ON THE CONVERGENCE OF THE FRACTIONAL STEP METHOD FOR HEAT CONDUCTIVITY EQUATIONS*

A. A. SAMARSKII
(Moscow)
(Received 9 June 1962)

1. The splitting method (fractional step method) is suggested in [1] for the numerical solution of the multi-dimensional heat conductivity equation with constant coefficients. In [2] the difference scheme for the heat conductivity equation with constant coefficients is written down for the case of two space variables and a rectangular region, and a proof of the convergence (in the mean) of this scheme with speed $O(\tau^2)$ with the additional condition $\tau/h^2 = \text{const.}$ is given. In this note we shall show that the modified fractional step method can also be used in the case of an arbitrary region and a non-uniform difference net. We prove the convergence in the mean with speed $O(h^2) + O(\tau)$ of the local one-dimensional schemes we examine; in contrast to [1], we shall not connect these schemes with the splitting of the multi-dimensional difference equation and this makes them suitable for an arbitrary region and quasi-linear parabolic equations. In contrast to [3], at each stage we use six-point homogeneous schemes which enables us, in particular, to take the scheme of [2] into account. As usual we give the argument for a family of homogeneous schemes characterised by the given pattern functionals of [3]-[5]. We use the methods developed in [3], [5] and [6]. Convergence is proved for arbitrary non-uniform nets.

2. Let $\mathcal{G}$ be an arbitrary two-dimensional region bounded by the contour $\Gamma$, $\partial = \mathcal{G} + \Gamma$, $x = (x_1, x_2)$ a point with the coordinates $x_1$ and $x_2$. In the cylinder $\bar{Q}_T = \mathcal{G} \times [0 \leq t \leq T]$ we look for the solution of the problem

$$\frac{\partial u}{\partial t} = \sum_{x=1}^{2} L_x u, \quad L_x u = \frac{\partial}{\partial x_2} \left( k_x(x, t) \frac{\partial u}{\partial x_2} \right), \quad (x, t) \in Q_T = \mathcal{G} \times (0 \leq t \leq T),$$

(1)
We shall assume: (1) the problem (1)-(3) has a unique solution $u = u(x, t)$, continuous in $\Omega_T$ (2) the following conditions are satisfied.

**Conditions A.** In the region $\Omega_T$ the functions $\partial^2 u / \partial x^2$, $\partial^2 k / \partial x^2$ satisfy the Lipschitz conditions for $x_{\alpha}$ and $x_{\alpha}$, $\partial u / \partial x_\alpha$, $\partial^2 u / \partial x^2_\alpha$, $\partial u / \partial t$, $k_{\alpha}$, $\partial k / \partial x_\alpha$ - the Lipschitz conditions for $t$ ($\alpha = 1, 2$).

Below we use only one property of the region $\Omega$ explicitly: the intersection with the region $\Omega$ of the straight line $L_\alpha$, drawn through any point $x \in \Omega$ parallel to the coordinate axis $Ox_\alpha$ consists of a finite number of intervals (see [3]). For simplicity we give our argument assuming that any straight line $L_\alpha$ intersects the contour $\Gamma$ twice.

3. Consider the arbitrary non-uniform rectangular net $\overline{\omega}_h = \{x_i \in \overline{\Omega}\}$, where $x_i = (x^{(i)}_1, x^{(i)}_2)$ is the node of the net with coordinates $x^{(i)}_1$ and $x^{(i)}_2$, $i_\alpha = 0, \pm 1, \pm 2, \ldots, \alpha = 1, 2$. The steps of the net $h^{(i)}_1 = x^{(i)}_1 - x^{(i-1)}_1$ and $h^{(i)}_2 = x^{(i)}_2 - x^{(i-1)}_2$ are functions of the coordinates $x^{(i)}_1$ and $x^{(i)}_2$ respectively. Let $\omega_h = \{x_i \in \Omega\}$ be the internal net region, $\gamma = \{x_i \in \Gamma\}$ the boundary net region. Through the point $x_i \in \omega_h$ we draw the straight line $L_\alpha$ parallel to the coordinate axis $Ox_\alpha$; the set of all nodes of the net $\overline{\omega}_h$ lying on $L_\alpha$, is called the chain $\mathcal{U}_\alpha$. Let $x_\alpha = x^{(a)}_\alpha \in \gamma$ and $x_R = x^{(b)}_\alpha \in \gamma$ be boundary points of the chain $\mathcal{U}_\alpha$, where $x_\alpha$ increases on going from $x_1$ to $x_r$, and $\gamma_\alpha = \{x_i \in \gamma\}$ is the set of the nodes $x_i$ of all chains $\mathcal{U}_\alpha$ in a given direction, $\gamma_\alpha = \{x_i \in \gamma\}$ is the set of nodes $x_i$. We put $\gamma_\alpha = \gamma_\alpha = \gamma_\alpha + \gamma_\alpha$, $\omega^{(a)} = \omega^{(a)} + \gamma_\alpha$, $\omega^{(a)} = \omega^{(a)} + \gamma_\alpha$. The segment $0 \leq \tau \leq T$ is divided into $K$ equal parts of length $\tau$ by the points $t_0 = 0$, $t_1 = \tau$, $t_K = T$, $\tau = T/K$ (uniform net with step $\tau$) and we introduce the intermediate (fractional) steps $t_{j + \tau/2} = t_j + \tau/2 = (j + \frac{1}{2}) \tau$, $j = 0, 1, \ldots, K - 1$. The net $\overline{\omega}_h$ with respect to time contains both integral and fractional steps, i.e. $\overline{\omega}_h = \{t_j \in [0 \leq t \leq T]\}$, where $t_j = j' \tau$, $j' = 0, \frac{1}{2}, 1, \ldots, l, l + \frac{1}{2}, \ldots, K - \frac{1}{2}, K$. The point $(x_i, t_j) \in \overline{Q}_T$ is a node of the space-time net $\overline{Q}$ and $\Omega = \{x_i, t_j\} \in \overline{Q}_T$ is the set of internal nodes of the net $\overline{Q}$; it is clear that $\overline{\omega}_h = \overline{\omega}_h \times \overline{\omega}_h$, $\Omega = \omega_h \times \omega_T$, where
\( \omega_{\tau} = \{ t_{j} \in (0 < t \leqslant T) \}. \)

We shall use the notation given in [3]. Let \( y = y(x, t_{j}) = y^{i*} \) be a net function given on \( \Omega. \) We shall write \( y(x \pm n \delta, t_{j}) = y(x \pm n \delta, t_{j}), \)

\[ y_{-} = \frac{1}{2} (y_{1} + y_{1-1}), \quad y_{+} = \frac{1}{2} (y_{1+1} + y_{1}). \]

\( y_{x} = \frac{1}{2} (y_{1}^{+} - y_{1}^{-})/h_{x}, \quad y_{x}^{2} = \frac{1}{2} (y_{1+1}^{+} + y_{1-1}^{-})/h_{x}^{+}. \)

We need the following sums and norms:

\[ (y, v) = \sum_{x, t} y_{x} v_{x}, \quad (y, v)_{\alpha} = \sum_{x, t} y_{x}^{(\alpha)} v_{x}, \]

\[ (y, v)_{\alpha} = \sum_{x, t} y_{x} v_{x}^{(\alpha)}, \quad \| y \| = \sqrt{(y, y)}, \quad \| y \|_{\alpha} = \sqrt{(y, y)_{\alpha}}, \]

\[ \| y \|_{3, \alpha} = \| y \|_{\alpha} \epsilon_{\alpha}, \quad \text{where } (\epsilon_{\alpha})_{x} = y, \]

\[ \| y \|_{3, \alpha} = \| y \|_{\alpha} \epsilon_{\alpha}, \quad \text{where } (\epsilon_{\alpha})_{x} = y, \]

\[ \eta_{x} = 0 \text{ for } x \in \Omega, \text{n}. \]

4. The operator \( L_{\alpha}u \) is replaced by the homogeneous uniform net scheme

\[ \Lambda_{\alpha}y = (a_{\alpha} y_{x}^{\alpha})^{*}, \]

having second order approximation on a uniform net (see [4] and [5]) so that \( \Lambda_{\alpha}u - L_{\alpha}u = O(h_{\alpha}^{2}). \)

The class of pattern functionals \( a_{\alpha} \) is described in [4]; in particular, for \( a_{\alpha} = k_{\alpha} (-h) \) we obtain the scheme used in [2]. We put the problem (1)-(3) in correspondence with the local one-dimensional scheme:

\[ y_{t_{j}} = \Lambda_{\alpha} y^{(o)} \text{ for } (x, t_{j+\alpha}) \in \Omega, \alpha = 1, 2; j = 0, 1, 2, \ldots, K - 1; \]

\[ y(x, t_{j+\alpha}) = u_{1}(x, t_{j+\alpha}), \quad y(x, 0) = u_{0}(x) \text{ for } x \in \Omega; \]

\[ (a_{\alpha}(x, t))^{*} > 0, \quad (x, t) \in \omega_{\alpha}^{(\alpha)} \times [0 < t < T], \]

where \( y^{(o)} = y + (1 - s) \hat{y}, \quad \hat{y} = y(x, t_{j+\alpha}), \quad \hat{y}^{(o)} = y(x, t_{j+\alpha}), \quad \hat{y}^{(o)} = y(x, t_{j+\alpha}), \)

\( y_{t_{j}} = (y - \hat{y})/\tau, \quad \Lambda_{\alpha}y = (a_{\alpha}(x, t)) y_{x}^{\alpha} \text{ for } t \in [t_{j}, t_{j+1}] \text{ and for definiteness we shall take } t' = t_{j+\alpha}. \) It is clear from (4)-(5) that at each moment \( t_{j+\alpha} \) we solve the first boundary problem for all chains \( H_{\alpha} \) in the given direction \( x_{\alpha}. \)
5. Let us now go on to study the convergence and accuracy of the scheme (4)-(6). Let \( u \) be the solution of problem (1)-(3), \( \gamma \) the solution of problem (4)-(6). For \( z = \gamma - u \) we obtain the conditions

\[
\begin{align*}
z(x, t_{j+1/2}) &= 0, \quad x \in \Gamma; \\
z(x, 0) &= 0, \quad x \in \Omega;
\end{align*}
\]

Suppose conditions A are satisfied. By analogy with (7) we find

\[
\Lambda u - L u - (\mu)_{x_{a}} + \Psi = O(\|h^2\|_{2a}), \quad \|\Psi\|_{2} = O(\|h^2\|_{2a})
\]

where \( \|h^2\|_{2a} = \sqrt{(1, h^4)}_{a} \) is the mean square value of \( h^2 \).

We put \( \omega_{L} \) in the form

\[
\omega_{L} = \omega_{L}^{0} + \omega_{L}^{*}, \quad \omega_{L}^{0} = \left( L u - \frac{1}{2} \frac{\partial u}{\partial t} \right), \quad \sum_{a=1}^{s} \omega_{L}^{0} = 0,
\]

\[
\omega_{L}^{*} = (\psi)_{x_{a}} + \psi, \quad \|\psi\|_{2} = O(\|h\|_{2a}) + O(\tau).
\]

6. With the estimate of the solution of the problem (7)-(9) we use the method of integral inequalities and a special method of summing the local errors \( \omega_{L} \), using the relation \( \omega_{L}^{0} + \omega_{L}^{*} = 0 \). Let us write down the basic integral identity with \( \sigma = \tau \). Let us multiply (7) by \( (z + \bar{z}) h \) and sum with respect to \( \omega_{h} \). Using Green's formula

\[
(z + \bar{z}, \Lambda z^{(1/2)}) = \frac{1}{2} V \bar{z}(z_{x_{a}} + z_{x_{a}})_{\omega_{a}} = \mathcal{I}
\]

we obtain the following identities:

\[
(\|z\|^2)_{L} - I_{1} + I_{2} = \sum_{a=1}^{s} (\psi_{a}, z + \bar{z})_*
\]

where

\[
(\|z\|^2)_{L} = (\|z^{i+1}\|^2 - \|z_{j-1}\|^2)_{\tau}.
\]

Putting \( \omega_{L}^{0} = 0 \) in (15) we see that the scheme (4)-(6) is stable with respect to the initial data (for any \( h_{\alpha} \) and \( \tau \):

\[
\|z(x, t_{j})\|_{\alpha} \ll \|z(x, 0)\|_{\alpha}, \quad (t_{j} = j \cdot \tau, \quad j = \frac{1}{2}, 1, \frac{3}{2}, \ldots, \quad K = \frac{1}{2}, \ldots, \quad K).
\]
7. Lemma 1. If $\psi_\alpha$ can be put in the form

$$\psi_\alpha = \psi_0^\alpha + (\mu_\alpha)_x^\alpha + \varphi_\alpha,$$

then for the solution of problem (7)-(9) for $\sigma = \frac{1}{2}$ and any $h_\alpha$ and $\tau$ we have the estimate

$$\|z(x, t_{j+1})\| \leq \tau \|\psi^0(x, t_{j+1})\| + \tau V^1 \|\psi(x, t_j)\| + \sqrt{1 + e} \sum_{j'=1}^{j+1} \tau Q^{j'},$$

where

$$\psi^0 = \psi_0^0 = -\psi_1^0,$$

$$\|\psi^0(x, t_j)\| = \left(\frac{1}{\tau} \sum_{j'=1}^j \tau \|\psi^0(x, t_{j'})\|^2\right)^{1/2},$$

$$Q^{j'} = \tau^2 \|\psi^0(x, t_{j'})\|^2 + \sum_{\alpha=1}^2 \left(\|\mu_\alpha(x, t_{j'})\|_{\alpha}^2 + \|\varphi(x, t_{j'})\|_{\alpha}^2\right).$$

We put (17) in (15) and make the following transformations:

$$\sum_{\alpha=1}^2 (\psi^0_\alpha, x + \hat{z})^* + (\psi^0, z^{j+1} - z^j) = \tau (\psi^0(x, t_{j+1}), z(x, t_{j+1}))_{\tau} - (\psi^0(x, t_{j+1}), z(x, t_j))_{\tau},$$

$$|((\mu_\alpha)_x\alpha, z + \hat{z})| = |(\mu_\alpha, x_{\alpha} + \hat{x})_\alpha| \leq \frac{1}{2} I_\alpha + \frac{1}{c_0} \|\mu_\alpha\|^2_{\alpha}.$$ 

Using the inequality $\tau |(\psi^0, z^j)^*| \leq \frac{\tau^2}{2c_0} \|\psi^0\|^2_{\tau}^2 + \frac{c_0}{2} \|z^j\|^2$, where $c_0$ is an arbitrary positive constant which we shall put equal to $1/\tau$ later. As a result we obtain the inequality

$$\|z(x, t_{j+1})\|^2 \leq (1 + 0.5c_0 \tau) \|z(x, t_j)\|^2 + 0.5rQ^{j+1} + \tau^2 (\psi^0(x, t_{j+1}), z(x, t_{j+1}))_{\tau}^*.$$

Using the inequality $\tau |(\psi^0, z^j)^*| \leq 0.5 \|z\|^2 + 0.5r^2 + 0.5t^2 \|\psi^0\|^2$, and also the initial condition $z(x, 0) = 0$ we find

$$\|z(x, t_{j+1})\|^2 \leq c_0 \tau \sum_{j'=1}^j \|z(x, t_{j'})\|^2 + \tau^2 \|\psi^0(x, t_{j+1})\|^2 + \sum_{j'=1}^j \tau Q^{j'}.$$ 

Let us now use the following simple lemma, the proof of which we omit:

**Lemma 2.** If $\rho(t_j)$ and $\sigma(t_j)$ are non-negative functions ($t_j = j\tau, j = 1, 2, \ldots$) then the inequality
\[
\rho(t_{j+1}) \leq c_0 \sum_{j'=1}^{j} \rho(t_{j'}) + \sigma(t_{j+1}), \quad j = 1, 2, \ldots, \quad (c_0 > 0)
\]
gives
\[
\rho(t_{j+1}) \leq \sigma(t_{j+1}) + c_0 \sum_{j'=1}^{j} \tau \rho(t_{j'}). \quad (20)
\]

It follows from the inequality (19), from Lemma 2, that
\[
\|z(x, t_{j+1})\| \leq \tau^2 \|\psi^0(x, t_{j+1})\|^2 + \tau^2 \varepsilon \|\psi^0(x, t_{j})\|^2 + (1 + \varepsilon) \sum_{j'=1}^{j+1} \tau Q^{j'},
\]
\[
\|\psi^0(x, t_j)\|^2 = \frac{1}{t_j} \sum_{j'=1}^{j} \tau \|\psi^0(x, t_{j'})\|^2.
\]

This proves Lemma 1.

8. Theorem 1. If conditions A are satisfied, then the difference scheme (5)-(7) converges in the mean with speed \(O(\|h^2\|_b + O(\tau))\) for \(\sigma = 0.5\) and as \(h, \tau\) tend independently to zero on any non-uniform space net:
\[
\|y(x, t_{j+\alpha/2}) - u(x, t_{j+\alpha/2})\| \leq M(\|h^2\|_b + \tau), \quad \alpha = 1, 2, \quad j = 0, 1, \ldots, \ K - 1, \quad (21)
\]
\[
\|h^2\|_b = \|h^2\|_{b_1} + \|h^2\|_{b_2}, \quad \|h^2\|_{b_\alpha} = \sqrt{(1, h^2)|_{z}}, \quad \alpha = 1, 2,
\]
where \(M\) is a positive constant which does not depend on the choice of net.

For the estimate (21) for \(\alpha = 2\) (on a whole step) follows from (18) since, from (10) and (12) we have \(Q^{j'} = O(\|h^2\|_b) + O(\tau^2).\) To estimate \(\|z(x, t_{j+\alpha/2})\|\) on a fractional step we use (18) with \(\alpha = 2\) and the identity (14) for \(\alpha = 1.\) Thus on fractional steps the scheme (5)-(7) has the same order of accuracy as on integral steps.

The convergence of the scheme (5)-(7) can be proved with considerably weaker requirements than conditions A.

Theorem 2. If the following conditions are satisfied then the scheme (4)-(6) converges in the mean for \(\sigma = 0.5\) and for any \(h, \tau\) so that
\[
\|y(x, t_{j+\alpha/2}) - u(x, t_{j+\alpha/2})\| \leq M(\|h^2\|_b + \tau^\gamma), \quad \gamma \leq 1, \quad (22)
\]
The conditions are (1) \(\partial k_\alpha/\partial x_\alpha, \partial^2 u/\partial x_\alpha^2\) satisfy in \(\mathcal{R}\) the Hölder conditions of order \(\gamma, 0 < \gamma \leq 1,\) (2) \(\partial u/\partial x_\alpha, \partial^2 u/\partial x_\alpha^2, \partial u/\partial t, k_\alpha, \partial k_\alpha/\partial x_\alpha\) satisfy in \(\mathcal{R}\)
the Hölder conditions of order $\gamma > 0$ with respect to $t$.

To prove Theorem 2 we have to use the fact that $\|\Phi_\alpha\|_a = O(h_\alpha^{\gamma}) + O(\tau^\gamma)$, $\|\mu_\alpha\|_a = O(h_\alpha^{1+\gamma}|a|)$ and make use of Lemma 1.

9. We have given all the preceding reasoning for the simplest equation of heat conductivity, (1). However, Theorems 1 and 2 remain true for equations of the general form

$$c(x,t)\frac{\partial u}{\partial t} = \sum_{a=1}^{2} L_a u - q(x,t)u; \quad L_a u = \frac{\partial}{\partial x_a} k_a(x,t) \frac{\partial u}{\partial x_a} + r_a(x,t) \frac{\partial u}{\partial x_a},$$

$$c(x,t)\frac{\partial u}{\partial t} = \sum_{a=1}^{2} L_a u + f(x,t,u); \quad L_a u = \frac{\partial}{\partial x_a} k_a(x,t,u) \frac{\partial u}{\partial x_a} + r_a(x,t,u) \frac{\partial u}{\partial x_a}.$$

The corresponding expressions for $A_{\alpha} \gamma$ are given in [3] and we shall not give them here. It is proved there that the local one-dimensional method is uniformly convergent for $\sigma = 1$. If $r_\alpha \neq 0, q_\alpha \neq 0$, then the estimate is true for $\tau < \tau_0$ where $\tau_0$ depends on $c_1$ and $\max|r_\alpha|, \max|q_\alpha|$.

10. The method of summation of local errors $\psi_\alpha$ used here, which leads to the $a$ priori estimate (18), cannot be extended to the case $p > 2$. In this case instead of (21) we have only been able to obtain the estimate

$$\|y - u\| \leq M(\sqrt{\tau} + h^2).$$

If we suppose that $\tau/h_2^2 \leq 2(1-\sigma)\max a_\alpha^{-1}$, then using the method of [3] we can prove the uniform convergence of the scheme (5)-(7) for $0 < \sigma < 1$ and for any $p > 1$ with speed $O(h^2) + O(\tau)$, where $h^2 = h_1^2 + h_2^2$, $h_3$ is the maximum value of the step $h_\alpha^{(i)}$ on the net $h^{(i+1)}$. For the case $\sigma = 1$ this estimate holds for any $h_\alpha$ and $\tau$ (Theorem 2 of [3]).

It must be stressed that we are using here boundary conditions without approximation (in contrast to [3]). This formulation of the difference boundary conditions always leads to different schemes on non-uniform nets. We have proved that in this case also second order accuracy is attained with respect to space.

Translated by R. Feinstein

REFERENCES