Karel Rektorys

On application of direct variational methods to the solution of parabolic boundary value problems of arbitrary order in the space variables

Czechoslovak Mathematical Journal, Vol. 21 (1971), No. 2, 318–339

Persistent URL: `http://dml.cz/dmlcz/101024`

Czechoslovak Mathematical Journal, 21 (96) 1971, Praha

ON APPLICATION OF DIRECT VARIATIONAL METHODS TO THE SOLUTION OF PARABOLIC BOUNDARY VALUE PROBLEMS OF ARBITRARY ORDER IN THE SPACE VARIABLES

KAREL REKTORYS, Praha

(Received January 15, 1970)

The Dirichlet problem for parabolic equations is treated in a way permitting the application of direct variational methods similar to those applied to elliptic problems.

1. INTRODUCTION

Direct variational methods commonly used to the solution of elliptic boundary value problems and based on the minimalization of corresponding functionals, are not used in parabolic problems, because functionals with similar properties do not exist. In fact, functionals have been constructed in the parabolic case, containing convolution integrals and permitting the formulation of analogous variational principles (see e.g. HLAVÁČEK [2]). However, although these principles are very important from the theoretical point of view, the convolution integrals, involved in the functionals, do not make it possible to apply such simple procedures for numerical solution of corresponding problems as is, for example, the Ritz method in the elliptic analogy. Thus, in this paper a method is developped, permitting the application of direct variational methods, and corresponding questions on convergence are clarified.

2. PRELIMINARY CONSIDERATIONS. AN EXAMPLE

To make clear the idea of the method, let us consider first a simple example of an equation of the second order. Let Ω be a bounded region in the N-dimensional Euclidean space E_N, with boundary $\dot{\Omega}$, let $x_1, \ldots, x_N$ be Cartesian coordinates of the point $x \in E_N$. Denote $Q = \Omega \times (0, T)$. Let the Dirichlet problem for a parabolic equation be given:

$$(1) \quad Au + \frac{\partial u}{\partial t} \equiv - \sum_{i,j=1}^{N} \frac{\partial}{\partial x_i}\left(a_{ij}(x) \frac{\partial u}{\partial x_j}\right) + \frac{\partial u}{\partial t} = f(x) \quad \text{in} \quad Q\,, \quad a_{ij} = a_{ji}\,,$$

$$(2) \qquad u(x, 0) = u_0(x)\,,$$

$$(3) \qquad u = 0 \quad \text{on} \quad \dot{\Omega} \times (0, T)\,.$$

Assumptions concerning the given data of the problem (the ellipticity of the operator A, etc.) will be specified in the next section for a more general case (see, in particular, Theorem 1, p. 329).

Denote

$$a(u, u) = \int_\Omega \sum_{i,j=1}^{n} a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} \, dx$$

and (v, u) the usual scalar product of (real) functions v and u in the region Ω, i.e.

$$(v, u) = \int_\Omega v(x) \, u(x) \, dx .$$

Let us divide the interval $[0, T]$ into p subintervals of the same length h (i.e. $ph = T$) and let us consider the following functionals:

(4)
$$\begin{aligned}
G_1(u) &= a(u, u) + \frac{1}{h}(u, u) - 2(f, u) - \frac{2}{h}(u_0, u), \\
G_2(u) &= a(u, u) + \frac{1}{h}(u, u) - 2(f, u) - \frac{2}{h}(u_1, u), \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
G_p(u) &= a(u, u) + \frac{1}{h}(u, u) - 2(f, u) - \frac{2}{h}(u_{p-1}, u).
\end{aligned}$$

Under well-known assumptions on $a(u, u)$, the functional $G_1(u)$ attains its minimum in a certain class of functions, satisfying condition (3) (let us denote this class of functions by V, see the following section) and the minimizing function u_1 is the (generalized) solution of the elliptic problem

$$Au + \frac{1}{h}(u - u_0) = f \quad \text{in} \quad \Omega, \quad u = 0 \quad \text{on} \quad \dot{\Omega}.$$

Similarly, the function u_2, minimizing the functional $G_2(u)$ in V, satisfies

$$Au_2 + \frac{1}{h}(u_2 - u_1) = f \quad \text{in} \quad \Omega, \qquad u_2 = 0 \quad \text{on} \quad \dot{\Omega}.$$

In general, the functional $G_i(u)$ $(1 \leqq i \leqq p)$ attains its minimum in V and the minimizing function satisfies (in the generalized sense)

(5)
$$Au_i + \frac{1}{h}(u_i - u_{i-1}) = f \quad \text{in} \quad \Omega, \qquad u_i = 0 \quad \text{on} \quad \dot{\Omega}.$$

For every $t_i = ih$ ($i = 1, \ldots, p$) being

$$\frac{1}{h}\,(u_i(x) - u_{i-1}(x)) \approx \frac{\partial u}{\partial t}\,(x, t_i)\,,$$

each of the functions $u_i(x)$ can be taken as an approximation, in the hyperplane $t = t_i$ ($x \in \Omega$), of the solution $u(x, t)$ of the problem (1)–(3). If wanted, an approximation $u_1(x, t)$ can be defined in the whole Q, for example as a function continuous and sectionally linear in t for every fixed $x \in \Omega$, assuming the values $u_i(x)$ at the points $t = t_i$. Thus

(6) $$u_1(x, t) = u_j(x) + \frac{t - t_j}{h}\,(u_{j+1}(x) - u_j(x)) \quad \text{for} \quad t_j \leqq t \leqq t_{j+1}\,,$$

where $t_j = jh$ ($j = 0, 1, \ldots, p - 1$).

Let us construct, in a similar way, a function $u_2(x, t)$, with the only difference that instead of dividing the interval $[0, T]$ into p subintervals of the length h as before, we divide it into $2p$ subintervals of the length $h_2 = h/2$. Going on in this way and dividing subsequently the interval $[0, T]$ into $4p, 8p, \ldots, 2^{n-1}p, \ldots$ subintervals, we construct a sequence of functions $u_n(x, t)$, defined in Q by the relations

(7) $$u_n(x, t) = u_j^n(x) + \frac{t - t_j^n}{h_n}\,(u_{j+1}^n(x) - u_j^n(x)) \quad \text{for} \quad t_j^n \leqq t \leqq t_{j+1}^n\,,$$

where $h_n = h/2^{n-1}$, $t_j^n = jh_n$ ($j = 0, 1, \ldots, p\,.\,2^{n-1} - 1$). (For $n = 1$, the notation $h_1 = h$, $t_j^n = t_j$, $u_j^1(x) = u_j(x)$ has been used.)

In this way we get a sequence $\{u_n(x, t)\}$ of approximate solutions of the problem (1)–(3).

In practice, the minima of functionals (4) are determined approximately, using some of the well-known direct methods, for example, the Ritz method: Let

$$\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x), \ldots$$

be a base in V. Let us take the first k members of this sequence and construct a function

(8) $$\tilde{u}_1(x) = a_1\varphi_1(x) + \ldots + a_k\varphi_k(x)\,,$$

determining the unknown coefficient a_i in a well-known way from the condition that $G_1(\tilde{u}_1)$ be minimal. Then substitute $\tilde{u}_1(x)$ instead of $u_1(x)$ into $G_2(u)$, denote this new functional by $\tilde{G}_2(u)$ and find $\tilde{u}_2(x)$ as a linear combination of $\varphi_1(x), \ldots, \varphi_k(x)$ so that $\tilde{G}_2(\tilde{u}_2)$ be minimal. Of course, $\tilde{u}_2(x)$ is different from $u_2(x)$, in general, because, first the functional $G_2(u)$ has been substituted by the functional $\tilde{G}_2(u)$ and, second, the minimizing element of this functional is again found approximatively.

Going on in this way, we come to the functions $\tilde{u}_3(x), \ldots, \tilde{u}_t(x)$. (The numerical process is very simple, see Section 6, p. 337, especially if a computer can be used.)

Having these functions, we construct a function $\tilde{u}_1(x, t)$ substituting $\tilde{u}_j(x)$ instead of $u_j(x)$ $(j = 1, 2, \ldots, p)$ into (6). Then, dividing the interval $[0, T]$ into $2p, 4p, \ldots$ subintervals, we come in the same way to the functions $\tilde{u}_2(x, t), \tilde{u}_3(x, t), \ldots$ and, generally, to the function

$$\tilde{u}_n(x, t) = \tilde{u}_j^n(n) + \frac{t - t_j^n}{h_n} (\tilde{u}_{j+1}^n(x) - \tilde{u}_j^n(x)) \quad \text{for} \quad t_j^n \leqq t \leqq t_{j+1}^n \tag{9}$$

which is of a similar form as the function (7).

Note that equations (5) and functions (7) are closely related with the so-called Rothe method (see ROTHE [4]). The expression $(u_i - u_{i-1})/h$ in (5) represents the "differential quotient" of u with respect to t; it seems plausible that for $n \to \infty$ (i.e. for $h_n \to 0$) the limiting function of the sequence $\{u_n(x, t)\}$ will have a derivative with respect to t and that it will be possible to establish this function as the required solution.

Of course, there are many questions concerning theoretical treating of our problem. The first is the concept of the solution itself, i.e. the question what a function or what an element of a properly chosen functional space is to be understood under a solution, especially in the case when an equation of higher than the second order is considered. Parabolic problems for higher-order equations have been treated by different authors (LIONS, BROWDER, LADYŽENSKAJA a.o.), even in nonlinear case. The concepts of the solution are slightly different in these works, according to the problems in question and to methods used by individual authors. Our concept of the solution is clarified in Theorem 1 (p. 329).

Further, there are two questions on convergence to be discussed. The first concerns the above mentioned convergence (for $n \to \infty$) of the "Rothe functions" $u_n(x, t)$ to the required solution. The second concerns the behavior of their approximations $\tilde{u}_n(x, t)$, constructed by the above explained direct method. In more detail, a question arises, whether it is possible to make $\tilde{u}_n(x, t)$ arbitrarily "close" to the solution of our problem (and in what a sense) if n and also k (in the linear combinations of the type (8)) are sufficiently large.

As to the first question: The Rothe method has been used several times to proofs of existence theorems, especially for parabolic problems of the second order. (See, e.g., ROTHE [4], LADYŽENSKAJA [5], or a surveyable paper by ILJIN, KALAŠNIKOV and OLEINIK [6], etc.). In these papers, mostly classical methods or methods based on fixed point theorems of the Schauder type are used. A basic work applying the Rothe method to equations of higher than the second order is the paper by LADYŽENSKAJA [7] (see also the paper by IBRAGIMOV [8]). Although the concept of her work is sufficiently general to cover a wide field of problems (and suggests also some ideas used in the present paper), we have chosen a rather different way, more appropriate for treating our problem, especially more suitable for answering the second question on convergence, stated above.

After these introductory considerations let us turn to more general investigations.

3. THE WEAK SOLUTION. GENERAL GIVEN DATA

Let Ω be a bounded region in E_N with a Lipschitz boundary (see e.g. NEČAS [1]; thus, the outward normal ν exists almost everywhere; examples of such regions are a circle, an annulus, a triangle, a cube, etc.). Let $a_{ij}(x)$ are bounded integrable functions in $\Omega, f(x) \in L_2(\Omega)$. We shall assume that a_{ij}, f are real. However, this assumption is not at all essential for basic results of this work.

Let

$$A = \sum_{|i|,|j| \leq k} (-1)^{|i|} D^i(a_{ij}(x) D^j) \tag{10}$$

be a differential operator of order $2k$; i, j are multiindices, $i = (i_1, \ldots, i_N)$, $|i| = i_1 + \ldots + i_N$, where $i_1, \ldots, i_N$ are nonnegative integers,

$$D^i = \frac{\partial^{|i|}}{\partial x_1^{i_1} \ldots \partial x_N^{i_N}};$$

similarly for j.

As usual, let $W_2^{(l)}(\Omega)$ be the Hilbert space of (real) functions which are square integrable in Ω with their generalized derivatives (taken in the sense of distributions) up to the order l including, with the scalar product

$$(v, u)_{W_2^{(l)}} = \sum_{|i| \leq l} (D^i v, D^i u),$$

where $(\ldots)$ is the scalar product in $L_2(\Omega)$. Let

$$V = E\left(v \in W_2^{(k)}(\Omega),\ v = \frac{\partial v}{\partial \nu} = \ldots = \frac{\partial^{k-1} v}{\partial \nu^{k-1}} = 0 \text{ on } \dot{\Omega}\right). \tag{11}$$

Let

$$a(v, u) = \sum_{|i|,|j| \leq k} \int_\Omega a_{ij} D^i v D^j u \, dx \tag{12}$$

be the bilinear form assotiated with the operator A and let us assume that this form satisfies the condition of V-ellipticity, i.e.

$$a(v, v) \geqq \alpha \|v\|^2_{W_2^{(k)}} \quad \text{for every} \quad v \in V \quad (\alpha = \text{const.} > 0). \tag{13}$$

As is well known from the theory of elliptic boundary value problems, (13) ensures the existence of just one solution $u \in V$ of the elliptic boundary value problem

$$Au = f, \quad u \in V \tag{14}$$

in the weak sense, i.e.

$$a(v, u) = (v, f) \quad \text{is satisfied for every} \quad v \in V.$$

Now, let the parabolic boundary value problem be given,

$$Au + \frac{\partial u}{\partial t} = f \quad \text{in} \quad Q \equiv \Omega \times (0, T), \tag{15}$$

$$u(x, 0) = u_0(x), \tag{16}$$

$$u = \frac{\partial u}{\partial \nu} = \ldots = \frac{\partial^{k-1} u}{\partial \nu^{k-1}} = 0 \quad \text{on} \quad \dot{\Omega} \times (0, T) \tag{17}$$

with A and f satisfying conditions introduced above. Moreover, let us first consider the case $u_0 = 0$ (cf., however, Section 5). We shall show in the present section that the given assumptions ensure the existence of a weak solution of the problem (15)–(17) in the sense given in Theorem 1 (p. 329) and that this weak solution is the limit (in the sense given later) of the sequence $\{u_n(x, t)\}$ of functions constructed above by the "Rothe method". In the next section, we shall be interested in the assumptions under which more smooth solutions of this problem can be obtained.

Thus, choose a fixed number p, denote $h = T/p$ and solve successively the following finite sequence of elliptic boundary value problems:

$$a(v, z_1) + \frac{1}{h}(v, z_1) = (v, f), \quad z_1 \in V, \tag{18}$$

$$a(v, z_2) + \frac{1}{h}(v, z_2) = (v, f) + \frac{1}{h}(v, z_1), \quad z_2 \in V, \tag{19}$$

. . . .

$$a(v, z_j) + \frac{1}{h}(v, z_j) = (v, f) + \frac{1}{h}(v, z_{j-1}), \quad z_j \in V, \tag{20}$$

. . . .

$$a(v, z_p) + \frac{1}{h}(v, z_p) = (v, f) + \frac{1}{h}(v, z_{p-1}), \quad z_p \in V. \tag{21}$$

Each of equations (18)–(21) should be fulfilled for all $v \in V$.

In consequence of (13) and of the fact that $1/h > 0$, a unique solution $z_1 \in V$ of (18) exists satisfying

$$\|z_1\|_V \leqq \text{const.} \|f\|_{L_2(\Omega)}. \tag{22}$$

Putting $v = z_1$ into (18),

$$a(z_1, z_1) + \frac{1}{h}(z_1, z_1) = (z_1, f),$$

and notig that $a(z_1, z_1) \geqq 0$, we get

$$\|z_1\|_{L_2(\Omega)} \leqq h\|f\|_{L_2(\Omega)}. \tag{23}$$

Similarly, z_1 being in $L_2(\Omega)$, a unique solution $z_2 \in V$ of (19) exists with

$$\|z_2\|_V \leqq \text{const.} \left\| f + \frac{1}{h} z_1 \right\|_{L_2(\Omega)} \leqq \text{const.} \|2f\|_{L_2(\Omega)} .$$

If we subtract (18) from (19),

$$a(v, z_2 - z_1) + \frac{1}{h}(v, z_2 - z_1) = \frac{1}{h}(v, z_1)$$

and put $v = z_2 - z_1$, we get in a similar way as in (23)

$$\|z_2 - z_1\|_{L_2(\Omega)} \leqq \|z_1\|_{L_2(\Omega)} \leqq h\|f\|_{L_2(\Omega)} . \tag{24}$$

Proceeding similarly, we have

$$\|z_{j+1} - z_j\|_{L_2(\Omega)} \leqq \|z_j - z_{j-1}\|_{L_2(\Omega)} \leqq h\|f\|_{L_2(\Omega)} \tag{25}$$

for $j = 2, 3, \ldots, p - 1$ (and also for $j = 1$, according to (24), noting that $z_0 = 0$).

From (23) and (25) it follows

$$\|z_j\|_{L_2(\Omega)} \leqq T\|f\|_{L_2(\Omega)} \tag{26}$$

for all $j = 1, \ldots, p$.

If we write (20) in the form

$$a(v, z_j) = (v, f) - \frac{1}{h}(v, z_j - z_{j-1})$$

and use (25), we come, similarly as in (22) to the result

$$\|z_j\|_V \leqq \text{const.} \left(\|f\|_{L_2(\Omega)} + \left\| \frac{z_j - z_{j-1}}{h} \right\|_{L_2(\Omega)} \right) \leqq K\|f\|_{L_2(\Omega)} , \tag{27}$$

where K does not depend on h.

Now, denote $t_j = jh$ ($j = 0, \ldots, p$) and construct in $\bar{Q} = \bar{\Omega} \times [0, T]$ the function

$$u_1(x, t) = z_j(x) + \frac{t - t_j}{h} \{z_{j+1}(x) - z_j(x)\} \quad \text{for} \quad t_j \leqq t \leqq t_{j+1} , \tag{28}$$

$j = 0, \ldots, p - 1$. Thus, in $\bar{Q}$, this function is piecewise linear in t and for every $t = t_j$ we have $u_1(x, t_j) = z_j(x)$. We shall also use the notation $u_1(x, t) = u_1(t)$, having on mind the mapping $t \to u_1(t)$ from the interval $[0, T]$ into V.

From the construction of the function (28) it is evident that, for every $t \in [0, T]$, we have

$$\frac{\partial u_1}{\partial x_1} = \frac{\partial z_1}{\partial x_1} + \frac{t - t_j}{h} \left\{ \frac{\partial z_{j+1}}{\partial x_1} - \frac{\partial z_j}{\partial x_1} \right\}, \quad t_j \leqq t \leqq t_{j+1} ,$$

and that the so constructed function is the generalized derivative (with respect to x_1) of the function u_1, taken as function of $x_1, \ldots, x_N, t$ in Q. In fact, the equation

$$\int_\Omega \varphi \frac{\partial u_1}{\partial x_1} \, \mathrm{d}x = - \int_\Omega \frac{\partial \varphi}{\partial x_1} u_1 \, \mathrm{d}x$$

being fulfilled for every $t \in [0, T]$ and for every function $\varphi(x)$ with compact support in Ω, it follows that

$$\int_Q \varphi \frac{\partial u_1}{\partial x_1} \, \mathrm{d}x \, \mathrm{d}t = - \int_Q \frac{\partial \varphi}{\partial x_1} u_1 \, \mathrm{d}x \, \mathrm{d}t \tag{29}$$

will be fulfilled for every $\varphi(x, t)$ with compact support in Q. A similar assertion holds for the derivatives

$$D^i u_1 = D^i z_j + \frac{t - t_j}{h} \{D^i z_{j+1} - D^i z_j\} \quad (|i| \leqq k) . \tag{30}$$

The same is true for the derivative

$$\frac{\partial u_1}{\partial t} = \frac{z_{j+1} - z_j}{h}, \quad t_j < t < t_{j+1} , \tag{31}$$

defined by (31) almost everywhere in Q.

Till now, we considered the division (denote it by d_1) of the interval $[0, T]$ into p subintervals of the length h by the points of division $t_j = jh$ ($j = 0, \ldots, p$). Let us consider a sequence of divisions $\{d_n\}$ of the interval $[0, T]$ into $p \,.\, 2^{n-1}$ subintervals of the length $h_n = T/(p \,.\, 2^{n-1})$ with points of division $t_j^n = jh_n$ ($j = 0, \ldots, p \,.\, 2^{n-1}$). Thus, $h = h_1$. Denote the functions z_j constructed above for the division d_1 by z_j^1 and functions, constructed similarly for the division d_n, by z_j^n. In the same way as before, we come to the estimates (26), (27) (valid independently of h_n) and (25) and to the functions

$$u_n(t) = u_n(x, t) = z_j^n(x) + \frac{t - t_j^n}{h_n} \{z_{j+1}^n(x) - z_j^n(x)\} \quad \text{for} \quad t_j^n \leqq t \leqq t_{j+1}^n , \tag{32}$$

with properties similar to the properties of the function (28).

Denote briefly $[0, T] = I$. Let H be a Hilbert space elements of which are functions defined in Ω. Let $L_2(I, H)$ be a set of functions $t \to W(t)$ from I into H (more precisely, $W(t) \in H$ for almost all $t \in I$, in general) which are square integrable in the interval I (in the Bochner sense; for details see, e.g., Wilcox [3]). $L_2(I, H)$ is a Hilbert space with the scalar product

$$(W_1, W_2)_{L_2(I,H)} = \int_0^T (W_1(t), W_2(t))_H \, \mathrm{d}t .$$

The integral

$$\int_0^t W(\tau)\,d\tau = w(t) \tag{33}$$

(in the Bochner sense) of a function $W(t) \in L_2(I, H)$ is defined by

$$(w(t), r)_H = \int_0^t (W(\tau), r)_H\,d\tau \tag{34}$$

for every $r \in H$. If $W(t) \in L_2(I, H)$, then $w(t)$ can be identified with a function representing a continuous mapping from I into H; thus, we can write

$$w(t) \in C^0(I, H)\,. \tag{35}$$

Moreover, $w(t)$ has a derivative $w'(t)$ almost everywhere in I (in the sense

$$\lim_{\Delta t \to 0} \left\| \frac{w(t + \Delta t) - w(t)}{\Delta t} - W(t) \right\|_H = 0)\,,$$

equal to $W(t)$.

It follows from (32) and (27) that the sequence $\{u_n(x, t)\}$ is uniformly bounded in $L_2(I, V)$. This space being complete and reflexif, a subsequence

$$\{u_{k_n}(t)\} \tag{36}$$

can be found, weakly convergent in $L_2(I, V)$ to a function

$$u(t) \in L_2(I, V)\,. \tag{37}$$

Denote

$$Z^n_{j+1}(x) = \frac{z^n_{j+1}(x) - z^n_j(x)}{h}$$

and

$$U_n(t) = U_n(x, t) = Z^n_{j+1}(x) \quad \text{for} \quad t^n_j \leqq t \leqq t^n_{j+1}\,, \tag{38}$$

$j = 0, 1, \ldots, p\,.\,2^n - 1$. According to (25), the sequence $\{U_n(t)\}$ is uniformly bounded in $L_2(I, L_2(\Omega))$, so that the sequence (36) can be choosen in such a way that, at the same time, the subsequence

$$\{U_{k_n}(t)\}$$

of the sequence $\{U_n(t)\}$ converges weakly in $L_2(I, L_2(\Omega))$ to a function

$$U(t) \in L_2(I, L_2(\Omega))\,.$$

Thus, the integral

$$\int_0^t U(\tau)\,d\tau = w(t)$$

exists (in the Bochner sense (34), where $H = L_2(\Omega)$). Because we have

$$\int_0^t U_n(\tau)\, d\tau = u_n(t)$$

(in $L_2(I, L_2(\Omega))$), as can be immediately seen from (32) and (38), it easily follows that

$$w(t) = u(t)$$

in $L_2(I, L_2(\Omega))$. Similarly as in (35), we can write

$$u(t) \in C^0(I, L_2(\Omega))$$

and

$$u'(t) = U(t)$$

in $L_2(I, L_2(\Omega))$.

Because, in consequence of (23), (25) and (32),

$$\|u_n(x, t)\|_{L_2(\Omega)} \leqq t\|f\|_{L_2(\Omega)}$$

for every $t \in I$ and for every n, we have

$$u(0) = 0 \tag{39}$$

in $C^0(I, L_2(\Omega))$.

To show in what a sense the function $u(t)$ satisfies the given differential equation, let us consider the system (18)–(21) (adding again the suffix n). A suitable linear combination of two of consecutive equations of this system gives

$$a\left(v, z_j^n + \frac{t - t_j^n}{h_n}\{z_{j+1}^n - z_j^n\}\right) + \left(v, Z_{j+1}^n + \frac{t - t_j^n}{h_n}\{Z_{j+1}^n - Z_j^n\}\right) = (v, f)$$

in the interval $t_j^n \leqq t \leqq t_{j+1}^n$, or

$$a(v, u_n(t)) + \left(v, U_n(t) + \frac{t - t_j^n}{h_n}\{Z_{j+1}^n - Z_j^n\}\right) = (v, f). \tag{40}$$

Let, first, v be independent of t in the interval $[0, T]$. Integrating (40) between the limits $[0, T]$, we have

(41)

$$\int_0^T a(v, u_n(t))\, dt + \int_0^T (v, U_n(t))\, dt + \sum_{t^n_j} \int_{t^n_j}^{t^n_{j+1}} \frac{t - t_j^n}{h_n}(v, Z_{j+1}^n - Z_j^n)\, dt = \int_0^T (v, f)\, dt.$$

The functions $u_n(t)$, $U_n(t)$ converging weakly in $L_2([0, T], V)$, or $L_2([0, T], L_2(\Omega))$ to $u(t)$, or $u'(t)$, respectively, the two first integrals converge to the integrals

$$\int_0^T a(v, u(t))\, dt, \quad \int_0^T (v, u'(t))\, dt.$$

In the third term in (41) we have (because $h_n = t^n_{j+1} - t^n_j$)

$$\int_{t^n_j}^{t^n_{j+1}} \frac{t - t^n_j}{h_n} (v, Z^n_{j+1} - Z^n_j) \, \mathrm{d}t = \frac{h_n}{2} (v, Z^n_{j+1} - Z^n_j) ,$$

and, consequently

$$\sum_{t^n_j} \int_{t^n_j}^{t^n_{j+1}} \frac{t - t^n_j}{h_n} (v, Z^n_{j+1} - Z^n_j) \, \mathrm{d}t = \frac{h_n}{2} (v, \{Z^n_2 - Z^n_1\} + \{Z^n_3 - Z^n_2\} + \ldots$$

$$\ldots + \{Z^n_{2^{n-1}p} - Z^n_{2^{n-1}p-1}\}) = \frac{h_n}{2} (v, Z^n_{2^{n-1}p} - Z^n_1) .$$

With $n \to \infty$, we have $h_n \to 0$ and $\|Z^n_j\|_{L_2(\Omega)} \leqq \|f\|_{L_2(\Omega)}$ according to (25). Thus, for every fixed $v \in V$, the third term in (41) converges to zero if $n \to \infty$ and we have

$$\int_0^T a(v, u(t)) \, \mathrm{d}t + \int_0^T (v, u'(t)) \, \mathrm{d}t = \int_0^T (v, f) \, \mathrm{d}t . \tag{42}$$

This result remains true for every function $v(t)$ which represents a continuous mapping from $[0, T]$ into V. These functions being dense in $L_2([0, T], V)$, (42) remains true for every $v(t)$ of this space:

$$\int_0^T a(v(t), u(t)) \, \mathrm{d}t + \int_0^T (v(t), u'(t)) \, \mathrm{d}t = \int_0^T (v(t), f) \, \mathrm{d}t , \quad v(t) \in L_2([0, T]\ V) . \tag{43}$$

Thus, we have established the existence of a function $u(t) \in L_2([0, T], V)$ (let us call this function a weak solution of the problem (15)–(17)), lying simultaneously in $C^0([0, T], L_2(\Omega))$, having, as function of this last space, a derivative $u'(t) \in L_2([0, T], L_2(\Omega))$ and satisfying the differential equation (15) in the sense (43) and the given initial condition in the sense (39).

Uniqueness: We shall show that there cannot exist two different functions $u_1(t)$, $u_2(t)$ having the same properties. Thus, let $u(t) = u_2(t) - u_1(t)$. We have

$$u(t) \in L_2([0, T], V) ,$$

$$u(t) \in C^0([0, T], L_2(\Omega)) ,$$

$$u'(t)_{L_2([0,T], L_2(\Omega))} \in L_2([0, T],\ L_2(\Omega)) ,$$

$$u(0) = 0 \quad \text{in} \quad C^0([0, T], L_2(\Omega)) , \tag{44}$$

(45)

$$\int_0^T a(v(t), u(t)) \, \mathrm{d}t + \int_0^T (v(t), u'(t))_{L_2(\Omega)} \, \mathrm{d}t = 0 \quad \text{for every} \quad v(t) \in L_2([0, T], V) .$$

Let $t^1 \in [0, T]$ be arbitrary. Let

$$(46) \qquad v(t) = \begin{cases} u(t) & \text{for} \quad 0 \leqq t \leqq t^1 , \\ 0 & \text{for} \quad t^1 < t \leqq T . \end{cases}$$

Thus, $v(t) \in L_2([0, T], V)$. Further, $v(t) \in C^0([0, t^1], L_2(\Omega))$ and

$$(47) \qquad \int_0^{t^1} (v(t), u'(t))_{L_2(\Omega)} \, dt = \int_0^{t^1} (u(t), u'(t))_{L_2(\Omega)} \, dt =$$

$$= \tfrac{1}{2}\{(u(t^1), u(t^1))_{L_2(\Omega)} - (u(0), u(0))_{L_2(\Omega)}\} = \tfrac{1}{2}\|u(t^1)\|^2_{L_2(\Omega)} .$$

The function (46), substituted into (45), fulfills

$$\int_0^{t^1} a(u(t), u(t)) \, dt + \int_0^{t^1} (u(t), u'(t))_{L_2(\Omega)} \, dt = 0 .$$

If we note that $a(u(t), u(t)) \geqq 0$ for all $t \in [0, t^1]$, we get according to (47),

$$\tfrac{1}{2}\|u(t^1)\|^2_{L_2(\Omega)} = 0 ,$$

or $u(t^1) = 0$ (in $L_2(\Omega)$). For t^1 was chosen arbitrarily in the interval $[0, T]$, we have $u(t) = 0$ in $[0, T]$ (and $u(x, t) = 0$ in Q almost everywhere).

Remark 1. It follows, in the usual way, by the uniqueness, that not only $\{u_{k_n}(t)\}$ but the whole sequence $\{u_n(t)\}$ converges weakly in $L_2([0, T], V)$ to the just described solution.

Thus, we have proved the following theorem:

Theorem 1. *Let $Q = \Omega \times (0, T)$, where Ω is a bounded region in E_N with a Lipschitz boundary $\dot{\Omega}$* (see Nečas [1]). *Let the boundary value problem*

$$Au + \frac{\partial u}{\partial t} = f \quad in \quad Q ,$$

$$u(x, 0) = 0 ,$$

$$u = \frac{\partial u}{\partial \nu} = \ldots = \frac{\partial^{k-1} u}{\partial \nu^{k-1}} = 0 \quad \text{on} \quad \dot{\Omega} \times (0, T)$$

be given, where

$$f(x) \in L_2(\Omega) ,$$

$$A = \sum_{|i|,|j| \leqq k} (-1)^{|i|} D^i(a_{ij}(x) D^j)$$

with $a_{ij}(x)$ bounded and measurable in Ω. Let the ellipticity condition be satisfied,

$$a(v, v) \geqq \alpha \|v\|^2_V \quad for\ every \quad v \in V , \quad \alpha = \text{const.} > 0 ,$$

where

$$a(v, u) = \sum_{|i|,|j| \leq k} \int_{\Omega} a_{ij} D^i v D^j u \, dx \, ,$$

$$V = E \left\{ v \in W_2^{(k)}(\Omega), \ v = \frac{\partial v}{\partial \nu} = \ldots = \frac{\partial^{k-1} v}{\partial \nu^{k-1}} = 0 \ \ (\textit{in the sense of traces}) \right\}.$$

Then there exists exactly one solution $u(t) = u(x, t)$ of this problem, in the following sense:

$$u(t) \in L_2([0, T], V) \, ,$$

$$u(t) \in C^0([0, T], L_2(\Omega)) \, ,$$

$$u'(t)_{L_2([0,T],L_2(\Omega))} \in L_2([0, T], L_2(\Omega)) \, ,$$

$$u(0) = 0 \quad \text{in} \quad C^0([0, T], L_2(\Omega)) \, ,$$

$$\int_0^T a(v(t), u(t)) \, dt + \int_0^T (v(t), u'(t)) \, dt = \int_0^T (v(t), f) \, dt$$

for every $v(t) \in L_2([0, T], V)$.

This solution is a weak limit in $L_2([0, T], V)$ of the sequence $\{u_n(t)\}$ of functions (32) constructed above.

Remark 2. It follows easily from the text following equation (28), coming to $n \to \infty$, that not only the functions $u_n(x, t)$, but the function $u(t)$ too, taken as function of $x_1, \ldots, x_N$, t in Q, has in Q square integrable generalized derivatives up to the order k with respect to the space variables and of the first order with respect to the variable t.

Remark 3. Moreover, u_n, $\partial u_n/\partial t$, $\partial u_n/\partial x_1, \ldots, \partial u_n/\partial x_N$ being uniformly bounded in $L_2(Q)$, it follows that the sequence $\{u_n(x, t)\}$ is bounded in $W_2^{(1)}(Q)$. Consequently, the sequence $\{u_n(x, t)\}$ being thus compact in $L_2(Q)$, the convergence in $L_2(Q)$ to $u(x, t)$ is the ordinary (not only weak) convergence.

4. MORE SMOOTH SOLUTIONS

Theorem 1 was proved under rather general conditions imposed on the given data of the problem. Imposing further assumptions, stronger results can be obtained.

Let the preceding assumptions be kept and assume, moreover, that

(48) $$f \in W_2^{(2k)} \cap V .$$

and that the operator A (i.j. the functions $a_{ij}(x)$) has such a property that

(48') $$a(v, f) = (v, Af) \quad \text{for all} \quad v \in V .$$

Divide equation (18) by h, denote, as before, $Z_1 = z_1/h$ and put (18) to the form

$$a(v, Z_1) + \frac{1}{h}(v, Z_1 - f) = 0 . \tag{49}$$

Taking (48′) into account, we can write

$$a(v, Z_1 - f) + \frac{1}{h}(v, Z_1 - f) = -(v, Af) .$$

Because $f \in V$, we can put here $v = Z_1 - f$, and noting the fact that $a(Z_1 - f, Z_1 - f) \geqq 0$, we get

$$\|Z_1 - f\|_{L_2(\Omega)} \leqq h\|Af\|_{L_2(\Omega)} . \tag{50}$$

If we subtract (18) from (19) and divide by h, we get, denoting $Z_2 = (z_2 - z_1)/h$ as before,

$$a(v, Z_2) + \frac{1}{h}(v, Z_2) = \frac{1}{h}(v, Z_1) . \tag{51}$$

Subtracting (49) from (51), we have

$$a(v, Z_2 - Z_1) + \frac{1}{h}(v, Z_2 - Z_1) = \frac{1}{h}(v, Z_1 - f) . \tag{52}$$

Putting $v = Z_2 - Z_1$ in (52), we get, because $a(Z_2 - Z_1, Z_2 - Z_1) \geqq 0$,

$$\|Z_2 - Z_1\|_{L_2(\Omega)} \leqq \|Z_1 - f\|_{L_2(\Omega)} \leqq h\|Af\|_{L_2(\Omega)} , \tag{53}$$

and generally,

$$\|Z_j - Z_{j-1}\|_{L_2(\Omega)} \leqq h\|Af\|_{L_2(\Omega)} .$$

This being verified for the division d_1 of the interval $[0, T]$ (see the text following equation (31)), i.e. for $h = h_1$, similar inequalities can be established, in the same way, for the divisions d_n ($n = 2, 3, \ldots$), giving

$$\|Z_j^n - Z_{j-1}^n\|_{L_2(\Omega)} \leqq h_n\|Af\|_{L_2(\Omega)} . \tag{54}$$

From equation (49) it follows, in virtue of (50),

$$\|Z_1\|_V \leqq \text{const. } \|Af\|_{L_2(\Omega)} .$$

In the same way, using (51) and (53), we get

$$\|Z_2\|_V \leqq \text{const. } \|Af\|_{L_2(\Omega)}$$

and quite similarly,

$$\|Z_j^n\|_V \leqq \text{const.}\ \|Af\|_{L_2(\Omega)} \tag{55}$$

in general.

Let us construct, as well as before, the functions

$$u_n(t) = u_n(x, t) = z_j^n(x) + \frac{t - t_j^n}{h_n}\{z_{j+1}^n(x) - z_j^n(x)\} = \tag{56}$$

$$= z_j^n(x) + (t - t_j^n)\, Z_{j+1}^n(x)\,, \quad t_j^n \leqq t \leqq t_{j+1}^n\,.$$

Because of (55) we have $u_n(t) \in C^0([0, T], V)$. Moreover, the set $\{u_n(t)\}$ is a set of functions equicontinuous as mappings from $[0, T]$ into V.

All assumptions of the preceding section being fulfilled, Theorem 1 is valid for the present case. However, in virtue of the supplementary assumptions (48) and (48′) and its consequence (55), the solution $u(t)$ has some other properties in this case:

In virtue of (55), the sequence $\{U_n(t)\}$ (see Section 3, eq. (38)) converges weakly to a function $U(t) \in L_2([0, T], V)$ which can be established, in the same way as in Section 3, to be the derivative $u'(t)$ in $L_2([0, T], V)$ of the function $u(t)$. It follows that

$$u(t) \in C^0([0, T], V)\,. \tag{57}$$

(This result can also be derived in an other way, using the fact that the functions $u_n(t)$ are equicontinuous from $[0, T]$ into V.) If we note that $z_0^n(x) = 0$ and use (55) in the form

$$\|z_j^n(x) - z_{j-1}^n(x)\|_V \leqq \text{const.}\ h_n\|Af\|_{L_2(\Omega)}\,,$$

we easily get

$$u(0) = 0$$

in $C^0([0, T], V)$ (not only in $C^0([0, T], L_2(\Omega))$ as in Section 3).

Moreover, for every $t \in [0, T]$, the sequence $\{u_n(t)\}$ is uniformly bounded in V and, consequently, compact in $W_2^{(k-1)}(\Omega)$. The set of functions $u_n(t)$ being equicontinuous from $[0, T]$ into V, it easily follows that the sequence $\{u_n(t)\}$ converges in $W_2^{(k-1)}(\Omega)$ (not only weakly) to $u(t)$ uniformly in the interval $[0, T]$.

Thus, in consequence of the supplementary assumption (48), the solution $u(t)$ has much stronger properties than before.

Imposing further assumptions on the given data of the problem, it is possible to get more and more strong results. If we require, for example, that the given problem satisfies the assumptions of $2k$-regularity (these assumptions concern, roughly speaking, some smoothness of a_{ij} and $\dot{\Omega}$; for details see Nečas [1]), then the solution obtained by the method explained above, satisfies the given differential equation almost everywhere in Q.

By imposing further assumptions, it is possible to come, in this way, to the classical solution of the problem.

Let an initial condition $u_0(x)$ be given. Let us first assume that

(58) $$u_0(x) \in W_2^{(2k)}(\Omega) \cap V$$

and that the coefficients of the operator A are such that

(59) $$Az \in L_2(\Omega) \quad \text{and} \quad a(v, z) = (v, Az)$$

holds for every $v \in V$ and every $z \in W_2^{(2k)}(\Omega) \cap V$. Especially, for u_0 satisfying (58), we have

(59′) $$Au_0 \in L_2(\Omega) \quad \text{and} \quad a(v, u_0) = (v, Au_0)\,.$$

Substitute $u(x, t) = u_0(x) + \tilde{u}(x, t)$. This substitution can be performed either in the given equation (15), giving

$$A\tilde{u} + \frac{\partial \tilde{u}}{\partial t} = f - Au_0\,,$$

or directly in (18) (substituting $z_1 = u_0 + \tilde{z}_1$), giving

$$a(v, \tilde{z}_1) + \frac{1}{h}(v, \tilde{z}_1) = (v, f) - a(v, u_0)\,,$$

or, in virtue of (59′),

$$a(v, \tilde{z}_1) + \frac{1}{h}(v, \tilde{z}_1) = (v, f) - (v, Au_0)\,.$$

Similarly, the substitution $z_2 = u_0 + \tilde{z}_2$ in (19) gives

$$a(v, \tilde{z}_2) + \frac{1}{h}(v, \tilde{z}_2) = (v, f) - (v, Au_0) + \frac{1}{h}(v, \tilde{z}_1)\,, \text{ etc.}$$

Thus, if (58), (59) are fulfilled, the problem is reduced to the problem treated above, with $g = f - Au_0 \in L_2(\Omega)$ instead of $f \in L_2(\Omega)$.

Let us examine a more general case. For $u_0(x) \neq 0$, equations (18)–(21) become

(60) $$a(v, z_1) + \frac{1}{h}(v, z_1) = (v, f) + \frac{1}{h}(v, u_0)\,,$$

$$a(v, z_2) + \frac{1}{h}(v, z_2) = (v, f) + \frac{1}{h}(v, z_1)\,,$$

etc. We have, putting subsequently $v = z_1, v = z_2$, etc.

$$\|z_1\|_{L_2(\Omega)} \leqq \|u_0\|_{L_2(\Omega)} + h\|f\|_{L_2(\Omega)},$$

$$\|z_2\|_{L_2(\Omega)} \leqq \|z_1\|_{L_2(\Omega)} + h\|f\|_{L_2(\Omega)} \leqq \|u_0\|_{L_2(\Omega)} + 2h\|f\|_{L_2(\Omega)},$$

etc. Thus we have (clearly independently of the division d_n)

$$\|z_j^n\|_{L_2(\Omega)} \leqq \|u_0\|_{L_2(\Omega)} + T\|f\|_{L_2(\Omega)}$$

and, according to (32) taking the form of the functions $u_n(x, t)$ into account,

$$\|u_n(x, t)\|_{L_2(\Omega)} \leqq \|u_0\|_{L_2(\Omega)} + T\|f\|_{L_2(\Omega)}$$

for all $t \in [0, T]$. Especially, for ${}^1u_n(x, t)$, ${}^2u_n(x, t)$, corresponding to the same f and to different initial conditions ${}^1u_0(x)$, ${}^2u_0(x)$, we have

$$\text{(61)} \qquad \|{}^2u_n(x, t) - {}^1u_n(x, t)\|_{L_2(\Omega)} \leqq \|{}^2u_0(x) - {}^1u_0(x)\|_{L_2(\Omega)}$$

for every $t \in [0, T]$. Making the square and integrating between 0 and T, we get

$$\text{(61')} \qquad \|{}^2u_n(x, t) - {}^1u_n(x, t)\|_{L_2([0,T], L_2(\Omega))} \leqq \sqrt{(T)}\ \|{}^2u_0(x) - {}^1u_0(x)\|_{L_2(\Omega)}.$$

These rather simple inequalities yield a number of important consequences:

First, let ${}^1u_0(x)$, ${}^2u_0(x)$ satisfy (58) and let A satisfy (59), so that the problems, corresponding to initial conditions ${}^1u_0(x)$, ${}^2u_0(x)$ can be converted into problems with zero initial conditions in the way explained at the beginning of this section. Let ${}^1u(t)$, ${}^2u(t)$ be solutions (in the sense of Theorem 1), corresponding to ${}^1u_0(x)$, ${}^2u_0(x)$, respectively. Having in mind that these solutions are weak limits of the sequences $\{{}^1u_n(t)\}$, $\{{}^2u_n(t)\}$, respectively, in $L_2([0, T], V)$, and consequently in $L_2([0, T], L_2(\Omega))$, we get, by (61')

$$\text{(62)} \qquad \|{}^2u(t) - {}^1u(t)\|_{L_2([0,T], L_2(\Omega))} \leqq \sqrt{(T)}\ \|{}^2u_0(x) - {}^1u_0(x)\|_{L_2(\Omega)}.$$

Now, let (59) be fulfilled and let $\{{}^iu(x)\}$ be a sequence of functions satisfying (58) and converging in $L_2(\Omega)$ to the given initial condition $u_0(x) \in L_2(\Omega)$. Let $\{{}^iu(t)\}$ be the sequence of corresponding solutions (in the sense of Theorem 1). By (62), $\{{}^iu(t)\}$ is a Cauchy sequence in $L_2([0, T], L_2(\Omega))$. Let $u(t)$ be its limit in this space, i.e.

$$\text{(63)} \qquad \lim_{i\to\infty} {}^iu(t) = u(t) \quad \text{in} \quad L_2([0, T], L_2(\Omega)).$$

(62) implies that, $u_0(x) \in L_2(\Omega)$ being given, the limit $u(t)$ is independent of the choice of the sequence $\{{}^iu(x)\}$ converging to $u_0(x)$ in $L_2(\Omega)$, if only this sequence satisfies the above required assumption, concerning condition (58). We shall call the function (63) a *generalized solution* of the problem (15)–(17).

If $u_0(x)$ satisfies (58), this generalized solution becomes the weak solution in the sense of Theorem 1.

Remark 4. The just defined generalized solution has a number of properties analogous to the properties of the weak solution of Theorem 1 or of that treated in [7]. These properties can be derived either by a more detailed treatment of the limit process (63) or by methods similar to those used in the work [7]. We shall state here only these which are of importance in the following text:

Let (59) be satisfied, let ${}^1u(x)$, ${}^2u(x)$ satisfy (58) and let ${}^1u(t)$, ${}^2u(t)$ be the corresponding solutions (thus having properties stated in Theorem 1). It follows immediately from (61):

$$\|{}^2u(t) - {}^1u(t)\|_{C^0([0,T],L_2(\Omega))} \leqq \|{}^2u(x) - {}^1u(x)\|_{L_2(\Omega)} . \tag{64}$$

Consequently, if $\{{}^iu(x)\}$ is a Cauchy sequence in $L_2(\Omega)$ (each of the functions ${}^iu(x)$ satisfying (58)), then the sequence $\{{}^iu(t)\}$ of the corresponding solutions is a Cauchy sequence in $C^0([0, T], L_2(\Omega))$. Thus:

Theorem 2. *The generalized solution is a function of* $C^0([0, T], L_2(\Omega))$.

Remark 5. Inequalities (62) or (64) express *continuous dependence of the solution on initial conditions* in $L_2([0, T], L_2(\Omega))$, and in $C^0([0, T], L_2(\Omega))$, respectively.

Remark 6. The generalized solution $u(t) = u(x, t)$ is a weak limit, in $L_2([0, T], L_2(\Omega))$ of the Rothe sequence $\{u_n(t)\} = \{u_n(x, t)\}$, corresponding to the given initial condition $u_0(x) \in L_2(\Omega)$.

To establish this assertion, it is sufficient to construct a sequence of functions ${}^iu(x)$, satisfying (58) and converging in $L_2(\Omega)$ to $u_0(x)$, then to write the difference $u(t) - u_n(t)$ in the form

$$u - u_n = (u - {}^iu) + ({}^iu - {}^iu_n) + ({}^iu_n - u_n) \tag{65}$$

and to take into account that, in $L_2([0, T], L_2(\Omega))$, ${}^iu$ converges to u, ${}^iu_n$ converges weakly to ${}^iu$ and the difference ${}^iu_n - u_n$ can be made arbitrarily small by (61′) if i is sufficiently large.

Moreover, we can state a stronger result, giving the ordinary (not only weak) convergence:

Theorem 3. *Let, for the problem* (15)–(17), *the assumptions of Theorem* 1 *and the condition* (59) *be fulfilled, let* $u_0(x) \in L_2(\Omega)$. *Let* $u_n(x, t)$ *be the Rothe functions* (32) *with* $z_0^n(x) = u_0(x)$, $u(x, t)$ *the generalized solution of the problem* (defined by (63)). *Then*

$$\lim_{n\to\infty} u_n(x, t) = u(x, t) \quad \text{in} \quad L_2(Q) .$$

Proof. We make use of Remark 3 (p. 330). All functions ${}^iu$ in (63) lying in $L_2(Q)$, we have by (63)

$$u(x, t) \in L_2(Q) .$$

Then our assertion follows by (65), because, in $L_2(Q)$, ${}^iu$ converges to u, ${}^iu_n$ converges to ${}^iu$ (according to Remark 3) and ${}^iu_n - u_n$ is arbitrarily small if i is sufficiently large.

The investigations of the next section are immediate consequences of the results of this section, especially of Theorem 3.

6. THE CONVERGENCE OF THE METHOD

In this section, we assume that $a_{ij} = a_{ji}$ in (10) (the symmetry of the problem) and that the operator A satisfies the condition (59). The problem (15)–(17) being symmetric, the solution of the problems (60) is equivalent to the problem of finding, in V, minima of the functionals

(66)
$$G_1(z) = a(z, z) + \frac{1}{h}(z, z) - 2(z, f) - \frac{2}{h}(z, u_0),$$

$$G_2(z) = a(z, z) + \frac{1}{h}(z, z) - 2(z, f) - \frac{2}{h}(z, z_1),$$

. . . .

$$G_k(z) = a(z, z) + \frac{1}{h}(z, z) - 2(z, f) - \frac{2}{h}(z, z_{k-1}),$$

$$G_r(z) = a(z, z) + \frac{1}{h}(z, z) - 2(z, f) - \frac{2}{h}(z, z_{r-1}),$$

where r is the number of subintervals of the interval $[0, T]$ (so that $rh = T$) and z_1, z_2 are minimizing elements (in V) of functionals $G_1(z)$, $G_2(z)$, etc. To these elements $z_1(x), z_2(x), \ldots, z_r(x)$ there corresponds a "Rothe function" of the type (28) (with $z_0(x) = u_0(x)$) which we denote here by

(67)
$$_ru(x, t).$$

Let z_1, z_2, etc., are determined approximately, using one of the usual direct methods. To make clear our idea, let us consider the Ritz method; of course, whichever other method can be chosen, having similar properties, concerning approximations investigated below.

Thus, let

$$\varphi_1(x), \varphi_2(x), \ldots, \varphi_j(x), \ldots$$

be a base in V and denote

(68)
$$z_k^{s_k} = \sum_{j=1}^{s_k} a_j^k \varphi_j(x).$$

(The suffix k corresponds to the functional G_k.) Choose s_1, substitute $z_1^{s_1}$ for z into $G_1(z)$ and determine a_j^1 $(j = 1, \ldots, s_1)$ from the condition that $G_1(z_1^{s_1})$ be minimal. It is well known that this problem leads to the solution of a system of linear algebraic equations which is ($a(v, v)$ being V-elliptic and h being positiv) uniquely solvable. Thus, s_1 being fixed, $z_1^{s_1}$ is uniquely determined.

Substitute $z_1^{s_1}$ for z_1 into $G_2(z)$. Let us denote this new functional by $\tilde{G}_2(z)$. Choosing s_2 and putting $z_2^{s_2}$ instead of z into $\tilde{G}_2(z)$, we get, in the same way as before, a_j^2 $(j = 1, \ldots, s_2)$ such that $z_2^{s_2}$ makes $\tilde{G}_2(z)$ minimal. In general, having $z_{k-1}^{s_{k-1}}$, we denote $G_k(z)$, with $z_{k-1}^{s_{k-1}}$ instead of z_{k-1}, by $\tilde{G}_k(z)$, choose s_k and determine uniquely $z_k^{s_k}$ to make the functional $\tilde{G}_k(z)$ minimal. In this way, we get the functions

$$z_1^{s_1}, z_2^{s_2}, \ldots, z_r^{s_r} . \tag{69}$$

Remark 7. As said above, the individual steps to determine the functions (69) lie in solving systems of linear equations for a_j^k $(j = 1, \ldots, s_k)$. Thus, in virtue of our method, just described, the (approximate) numerical solution of our problem is converted into the solution of a finite number of linear algebraic systems. Thus the terminology "direct method" is justified. Note that in converting the functionals $G_k(z)$ into $\tilde{G}_k(z)$, the terms $a(z, z)$, $(1/h)(z, z)$ remain unchanged. Consequently, if $s_1 = s_2 = \ldots = s_r$ is chosen, the left hand sides of all the systems are the same, so that in this case the numerical process is particularly simple, especially if a computer is at hand.

Now, let the functions (69) be found. Let us construct for them the corresponding "Rothe function" in the same way as we have constructed the function ${}_r u(x, t)$ for the functions $z_1(x), \ldots, z_r(x)$, with $z_k^{s_k}(x)$ substituted for $z_k(x)$. Denote this function by

$${}^{s_1, \ldots, s_r}_{\quad\; r}u(x, t) . \tag{70}$$

A question arises, of course, how "close" is the function (70) to the function (67) and, in particular, to the solution of our problem. We shall show that

$$\left\| u(x, t) - {}^{s_1, \ldots, s_r}_{\quad\; r}u(x, t) \right\|_{L_2(Q)} \tag{71}$$

can be made arbitrarily small if r and $s_1, \ldots, s_r$ are sufficiently large.

Thus, let $\varepsilon > 0$ be given. Denote, as before, $d_1, d_2, \ldots$ the sequence of divisions of the interval $[0, T]$ into subintervals of the lengths $h_1, h_2, \ldots$ (cf. the text following equation (31)). According to Theorem 3 (p. 335), $\varepsilon/2$ being given, it is possible to find n_0 such that for every $n > n_0$ the inequality

$$\left\| u(x, t) - u_n(x, t) \right\|_{L_2(Q)} < \frac{\varepsilon}{2} \tag{72}$$

is satisfied.

Thus, let such an n be fixed. Denote briefly the corresponding h_n by h and let $r = T/h$. So we are in the notation used in (66) with h and r fixed and with (67) satisfying

$$\|u(x, t) - {}_r u(x, t)\|_{L_2(Q)} < \frac{\varepsilon}{2}. \tag{73}$$

Let us construct the functions (69) and (70) in the way explained in the text following (68). We have to show that it is possible to choose the numbers $s_1, \ldots, s_r$ in such a way that

$$\|{}_r u(x, t) - {}^{s_1,\ldots,s_r}_{\quad r} u(x, t)\|_{L_2(Q)} < \frac{\varepsilon}{2}. \tag{74}$$

Taking the form of the functions ${}_r u(x, t)$, ${}^{s_1,\ldots,s_r}_{\quad r} u(x, t)$ into account, it is sufficient to show that

$$\|z_k(x) - z_k^{s_k}(x)\|_{L_2(\Omega)} < \frac{\varepsilon}{2\sqrt{T}} \tag{75}$$

for every $k = 1, \ldots, r$.

Let s_1 be sufficiently large so that the difference between $z_1(x)$ and $z_1^{s_1}(x)$ be smaller than δ_1 in $L_2(\Omega)$. (This can be always realized, even in V.) If $z_2(x)$, or $z_2^*(x)$ are minimizing elements of functionals $G_2(z)$, or $\tilde{G}_2(z)$, respectively, then, according to (61),

$$\|z_2^*(x) - z_2(x)\|_{L_2(\Omega)} < \delta_1 . \tag{76}$$

Similarly, if s_2 is sufficiently large so that the difference between $z_2^*(x)$ and $z_2^{s_2}(x)$ is smaller than δ_2 in $L_2(\Omega)$, then

$$\|z_3^*(x) - z_3(x)\|_{L_2(\Omega)} < \delta_1 + \delta_2 ,$$

where $z_3(x)$, or $z_3^*(x)$ are minimizing elements of the functionals $G_3(z)$, or $\tilde{G}_3(z)$, respectively. Going on in this way, we come to the following conclusion: Let $s_1, \ldots, s_r$ be chosen in such a way that for every $k = 1, \ldots, r$ we have

$$\|z_k^*(x) - z_k^{s_k}(x)\|_{L_2(\Omega)} < \frac{\varepsilon}{2r\sqrt{T}} \tag{77}$$

(while $z_1^*(x) = z_1(x)$, of course). Then (74) is fulfilled. Coming back to (73) and summarizing, we have: If, in the Rothe method, n is sufficiently large and if, in minimizing functionals $\tilde{G}_k(z)$, a sufficiently number of terms in linear combination (68) is taken, then the difference between the generalized solution of our problem and the Rothe function ${}^{s_1,\ldots,s_r}_{\quad r} u(x, t)$, constructed by the above described direct method, can be made arbitrarily small in $L_2(Q)$.

This result is briefly expressed in the following

Theorem 4. *The direct method, described above, is convergent.*

Remark 8. It follows from the procedure described in this section that the convergence of our direct method can be examined also in other metrics than in $L_2(Q)$. Especially, if (59) is fulfilled and if f and u_0 are so smooth that $f(x) - Au_0(x) \in \in W_2^{(2k)} \cap V$, then, according to Section 4 and the preceding section, the convergence can be examined, for example, in $C^0([0, T], W_2^{k-1} \cap V)$. If the given data of the problem are so smooth that the (not only weak) convergence of the Rothe sequence $\{u_n(x, t)\}$ to $u(x, t)$ in a very „smooth" metric is justified, it is preferable, in order to improve the convergence of the method, to use functionals of a more fine structure than are functionals of the type (66) (for example, to use functionals of types applied in the so-called Courant method, or others).

Remark 9. Obviously, the just described method can be generalized, without essential difficulties, for time-dependent operators and for problems other than the Dirichlet problem.

References

[1] *Nečas, J.:* Les méthodes directes en théorie aux équations elliptique. Edition de l'Academie tchécoslovaque des sciences, Praha 1967.

[2] *Hlaváček, I.:* Variational Principles for Parabolic equations. Aplikace matematiky *14* (1969), 4, Praha 1969.

[3] *Wilcox, C. H.:* Linear hyperbolic equations. Arch Rational Mech. and Anal., Vol. *10*, Number 5, 1962.

[4] *Rothe, E.:* Zweidimensionale parabolische Randwertaufgaben als Grenzfall eindimensionaler Randwertaufgaben. Math. Ann. *102* (1930).

[5] *Ladyženskaja O. A.:* Solution of the First Boundary Value Problem for Quasilinear Parabolic Equations. Trudy Mosk. Mat. Obš. *7* (1958). (In Russian.)

[6] *Iljin, A. M., Kalašnikov, A. S., Oleinik, O. A.:* Linear Equations of the Second Order of Parabolic Type. Uspechi *XVII* (1962), 3. (In Russian.)

[7] *Ladyženskaja, O. A.:* On the Solution of Nonstationary Operator Equations. Mat. Sbornik *39* (*81*), No. 4, 1956. (In Russian.)

[8] *Ibragimov, Š. I.:* On the Analogy of the Method of Lines for Differential Equations in Abstract Spaces. Dokl. Azerb. Ak. Nauk *XXI*, No. 6, 1965. (In Russian.)

Author's address: Praha 2, Trojanova 13, ČSSR (České vysoké učení technické).