check{w}, \qquad w|_\gamma = 0, \qquad w^{(0)}(x) = v_0(x) - y(x). \tag{15}$$

We apply Fourier's method for finding the solution of problem (15). Let $\mu_\alpha = \mu_{k_\alpha}(x)$ and $\lambda_\alpha = \lambda_{k_\alpha}$, $k_\alpha = 1, \ldots, N_\alpha - 1$, $\alpha = 1, \ldots, p$, be the eigenfunctions and eigenvalues of the one-dimensional Sturm-Liouville difference problem

$$\Lambda_\alpha \mu_\alpha + \lambda_\alpha \mu_\alpha = 0, \qquad \mu_\alpha(0) = \mu_\alpha(l_\alpha) = 0. \tag{16}$$

The problem

$$\Lambda\mu + \lambda\mu = 0, \qquad \mu|_\gamma = 0 \tag{17}$$

now has the solution

$$\mu = \mu_k(x) = \prod_{\alpha=1}^{p} \mu_{k_\alpha}(x_\alpha), \qquad \lambda = \lambda_k = \sum_{\alpha=1}^{p} \lambda_{k_\alpha}, \qquad k = (k_1, \ldots, k_p).$$

The eigenvalues of problem (16) are easily obtained:

$$\lambda_{k_\alpha} = \frac{4}{h_\alpha^2} \sin^2 \frac{k_\alpha \pi h_\alpha}{2l_\alpha}, \qquad k_\alpha = 1, \ldots, N_\alpha - 1, \tag{18}$$

but we shall only require the maximum and minimum of them in what follows.

We shall seek the solution of problem (16) in the form

$$w = w^{(n+1)} = \sum_k a_{k,n+1}\mu_k(x), \qquad \check{w} = \sum_k a_{k,n}\mu_k(x). \tag{19}$$

Substituting (19) in (15) and recalling that the functions $\mu_h(x)$ are orthogonal, we get

$$a_{k,\,n+1} = \rho_{k,\,n+1} a_{k,\,n}, \tag{20}$$

where

$$\rho_{k,n+1} = 1 - \tau\left[\lambda - \frac{\theta}{12} \sum_{\alpha=1}^{p} h_\alpha^2 \sum_{\beta \neq \alpha}^{1-p} \lambda_\alpha \lambda_\beta\right] \prod_{\alpha=1}^{p} (1 + \sigma\tau\lambda_\alpha)^{-1}. \tag{21_θ}$$

We obtain from (20)

$$a_{k,n+1} = a_{k,0} \prod_{s=1}^{n+1} \rho_{k,s},$$

and hence, from (19)

$$\begin{gathered} w^{(n+1)} = \sum_k a_{k,0} \prod_{s=1}^{n+1} \rho_{k,s}\mu_k, \\ \|w^{(n+1)}\| = \left(\sum_k \left[a_{k,0} \prod_{s=1}^{n+1} \rho_{k,s}\mu_k\right]^2, 1\right)^{1/2} \leqslant R_{n+1} \|w^{(0)}\|, \end{gathered} \tag{22}$$

where

$$R_{n+1} = \max_k \prod_{s=1}^{n+1} \rho_{k,s}. \tag{23}$$

Theorem 3. When conditions (4) are satisfied, the iterational process (13) with $p = 2, 3$ is convergent in the mean whatever the parameters τ_h satisfying

$$0 < c_1 \leqslant \tau_n \leqslant c_2, \tag{24}$$

where c_1 and c_2 are constants independent of n.

From (22), to prove the theorem we have to show that $R_{n+1} \to 0$ as $n \to \infty$. But to do this, it is sufficient to show that

$$|\rho_{k,s}| < \rho < 1, \tag{25}$$

where ρ is a constant independent of n, since we then have from (23) $R_{n+1} \leqslant \rho^{n+1}$. By (18),

$$h_\alpha^2 \lambda_\alpha \lambda_\beta < 4\lambda_\beta \text{ and } \frac{1}{12} \sum_{\alpha=1}^{p} h_\alpha^2 \sum_{\beta \neq \alpha}^{1-p} \lambda_\alpha \lambda_\beta < \frac{p-1}{3} \sum_{\alpha=1}^{p} \lambda_\alpha = \frac{p-1}{3} \lambda.$$

It follows from this, and (21_θ), that

$$\rho_{k,s} < 1 - \left(1 + \theta \frac{1-p}{3}\right) \tau\lambda \prod_{\alpha=1}^{p} (1 + \sigma\tau\lambda_\alpha)^{-1}, \tag{26_θ}$$

$$\rho_{k,s} > 1 - \tau\lambda \prod_{\alpha=1}^{p} (1 + \sigma\tau\lambda_\alpha)^{-1} > 1 - \frac{\tau\lambda}{1 + \sigma\tau\lambda}. \tag{27}$$

On now using condition (24), we find that

$$|\rho_{k,s}| < \rho,$$

where

$$\rho = \max \left\{ \left| 1 - \left(1 + \theta \frac{1-p}{3}\right) c_1 \lambda \prod_{\alpha=1}^{p} (1 + \sigma c_2 \lambda_\alpha)^{-1} \right|, \ \left| 1 - \frac{c_2\lambda}{1 + \sigma c_2 \lambda} \right| \right\},$$

i.e. ρ is independent of n. Recalling (4), we find that in fact $\rho < 1$. (Theorem 3 was proved for $\theta = 0$ in [10].)

Note 2. The iterational scheme for problem (5*) - (6*) is

$$A v_{\bar{t}} = \Lambda^* \check{v} + \varphi^*, \qquad v|_\gamma = g(x), \qquad v^{(0)}(x) = v_0(x), \tag{13*}$$

while the corresponding function is

$$\rho^*_{k,n+1} = 1 - \tau \left[\lambda - \frac{h^2}{6} \sum_{\alpha=1}^{p-1} \sum_{\beta=\alpha+1}^{p} \lambda_\alpha \lambda_\beta + \frac{h^4}{30} \sum_{\alpha=1}^{p-2} \sum_{\beta=\alpha+1}^{p-1} \sum_{\gamma=\beta+1}^{p} \lambda_\alpha \lambda_\beta \lambda_\gamma \right] \prod_{\alpha=1}^{p} (1 + \sigma\tau\lambda_\alpha)^{-1}; \tag{21*}$$

Recalling (18), it can easily be seen from (21*) that, with (4) and (24), the upper bound for the function $\rho^*_{k,n+1}$ is of the same form as for the function $\rho_{k,n+1}$ with $\theta = 1$. Theorem 1 therefore holds for scheme (13*) also.

5. To estimate the rate of convergence (number of iterations) of the iterational process (13), we require a more exact upper bound for $|\rho_{k,\,n+1}|$.

Lemma 1. For the function $\rho_{k,\,n+1}$, defined by (21_θ) with

$$\sigma \geqslant \sigma_{p,\theta}, \qquad \sigma_{2,1} = \tfrac{5}{9}, \qquad \sigma_{3,1} = \tfrac{1}{2}, \qquad \sigma_{p,0} = \tfrac{1}{2}\left[1 + \left(\frac{p-1}{p}\right)^{p-1}\right] \tag{28}$$

we have

$$|\rho_{k,\,n+1}| < \bar{\rho}(a), \tag{29}$$

where

$$0 < \bar{\rho}(a) = 1 - \frac{1}{\sigma}\left(1 + \theta\frac{1-p}{3}\right)\frac{pa}{(1+a)^p}, \qquad a = \frac{\sigma\tau\lambda}{p} > 0. \tag{30}$$

For, by the theorem on the arithmetic mean and geometric mean (see [14], p. 29), we have

$$\prod_{\alpha=1}^{p}(1 + \sigma\tau\lambda_\alpha) \leqslant \left(1 + \frac{\sigma\tau\lambda}{p}\right)^p.$$

We find from this and (26_θ) that

$$\rho_{k,n+1} < 1 - \left(1 + \theta\frac{1-p}{3}\right)\frac{\tau\lambda}{\left(1 + \frac{\sigma\tau\lambda}{p}\right)^p} = \bar{\rho}(a)$$

whatever the positive σ_p.

It follows from (27) that, to complete the proof of Lemma 1, we have to show that

$$-1 + \left(1 + \theta\frac{1-p}{3}\right)\frac{\tau\lambda}{\left(1 + \frac{\sigma\tau\lambda}{p}\right)^p} < 1 - \frac{\tau\lambda}{1 + \sigma\tau\lambda}. \tag{31}$$

To this end, we consider the function

$$F_{p,\theta} = 2\sigma - \frac{pa}{1+pa} - \left(1 + \theta\frac{1-p}{3}\right)\frac{pa}{(1+a)^p},$$

the fact that this is positive being equivalent to (31). We transform $F_{p,1}$ to the form

$$F_{2,1} = 2\sigma - \frac{10}{9} + \frac{2}{9}\,\frac{a^3 - 5a^2 + 5a + 5}{(a+1)^2(2a+1)}, \qquad F_{3,1} = 2\sigma - 1 + \frac{a^3 + 2a + 1}{(a+1)^3(3a+1)}.$$

Given (28), the fact that $F_{3,1}$ is positive is now obvious. Given (28), the fact that $F_{2,1}$ is positive is equivalent to the numerator being positive, and this can easily be proved by considering its minimum. When investigating $F_{p,0}$, we shall be satisfied with a crude estimate. In fact, we shall estimate separately $pa/(1+a)^p$ and $pa/(1+pa)$. Now,

$$F_{p,0} > 2\sigma - 1 - \left(\frac{p-1}{p}\right)^{p-1},$$

and the lemma follows from this and (28).

The expression involving $\rho_{k,\,n+1}$, established by Lemma 1 holds under strcnger restrictions on σ (except for the case $p = 3$, $\theta = 1$) then does Theorem 3. For $p = 2$ it is possible to obtain an estimate rather different from (29) for $\rho_{k,\,n+1}$, which holds for $\sigma \geqslant ½$.

Lemma 2. Given the function

$$\rho(a_1, a_2) = 1 - \varkappa\frac{a_1 + a_2 - \alpha a_1 a_2}{(1+a_1)(1+a_2)}, \qquad \varkappa \geqslant 0, \quad \alpha \geqslant 0, \quad a_\alpha > 0.$$

If the condition

$$\varkappa \leqslant 2, \tag{32}$$

is satisfied, we have

$$\rho(a_\alpha, a_\alpha) \geqslant 0, \qquad \rho^2(a_1, a_2) \leqslant \rho(a_1, a_1)\rho(a_2, a_2). \tag{33}$$

The first inequality may be proved immediately

$$\rho(a_\alpha, a_\alpha) \geqslant \frac{1 - 2(\varkappa - 1)a_\alpha + a_\alpha^2}{(1+a_\alpha)^2} \geqslant 0 \quad \text{for } \varkappa \leqslant 2.$$

We consider the difference

$$\rho(a_1, a_1)\,\rho(a_2, a_2) - \rho^2(a_1, a_2) = \frac{J}{(1+a_1)^2(1+a_2)^2},$$

where

$$J = [(1+a_1)^2 + \varkappa\alpha a_1^2 - 2\varkappa a_1][(1+a_2)^2 + \varkappa\alpha a_2^2 - 2\varkappa a_2] -$$

$$-\left[(1+a_1)(1+a_2)+\varkappa\alpha a_1 a_2-\varkappa(a_1+a_2)\right]^2.$$

Removing the brackets and collecting like terms, we get

$$J=\varkappa(2-\varkappa+\alpha)(a_1-a_2)^2.$$

This leads us to (33), provided (32) is satisfied.

A fairly simple corollary of Lemma 2 is

Lemma 3. If $p = 2$, we have for the function $\rho_{k,n+1}$ defined by (21_θ), provided condition (4) is satisfied

$$(\rho_{k,n+1})^2 \leqslant \prod_{\alpha=1}^{2} \bar{\rho}(a_\alpha), \tag{34}$$

where

$$\bar{\rho}(a_\alpha)=1-\frac{1}{\sigma}\left(1-\frac{\theta}{3}\right)\frac{2a_\alpha}{(1+a_\alpha)^2}, \qquad a_\alpha=\sigma\tau\lambda_\alpha. \tag{35}$$

For, it follows from Lemma 2 that

$$(\rho_{k,n+1})^2 \leqslant \prod_{\alpha=1}^{2}\rho_{k_\alpha,n+1}, \qquad \rho_{k_\alpha,n+1}=1-\frac{2\tau\lambda_\alpha-\theta\frac{h_\alpha^2}{6}\tau\lambda_\alpha^2}{(1+\sigma\tau\lambda_\alpha)^2},$$

since $\rho_{k_\alpha,n+1}>0$ for $\sigma>\frac{1}{2}$. But we have, by (18),

$$\frac{h^2}{6}\tau\lambda_\alpha^2 \leqslant \frac{2}{3}\tau\lambda_\alpha \text{ and } \rho_{k_\alpha,n+1}\leqslant\bar{\rho}(a_\alpha).$$

(This lemma was proved in [12] for $\theta = 0$ and $\sigma = 1$.)

Finally, we require

Lemma 4. Given

$$0<m<M. \tag{36}$$

The maximum of the function $\bar{\rho}(a)$ defined by (30) and (35) in the interval $[m, M]$ is now equal to

$$\begin{aligned}\rho_p&=\max_{m\leqslant a\leqslant M}\bar{\rho}(a)=\\&=\max\left[1-\frac{1}{\sigma}\left(1+\theta\frac{1-p}{3}\right)\frac{pm}{(1+m)^p},\ 1-\frac{1}{\sigma}\left(1+\theta\frac{1-p}{3}\right)\frac{pM}{(1+M)^p}\right].\end{aligned} \tag{37}$$

For, it follows from

$$\bar{\rho}'(a) = \frac{1}{\sigma}\left(1 + \theta\frac{1-p}{3}\right) p \frac{(p-1)a-1}{(1+a)^{p+1}} = \begin{cases} \leqslant 0 \text{ for } a \leqslant \frac{1}{p-1}, \\ \geqslant 0 \text{ for } a \geqslant \frac{1}{p-1} \end{cases}$$

that $\bar{\rho}(a)$ takes its maximum value at either the left- or the right-hand end of the interval $[m, M]$.

6. We shall now estimate the rate of convergence of the iterational process (13). To be more precise, we shall find a sequence of iterational parameters $\{\tau_n\}$ such that a "reasonably high" rate of convergence is obtained. It follows from Theorem 3 that the parameter τ_n may vary within fairly wide limits. We shall therefore try to find a sequence $\{\tau_n\}$ such that, given any value of λ, there is at least one value of τ such that $|\rho_{k,\,n+1}| < \rho < 1$, where ρ is independent of both n and $|h|$. If we then perform the cycle of iterations (13) with the given system of parameters, we shall obtain, in view of (22) - (23), a ρ^{-1} times reduction in the norm of the error. It is desirable for the total number of parameters in the sequence $\{\tau_n\}$ to be "not very great" (obviously, in the worst case we can avoid a number of parameters equal to the number of distinct eigenvalues λ), i.e. for one parameter τ to be "stipulated" by a whole series of eigenvalues and not just one. In fact, let the sequence of intervals $(\xi_{(n-1)}, \xi_{(n)})$, $n = 1, \ldots, n_0$, cover the interval $[\lambda_1, \lambda_{N-1}]$, where

$$\xi_{(0)} = \lambda_1, \quad \xi_{(n_0-1)} < \lambda_{N-1}, \quad \xi_{(n_0)} \geqslant \lambda_{N-1}, \tag{38}$$

the coordinates $\xi_{(n)}$ and the number n_0 being subject to definition. Let τ_n "stipulate" the λ_k which satisfy

$$\xi_{(n-1)} \leqslant \lambda_k \leqslant \xi_{(n)}, \tag{39}$$

i.e. for the k given by (39), the functions $\rho_{k,\,n+1}$ satisfy (25) with a ρ independent of either n or $|h|$. This means in our case, by Lemmas 1 and 4, that

$$pm \leqslant \tau_n\sigma\xi_{(n-1)} \leqslant \tau_n\sigma\lambda_k \leqslant \tau_n\sigma\xi_{(n)} \leqslant pM,$$

where $m < M$ are positive constants independent of either n or $|h|$. If m and M are chosen, let

$$pm = \sigma\tau_n\xi_{(n-1)}, \qquad pM = \sigma\tau_n\xi_{(n)}. \tag{40}$$

It now follows from this and (38) that

$$\xi_{(n)} = q^{-n}\lambda_1, \qquad \tau_n = \frac{pm}{\sigma}\lambda_1^{-1}q^{n-1}, \qquad q = \frac{m}{M}, \tag{41}$$

$$\lg\frac{\lambda_1}{\lambda_{N-1}}\lg^{-1}q \leqslant n_0 < \lg\frac{\lambda_1}{\lambda_{N-1}}\lg^{-1}q + 1. \tag{42}$$

Using Lemma 4, we arrive from (41), (42) and (22) - (23) at

Lemma 5. If a cycle of n_0 iterations is carried out in accordance with method (13) with a system of parameters $\{\tau_n\}$ given by (41), then, if conditions (28) are satisfied,

$$\| z^{(n_0)} \| \leqslant \rho_p \| z^{(0)} \|, \tag{43}$$

where ρ_p is given by (37).

A simple consequence of Lemma 5 is

Theorem 4. In order to reduce the norm L_2 of the error $\|z^{(0)}\|$ by a factor $1/\varepsilon$ with the aid of method (13), it is sufficient, if conditions (28) are satisfied, to perform a cycle of n_0 iterations with the system of parameters $\{\tau_n\}$ given by (41) k_0 times, where n_0 is given by (42), and k_0 by

$$k_0 \geqslant \lg \varepsilon \lg^{-1} \rho_p. \tag{44}$$

The following asymptotic formula holds here for the total number of iterations $\nu = n_0 k_0$:

$$\nu \asymp \nu_0 \lg\frac{\lambda_1}{\lambda_{N-1}} \lg \varepsilon, \qquad \nu_0 = \frac{1}{\lg q \lg \rho_p}. \tag{45}$$

Note 3. We have by (18):

$$\frac{\lambda_1}{\lambda_{N-1}} = \frac{\sum\limits_{\alpha=1}^{p} \frac{1}{h_\alpha^2}\sin^2\frac{\pi h_\alpha}{2l_\alpha}}{\sum\limits_{\alpha=1}^{p} \frac{1}{h_\alpha^2}\cos^2\frac{\pi h_\alpha}{2l_\alpha}}.$$

If $\bar{D}$ is the p-dimensional cube with side l and the mesh ω_h is square, i.e. $h_\alpha = h$, $\alpha = 1, \ldots, p$, then

$$\frac{\lambda_1}{\lambda_{N-1}} = \mathrm{tg}^2\frac{\pi h}{2l} = O(h^2) \text{ and } \lg\frac{\lambda_1}{\lambda_{N-1}} = O(\lg h).$$

The constructions used in the proofs of Lemma 5 and Theorem 4 are based on Lemma 1 and therefore hold only if conditions (28) are satisfied.

TABLE 1

p	θ	σ	ν	m	q	ρ
2	1	$\frac{1}{2}$	3.425	0.277	0.0766	0.547
		$\frac{5}{9}$	3.939	0.296	0.0710	0.601
	0	$\frac{1}{2}$	1.707	0.415	0.1719	0.171
		$\frac{3}{4}$	3.425	0.277	0.0766	0.547
3	1	$\frac{1}{2}$	10.577	0.135	0.0850	0.815
	0	$\frac{13}{18}$	4.432	0.153	0.1070	0.586

TABLE 2

p	θ	σ	ν	m	q	ρ
2	1	1	7.958	0.237	0.0563	0.793
	0	1	4.957	0.254	0.0645	0.677
3	1	1	22.258	0.128	0.0786	0.911
	0	1	6.648	0.142	0.0938	0.714

We can prove with $p = 2$, from Lemma 3, and by analogy with [11]:

Theorem 5. In the case $p = 2$, in order to reduce the norm L_2 of the error $\|z^{(0)}\|$ by a factor $1/\varepsilon$ with the aid of method (13), it is sufficient, given any $\sigma \geqslant 0.5$, to carry out a cycle of n_0 iterations with the system of parameters

$$\tau_n = \frac{m}{\sigma} c_* q^{n-1} \tag{46}$$

k_0 times, where k_0 is given by (44), while

$$\lg \frac{c^*}{c_*} \lg^{-1} \frac{1}{q} \leqslant n_0 < \lg \frac{c^*}{c_*} \lg^{-1} \frac{1}{q} + 1 \tag{47}$$

and

$$c_* = \min_{k_\alpha} \lambda_{k_\alpha}, \qquad c^* = \max_{k_\alpha} \lambda_{k_\alpha}.$$

The following asymptotic formula holds here for the total number of iterations $\nu = n_0 k_0$:

$$\nu = \nu_0 \lg \frac{c^*}{c_*} \lg \frac{1}{\varepsilon}, \qquad \nu_0 = \frac{1}{\lg q \lg \rho_p}. \tag{48}$$

Notice that, in a square region and on a square mesh, (46) is the same as (41), (47) as (42) and (48) as (45).

Note 4. Using Note 1, it is easily shown that Theorem 4 also holds for the iterational scheme (13*).

We now consider the minimization of the coefficient ν_0. Using (37) and (41), it is clear from (45) and (48) that, with θ fixed, ν_0 is a function of the three variables m, M and σ. Since q and ρ_p are always less than unity, ν_0 will decrease with q and ρ_p. Hence, if ρ_p is fixed, ν will be a minimum if q is a minimum. But it follows from (41) that q is a minimum if the first and second terms on the right-hand side of (37) are the same, i.e. $m / (1+m)^p = M / (1+M)^p$. Hence

$$M_2 = \frac{1}{m}, \quad M_3 = \frac{\sqrt{(3+m)^2 + 4/m} - (3+m)}{2}.$$

It is clear from (37) that ρ_p is an increasing function with respect to σ. Hence, for ν_0 to be as small as possible, σ must also be a minimum. After M and σ have been fixed, ν_0 remains a function of m only and its minimum can be found numerically to any degree of accuracy. Table 1 gives the numerical values of the parameters occurring in ν_0, optimum with respect to m for minimum σ.

Table 2 gives for comparison the same parameters for $\sigma = 1$. (The values of the parameters were obtained in [12] for $p = 2$ and $\theta = 0$, and in [4] for $p = 2$, 3 and $\theta = 1$. The numbers quoted there correspond to natural logarithms in (45) and (47), whereas we use logarithms to base 10.)

Translated by D.E. Brown

REFERENCES

1. MIKELADZE, Sh.E., Numerical integration of partial differential equations, *Izv. Akad. Nauk SSSR, Ser. Mat.*, No. 6, 819-842, 1934.

2. MIKELADZE, Sh.E., *Numerical methods of integrating partial differential equations* (Chislennye metody integrirovaniya differentsial'nykh uravnenii s chastnymi proizvodnymi). Moscow, Akad. Nauk SSSR, 1936.

3. MIKELADZE, Sh.E., Numerical solution of the differential equation $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 + \partial^2 u/\partial z^2 = f(x, y, z)$. *Dokl. Akad. Nauk SSSR*, 14, No. 4, 177-180, 1937.

4. SAMARSKII, A.A. and ANDREEV, V.B., Difference scheme of high accuracy for an elliptic equation with several spatial variables. *Zh. Vych. Mat.* 3, No. 6, 1006-1013, 1963.

5. GREENSPAN, D., On "best" nine-point Laplace difference analogues on rectangular grids, *Quart. J. Mech. Appl. Math.*, 12, No. 1, 111 - 116, 1959.

6. SAMARSKII, A.A., Locally uniform difference schemes on nonuniform meshes. *Zh. Vych. Mat.* 3, No. 3, 431-466, 1963.

7. SAMARSKII, A.A., Schemes of improved accuracy for the multi-dimensional equation of heat conduction, *Zh. Vych. Mat.* 3, No. 5, 812-840, 1963.

8. BADAGADZE, V.V., Numerical solution of second order elliptic equations, *Soobshch. Akad. Nauk Gruz SSR*, 30, No. 6, 689-696, 1963.

9. MIKELADZE, Sh.E., Numerical solution of Laplace and Poisson equations, *Izv. Akad. Nauk SSSR, Ser. Mat.*, No. 2, 271-292, 1938.

10. D'YAKONOV, E.G., Solution of some multi-dimensional problems of mathematical physics by means of the mesh method, Diss. kand. fiz-matem. n., Matem. in-t Akad. Nauk SSSR, Moscow, 1962.

11. PEACEMAN, D.W. and RACHFORD, H.H., The numerical solution of parabolic and elliptic differential equations, *J. Soc. Industr. Appl. Math.*, 3, No. 1, 28-41, 1955.

12. DOUGLAS, H. and RACHFORD, H.H., On the numerical solution of heat conduction problems in two and three space variables, *Trans. Amer. Math. Soc.*, 82, No. 2, 421-439, 1956.

13. DOUGLAS, J., Alternating direction methods for three space variables, *Numer. Math.*, 4, No. 1, 41-63, 1962.

14. HARDY, G.H., LITTLEWOOD, D. and POLYA, G., *Inequalities*, Cambridge Univ. Press, 1934.