Ivan Hlaváček

On a semi-variational method for parabolic equations. I

Aplikace matematiky, Vol. 17 (1972), No. 5, 327–351

Persistent URL: http://dml.cz/dmlcz/103427

ON A SEMI-VARIATIONAL METHOD FOR PARABOLIC EQUATIONS I

Ivan Hlaváček

(Received July 12, 1971)

INTRODUCTION

The popularity of variational methods has grown during the past decade mostly due to the finite element method, which combines the versatility of the variational approach with the computational advantage of band matrices, the main feature of finite-difference procedures. The present paper aims at a further development of the finite element technique, when applied to mixed problems for parabolic equations.

Much work has been done on a special Galerkin-type procedure, derived by an analogy with the well-known Crank-Nicholson finite difference scheme [1], [4]. We propose here a sequence of approximations with gradually increased accuracy in time, generalizing the Crank-Nicholson-Galerkin procedure.

In Section 1. a heuristic approach is shown for the derivation of the Crank-Nicholson-Galerkin procedure and then applied to derive the second and third approximation. In Section 2 we prove the convergence of the second approximation and its stability with respect to the initial condition. In Section 3 we show that, in case of ordinary differential equations, the sequence of approximations leads to the sequence of Padé approximations of the exponential function. In Section 4 a numerical example is presented for a parabolic equation.

In Part II of this paper some further properties of the semi-variational approximations will be shown, namely their independence of the choice of polynomial basis in time and a close relation between them and the Padé approximations for general n. Moreover, the cases of inhomogeneous mixed problems and some more general abstract equations will be considered.

1. SEMI-VARIATIONAL APPROXIMATIONS OF SOLUTION

In the present section an algorithmus will be presented, which enables to construct a sequence of numerical procedures for approximate solutions of the given problem

(1.1), (1.2). The first approximation coincides with the well-known Crank-Nicholson-Galerkin procedure and consequently it is second order correct in time [1]. The second approximation is shown to be fourth order correct in time (in Section 2) and the third approximation may be expected sixth order correct, according to the case of ordinary differential equations. Each approximation will be derived in two steps, both of the projection type. As the first step is based on a variational formulation of the problem under consideration, we call the method semi-variational.

In order to explain briefly the main features of the derivation, let us consider an ordinary differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}t} + Ay = 0\,, \quad A = \text{const (real)}\,, \quad 0 \leqq t \leqq \tau\,,$$

with the initial condition

$$y(0) = \varphi_0\,.$$

Let us set

$$y^{(n-1)}(t) = \sum_{i=0}^{n-1} \alpha_i N_i^{(n-1)}(t)\,,$$

n positive integer, where $\{N_i^{(n-1)}(t)\}_0^{n-1}$ denotes the Lagrangian interpolation polynomial basis of polynomials of degree $n - 1$, i.e.,

$$N_i^{(n-1)}\left(\frac{k\tau}{n-1}\right) = \delta_{ik}\,, \quad 0 \leqq i, k \leqq n - 1\,,$$

δ_{ik} is the Kronecker's delta, $N_i^{(n-1)}(t)$ are polynomials of degree $n - 1$.

The coefficients $\{\alpha_i\}_0^{n-1}$ will be determined by means of the variational condition

$$\int_0^\tau \left(y^{(n-1)}(t) + \int_0^t A\, y^{(n-1)}(z)\,\mathrm{d}z - \varphi_0\right) \delta y^{(n-1)}(\tau - t)\,\mathrm{d}t = 0\,.$$

This condition follows from the "integral convolution principle" (see [9], [2] – Th. 1 or [3]): if we define a functional

$$\mathscr{F}(y) = \int_0^\tau \left(y(t) + \int_0^t A\, y(z)\,\mathrm{d}z - 2\varphi_0\right) y(\tau - t)\,\mathrm{d}t$$

on a sufficiently large class $\mathscr{K}$ of admissible functions, then $\mathscr{F}(y)$ attains its stationary value on $\mathscr{K}$, if and only if y is a solution to the initial-value problem under consideration.

The approximation $y^{(n-1)}$, however, does not satisfy the initial condition exactly. Therefore we construct another approximation

$$y^{(n)}(t) = \sum_{i=0}^{n} \beta_i N_i^{(n)}(t)\,,$$

which is polynomial of degree n, satisfies the initial condition

$$y^{(n)}(0) = \varphi_0$$

and the following conditions

$$\int_0^\tau [y^{(n)}(t) - y^{(n-1)}(t)]\, N_i^{(n-1)}(t)\, \mathrm{d}t = 0\,, \quad i = 0, 1, \ldots, n-1\,.$$

These group of n equations means that the $L_2(0, \tau)$ projection of $y^{(n)}$ into the subspace of polynomials of degree at most $n - 1$ coincides with $y^{(n-1)}$. The function $y^{(n)}$ will be referred to as the n-th semi-variational approximation on $\langle 0, \tau \rangle$.

In the following we shall apply the main idea of the derivation to an abstract parabolic equation in a Hilbert space and construct the first three approximations (for $n = 1, 2, 3$) in detail.

Let a real Hilbert space H with the scalar product (u, v) and the norm $|u| = (u, u)^{1/2}$ be given. Let us consider the equation

$$\text{(1.1)} \qquad \frac{\mathrm{d}u}{\mathrm{d}t} = A\, u(t) = f(t)\,, \quad 0 < t \leqq T$$

with the initial condition

$$\text{(1.2)} \qquad u(0) = \varphi_0\,,$$

where $u(t)$ and $f(t)$ are mappings of the interval $\langle 0, T \rangle$ into H, $\varphi_0 \in H$, A is a linear symmetric and positive definite operator in H, which does not depend on t.

Assume, that a Hilbert space $\mathscr{V}$ with the norm $\|u\|$, a bilinear form $[u, v]_A$ continuous and symmetric on $\mathscr{V} \times \mathscr{V}$ and positive constants c, α, C_0 exist such that

$$\text{(1.3)} \qquad \mathscr{V} \subset H\,, \quad u \in \mathscr{V} \Rightarrow \|u\| \geqq c|u|\,,$$

the domain $D(A)$ of the operator A is a subset of $\mathscr{V}$,

$$\text{(1.4)} \qquad \begin{gathered} u, v \in D(A) \Rightarrow (Au, v) = [u, v]_A \\ u \in \mathscr{V} \Rightarrow \alpha\|u\|^2 \leqq [u, u]_A \leqq C_0\|u\|^2\,. \end{gathered}$$

First approximation

Consider a finite-dimensional subspace $\mathscr{M}$ of $\mathscr{V}$, spanned by elements $v_1, v_2, v_3, \ldots$ $\ldots, v_N$. Let us have a fixed $\tau > 0$ and set

$$\text{(1.5)} \qquad u^{(0)} = N^{(0)}(t)\, V_0\,, \quad 0 \leqq t \leqq \tau\,,$$

where

$$N^{(0)}(t) = 1\,, \quad V_0 = \sum_{i=1}^{N} a_i v_i \in \mathscr{M}\,.$$

Let the function $f(t)$ be approximated similarly by

$$f^{(0)}(t) = N^{(0)}(t)\,\tfrac{1}{2}[f(0) + f(\tau)]\,. \tag{1.6}$$

The formula (1.6) may be obtained setting

$$\begin{gathered} f^{(0)}(t) = \text{const.}\,, \\ f^{(1)}(t) = N_0^{(1)}(t)\,f(0) + N_1^{(1)}(t)\,f(\tau)\,, \\ N_0^{(1)}(t) = 1 - t/\tau\,, \quad N_1^{(1)}(t) = t/\tau \end{gathered} \tag{1.7}$$

and

$$\int_0^\tau (f^{(1)} - f^{(0)})\,N^{(0)}(t)\,\mathrm{d}t = 0\,. \tag{1.8}$$

In order to determine the coefficients a_i of V_0, we use the variational condition

$$\int_0^\tau u^{(0)} + \int_0^t A\,u^{(0)}(z)\,\mathrm{d}z - \varphi_0 - \int_0^t f^{(0)}(z)\,\mathrm{d}z\,, \quad \delta u^{(0)}(\tau - t)\,\mathrm{d}t = 0 \tag{1.9}$$

(see [2] – Th. 1). The term with the operator A will be replaced by

$$\int_0^\tau \left[\int_0^t u^{(0)}(z)\,\mathrm{d}z\,, \quad \delta u^{(0)}(\tau - t)\right]_A \mathrm{d}t\,,$$

so that $u^{(0)}(z)$ need not belong to the domain $D(A)$, but to $\mathscr{V}$ only (cf. also [3]).

Inserting

$$\delta u^{(0)}(\tau - t) = v_j\,, \quad j = 1, 2, \ldots, N \tag{1.10}$$

and integrating, we obtain the following system of linear equations for a_i

$$\tau \sum_{i=1}^N a_i(v_i, v_j) + \frac{\tau^2}{2} \sum_{i=1}^N a_i[v_i, v_j)]_A = \tau(\varphi_0, v_j) + \frac{\tau^2}{4}(f(0) + f(\tau), v_j)\,. \tag{1.11}$$

Let us denote

$$\begin{gathered} (v_i, v_j) = G_{ij}\,, \quad [v_i, v_j]_A = \mathscr{A}_{ij}\,, \\ (\varphi_0, v_j) = \omega_{0j} \quad (f(t), v_j) = F_j(t)\,, \\ i, j = 1, 2, \ldots, N\,. \end{gathered} \tag{1.12}$$

Then the system (1.11) may be rewritten in the matrix form[1])

$$\left(G + \frac{\tau}{2}\mathscr{A}\right)\boldsymbol{a} = \omega_0 + \frac{\tau}{4}(\mathbf{F}(0) + \mathbf{F}(\tau))\,. \tag{1.13}$$

---

[1]) We denote

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}, \quad \omega_0 = \begin{pmatrix} \omega_{01} \\ \vdots \\ \omega_{0N} \end{pmatrix}, \quad \mathbf{F}(t) = \begin{pmatrix} F_1(t) \\ \vdots \\ F_N(t) \end{pmatrix}, \quad \bar{\boldsymbol{u}}_1 = \begin{pmatrix} \bar{u}_{11} \\ \vdots \\ \bar{u}_{1N} \end{pmatrix}.$$

The second step starts with the replacing $u^{(0)}(t)$ by the linear approximation

$$u^{(1)}(t) = U_0 + N_1^{(1)}(t)\,\bar{U}_1\,, \tag{1.14}$$

where

$$U_0 = \sum_{i=1}^{N} w_{0i} v_i\,, \quad \bar{U}_1 = \sum_{i=1}^{N} \bar{u}_{1i} v_i\,. \tag{1.15}$$

The initial condition (1.2) is now employed in the projection form

$$(U_0, v_j) = (\varphi_0, v_j)\,, \quad j = 1, 2, \ldots, N\,, \tag{1.16}$$

which may be rewritten as follows

$$G\mathbf{w}_0 = \boldsymbol{\omega}_0\,. \tag{1.17}$$

The coefficients $\bar{u}_{1i}$ will be determined from the projection condition

$$\int_0^\tau (u^{(1)} - u^{(0)}\,, \quad \delta u^{(0)})\,\mathrm{d}t = 0\,. \tag{1.18}$$

Inserting (1.14), (1.5) and (1.10), we obtain

$$\int_0^\tau (U_0 + N_1^{(1)}(t)\,\bar{U}_1, v_j)\,\mathrm{d}t = \int_0^\tau (V_0, v_j)\,\mathrm{d}t\,.$$

The integration leads to the system

$$\tau(U_0, v_j) + \frac{\tau}{2}(\bar{U}_1, v_j) = \tau(V_0, v_j)\,,$$

which may be rewritten in the following matrix form

$$\tfrac{1}{2}G\bar{\mathbf{u}}_1 = G\mathbf{a} - G\mathbf{w}_0\,, \tag{1.19}$$

if we use also (1.17). From (1.19) we conclude

$$\mathbf{a} = \mathbf{w}_0 + \tfrac{1}{2}\bar{\mathbf{u}}_1\,. \tag{1.20}$$

If we substitute for $\mathbf{a}$ and ω_0 in (1.13), we obtain

$$G\left(\frac{1}{\tau}\bar{\mathbf{u}}_1\right) + \mathscr{A}(\mathbf{w}_0 + \tfrac{1}{2}\bar{\mathbf{u}}_1) = \tfrac{1}{2}(\mathbf{F}(0) + \mathbf{F}(\tau))\,,$$

which is equivalent to

$$\left(\frac{1}{\tau}(u^{(1)}(\tau) - U_0), v_j\right) + \tfrac{1}{2}[u^{(1)}(\tau) + U_0, v_j]_A = \tfrac{1}{2}(f(0) + f(\tau), v_j)\,. \tag{1.21}$$

Obviously, the formula (1.21) is just that of Crank-Nicholson-Galerkin approximation[1]) [1], [4] for the first step.

If the coefficients $\mathbf{w}_m$ of the expansion

$$u^{(1)}(m\tau) = \sum_{i=1}^{N} w_{mi} v_i$$

are known, the next step is to solve the system

$$\left(G + \frac{\tau}{2} \mathscr{A}\right) \boldsymbol{a}_m = G\mathbf{w}_m + \frac{\tau}{2} \mathbf{F}_m^{(0)}, \tag{1.22}$$

where

$$F_{mj}^{(0)} = \tfrac{1}{2}(f(m\tau) + f(m\tau + \tau), v_j).$$

Then the coefficients of $u^{(1)}(m\tau + \tau)$ are given by the formula

$$\mathbf{w}_{m+1} = 2\boldsymbol{a}_m - \mathbf{w}_m .$$

Second approximation

Let us apply the approach, which has been used to the derivation of the first approximation, to polynomials of the first and second degree in t, instead of those of zero and first degree, respectively. Thus we set

$$u^{(1)} = N_0^{(1)}(t)\, V_0 + N_1^{(1)}(t)\, V_1 , \tag{1.23}$$

where $N_0^{(1)}$, $N_1^{(1)}$ are given in (1.7),

$$V_0 = \sum_{i=1}^{N} a_{0i} v_i , \quad V_1 = \sum_{i=1}^{N} a_{1i} v_i \tag{1.24}$$

and v_i, $(i = 1, 2, \ldots, N)$ is the base of the subspace $\mathscr{M}$ of $\mathscr{V}$.

The function $f(t)$ will be approximated similarly by a linear function

$$f^{(1)}(t) = N_0^{(1)}(t) \tilde{f}_0 + N_1^{(1)}(t) \tilde{f}_1 , \tag{1.25}$$

where

$$\tilde{f}_0 = \frac{1}{3}\left[2f(0) + 2f\left(\frac{\tau}{2}\right) - f(\tau)\right], \tag{1.26}$$

$$\tilde{f}_1 = \frac{1}{3}\left[-f(0) + 2f\left(\frac{\tau}{2}\right) + 2f(\tau)\right].$$

---

[1]) Some authors replace the average in the definition of $f^{(0)}$ (1.6) by f evaluated at $t = \tau/2$. The error estimation, however, remains of the same order in t.

The formulas (1.26) may be obtained setting

$$(1.27)\qquad f^{(2)} = N_0^{(2)}(t)\, f(0) + N_{1/2}^{(2)}(t)\, f\left(\frac{\tau}{2}\right) + N_1^{(2)}(t)\, f(\tau)\,,$$

$$N_0^{(2)}(t) = 1 - 3t/\tau + 2(t/\tau)^2\,,$$

$$N_{1/2}^{(2)}(t) = 4(t/\tau)\,(1 - t/\tau)\,, \quad N_1^{(2)}(t) = 2(t/\tau)^2 - t/\tau$$

(Lagrange parabolic interpolation) and

$$\int_0^\tau (f^{(2)} - f^{(1)})\, N_k^{(1)}(t)\, \mathrm{d}t = 0\,, \quad k = 0, 1\,.$$

Thus $f^{(1)}(t)$ is the $L_2(0, \tau)$ projection of the Lagrangian second degree interpolate of $f(t)$ into the subspace of linear functions.

In order to determine the coefficients a_{0i}, a_{1i} of V_0 and V_1, we use the variational condition

$$(1.28)\qquad \int_0^\tau \left(\left(u^{(1)}(t) + \int_0^t A\, u^{(1)}(z)\, \mathrm{d}z - \varphi_0 - \int_0^t f^{(1)}(z)\, \mathrm{d}z\,,\ \delta u^{(1)}(\tau - t)\right)\right) \mathrm{d}t = 0$$

and replace the term with the operator A by

$$\int_0^\tau \left[\int_0^t u^{(1)}(z)\, \mathrm{d}z\,,\ u^{(1)}(\tau - t) \right]_A \mathrm{d}t\,.$$

Remark I.1.1. If some non-homogeneous boundary conditions, associated with the differential operator A, are assigned by means of a function $g(t)$ on a part Γ_g of the boundary, we proceed with $g(t)$ in a manner quite similar to that used with $f(t)$, adding an integral of the form

$$-\int_0^\tau \int_{\Gamma_g} \int_0^t g^{(1)}(z)\, \mathrm{d}z\, \delta u^{(1)}(\tau - t)\, \mathrm{d}\Gamma\, \mathrm{d}t$$

on the left-hand side of (1.28) (cf. "α-integral convolution principle" in [3]). We shall consider such cases in Part II. thoroughly.

We have

$$(1.29)\qquad \delta u^{(1)}(t) = N_k^{(1)}(t)\, v_j\,; \quad k = 0, 1\,; \quad j = 1, 2, \ldots, N$$

$$N_0^{(1)}(\tau - t) = N_1^{(1)}(t)\,, \quad N_1^{(1)}(\tau - t) = N_0^{(1)}(t)\,.$$

Inserting (1.29), (1.23), (1.25) and (1.26) into (1.28) and integrating, we are led to the following two systems (for $k = 0$ and $k = 1$)

$$(1.30)\quad \left(\frac{\tau}{3} V_0 + \frac{\tau}{6} V_1, v_j\right) + \left[\frac{\tau^2}{8} V_0 + \frac{\tau^2}{24} V_1, v_j\right]_A = \left(\frac{\tau}{2} \varphi_0 + \frac{\tau^2}{8} \tilde{f}_0 + \frac{\tau^2}{24} \tilde{f}_1, v_j\right),$$

$$(1.31)\quad \left(\frac{\tau}{6} V_0 + \frac{\tau}{3} V_1, v_j\right) + \left[\frac{5}{24} \tau^2 V_0 + \frac{\tau^2}{8} V_1, v_j\right]_A = \left(\frac{\tau}{2} \varphi_0 + \frac{5}{24} \tau^2 \tilde{f}_0 + \frac{\tau^2}{8} \tilde{f}_1, v_j\right).$$

The sum of (1.30) and (1.31) yields

$$\frac{1}{2}(V_0 + V_1, v_j) + \left[\frac{\tau}{3} V_0 + \frac{\tau}{6} V_1, v_j\right]_A = \left(\varphi_0 + \frac{\tau}{3}\tilde{f}_0 + \frac{\tau}{6}\tilde{f}_1, v_j\right) \tag{1.32}$$

and their difference yields

$$(V_1 - V_0, v_j) + \frac{\tau}{2}[V_0 + V_1, v_j]_A = \frac{\tau}{2}(\tilde{f}_0 + \tilde{f}_1, v_j)\,. \tag{1.33}$$

We may insert

$$\tfrac{1}{3}V_0 + \tfrac{1}{6}V_1 = \tfrac{1}{4}(V_1 + V_0) - \tfrac{1}{12}(V_1 - V_0) \tag{1.34}$$

into (1.32) to obtain

$$\frac{1}{2}(V_0 + V_1, v_j) + \frac{\tau}{2}\left[\frac{1}{2}(V_1 + V_0) - \frac{1}{6}(V_1 - V_0), v_j\right]_A = \tag{1.35}$$

$$= \left(\varphi_0 + \frac{\tau}{3}\tilde{f}_0 + \frac{\tau}{6}\tilde{f}_1, v_j\right).$$

Introducing vectors c and b by means of the relations

$$\tfrac{1}{2}(V_0 + V_1) = \sum_{i=1}^{N} c_i v_i = \sum_{i=1}^{N} \tfrac{1}{2}(a_{0i} + a_{1i})\, v_i\,,$$

$$V_1 - V_0 = \sum_{i=1}^{N} b_i v_i = \sum_{i=1}^{N} (a_{1i} - a_{0i})\, v_i\,,$$

the systems (1.35) and (1.33) may be written in the following matrix form

$$\left(G + \frac{\tau}{2}\mathscr{A}\right)\mathbf{c} - \frac{\tau}{12}\mathscr{A}\mathbf{b} = \omega_0 + \frac{\tau}{6}\mathbf{F}(0) + \frac{\tau}{3}\mathbf{F}\left(\frac{\tau}{2}\right), \tag{1.36}$$

$$\tau\mathscr{A}\mathbf{c} + G\mathbf{b} = \frac{\tau}{6}\left[\mathbf{F}(0) + 4\,\mathbf{F}\left(\frac{\tau}{2}\right) + \mathbf{F}(\tau)\right],$$

where G, $\mathscr{A}$, $\mathbf{F}(t)$ and ω_0 were defined in (1.12).

Next let us replace $u^{(1)}(t)$ by the quadratic approximation

$$u^{(2)} = U_0 + N_{1/2}^{(2)}(t)\,\bar{U}_{1/2} + N_1^{(2)}(t)\,\bar{U}_1\,, \tag{1.37}$$

where U_0 and $\bar{U}_1$ were defined in (1.15), (1.16), $N_{1/2}^{(2)}(t)$ and $N_1^{(2)}(t)$ in (1.27),

$$\bar{U}_{1/2} = \sum_{i=1}^{N} \bar{u}_{1/2\,i} v_i\,.$$

The coefficients $\bar{u}_{1/2i}$ and $\bar{u}_{1i}$ will be determined by means of the projection condition

$$\int_0^\tau (u^{(2)} - u^{(1)}, \ \delta u^{(1)}) \, dt = 0 . \tag{1.38}$$

Inserting (1.37), (1.23) and (1.29) into (1.38), we obtain the following conditions ($k = 0, 1$)

$$\int_0^\tau (U_0 + N_{1/2}^{(2)}(t) \, \bar{U}_{1/2} + N_1^{(2)}(t) \, \bar{U}_1, N_k^{(1)}(t) \, v_j) \, dt =$$
$$= \int_0^\tau (N_0^{(1)}(t) \, V_0 + N_1^{(1)}(t) \, V_1 , \ N_k^{(1)}(t) \, v_j) \, dt ,$$

which yield two systems of equations

$$\left(\frac{\tau}{2} u_0 + \frac{\tau}{3} \bar{U}_{1/2}, v_j\right) = \left(\frac{\tau}{3} V_0 + \frac{\tau}{6} V_1, v_j\right), \tag{1.39}$$
$$\left(\frac{\tau}{2} u_0 + \frac{\tau}{3} \bar{U}_{1/2} + \frac{\tau}{6} \bar{U}_1, v_j\right) = \left(\frac{\tau}{6} V_0 + \frac{\tau}{3} V_1, v_j\right).$$

From (1.39) we obtain by subtraction

$$(\bar{U}_1, v_j) = (V_1 - V_0, v_j) ,$$

consequently

$$\bar{\boldsymbol{u}}_1 = \boldsymbol{b} . \tag{1.40}$$

Using the identity (1.34), the first equation of (1.39) may be rewritten in the form

$$(\tfrac{1}{2} u_0 + \tfrac{1}{3} \bar{U}_{1/2}, v_j) = (\tfrac{1}{4}(V_1 + V_0) - \tfrac{1}{12}(V_1 - V_0), v_j) ,$$

which is equivalent to

$$G \, \bar{\boldsymbol{u}}_{1/2} = \tfrac{3}{2} G \boldsymbol{c} - \tfrac{1}{4} G \boldsymbol{b} - \tfrac{3}{2} \omega_0 ,$$

consequently

$$\bar{\boldsymbol{u}}_{1/2} = \tfrac{3}{2} \boldsymbol{c} - \tfrac{1}{4} \boldsymbol{b} - \tfrac{3}{2} \boldsymbol{w}_0 .$$

The same procedure may be repeated in the following intervals

$$\langle \tau, 2\tau \rangle, \langle 2\tau, 3\tau \rangle, \ldots$$

If the coefficients $\boldsymbol{w}_m$ of the expansion

$$u^{(2)}(m\tau) = \sum_{i=1}^{N} w_{mi} v_i , \quad m = 0, 1, \ldots$$

are known, the next step is to solve the system

$$(1.41)\qquad G\mathbf{c}_m - \left(\frac{\tau}{12}\mathscr{A} + \frac{1}{2}G\right)\mathbf{b}_m = G\mathbf{w}_m + \frac{\tau}{12}\left[\mathbf{F}(m\tau) - \mathbf{F}(m\tau + \tau)\right],$$

$$\frac{1}{\tau}G\mathbf{b}_m + \mathscr{A}\mathbf{c}_m = \frac{1}{6}\left[\mathbf{F}(m\tau) + 4\,\mathbf{F}\left(m\tau + \frac{\tau}{2}\right) + \mathbf{F}(m\tau + \tau)\right],$$

which follows from an analogy of (1.36) by elimination of $\mathscr{A}\mathbf{c}$ in the first equation.

Then the coefficients of $u^{(2)}(m\tau + \tau)$ are given by the formula

$$(1.42)\qquad \mathbf{w}_{m+1} = \mathbf{w}_m + \mathbf{b}_m$$

and

$$(1.43)\qquad u^{(2)}(t) = \sum_{i=1}^{N}\left[w_{mi} + N_{1/2}^{(2)}(t)\left(\tfrac{3}{2}c_{mi} - \tfrac{1}{4}b_{mi} - \tfrac{3}{2}w_{mi}\right) + N_1^{(2)}(t)\,b_{mi}\right]v_i$$

holds in the interval $m\tau \leqq t \leqq m\tau + \tau$.

Remark 1.2. If G and $\mathscr{A}$ are band matrices, with the band width s, the system (1.41) can be rearranged easily so that the resulting matrix will also turn out to be a band matrix and its band width equals $2s + 1$.

Third approximation

Let us consider the approximate solution of the form

$$(1.44)\qquad u^{(2)} = N_0^{(2)}(t)\,V_0 + N_{1/2}^{(2)}(t)\,V_1 + N_1^{(2)}(t)\,V_2\,,$$

where

$$V_2 = \sum_{i=1}^{N} a_{2i}v_i$$

and V_0, V_1 were given in (1.24), $N_0^{(2)}, N_{1/2}^{(2)}, N_1^{(2)}$ in (1.27).

The function $f(t)$ will be approximated by a quadratic function

$$(1.45)\qquad f^{(2)}(t) = N_0^{(2)}(t)\tilde{f}_0 + N_{1/2}^{(2)}(t)\tilde{f}_1 + N_1^{(2)}(t)\tilde{f}_2\,,$$

where

$$(1.46)\qquad \tilde{f}_0 = \frac{1}{40}\left[31\,f(0) + 27\,f\left(\frac{\tau}{3}\right) - 27\,f(\tfrac{2}{3}\tau) + 9\,f(\tau)\right],$$

$$\tilde{f}_1 = \frac{1}{16}\left[-f(0) + 9\,f\left(\frac{\tau}{3}\right) + 9\,f(\tfrac{2}{3}\tau) - f(\tau)\right],$$

$$\tilde{f}_2 = \frac{1}{40}\,9\left[f(0) - 27\,f\left(\frac{\tau}{3}\right) + 27\,f(\tfrac{2}{3}\tau) + 31\,f(\tau)\right].$$

The formulas (1.46) may be obtained by setting

$$f^{(3)} = N_0^{(3)}(t)\,f(0) + N_{1/3}^{(3)}(t)\,f\left(\frac{\tau}{3}\right) + N_{2/3}^{(3)}(t)\,f(\tfrac{2}{3}\tau) + N_1^{(3)}(t)\,f(\tau)$$

and

$$\int_0^\tau (f^{(3)} - f^{(2)})\,N_k^{(2)}(t)\,\mathrm{d}t = 0\,; \quad k = 0, \tfrac{1}{2}, 1\,,$$

where

(1.47)
$$N_0^{(3)}(t) = 1 - \frac{11}{2}\frac{t}{\tau} + 9\left(\frac{t}{\tau}\right)^2 - \frac{9}{2}\left(\frac{t}{\tau}\right)^3,$$

$$N_{1/3}^{(3)}(t) = \frac{9}{2}\frac{t}{\tau}\left(2 - 5\frac{t}{\tau} + 3\left(\frac{t}{\tau}\right)^2\right),$$

$$N_{2/3}^{(3)}(t) = -\frac{9}{2}\frac{t}{\tau}\left(1 - 4\frac{t}{\tau} + 3\left(\frac{t}{\tau}\right)^2\right),$$

$$N_1^{(3)}(t) = \frac{9}{2}\frac{t}{\tau}\left(\frac{2}{9} - \frac{t}{\tau} + \left(\frac{t}{\tau}\right)^2\right).$$

Thus $f^{(2)}(t)$ is the $L_2(0, \tau)$ projection of the Lagrangian third degree interpolate of $f(t)$ into the subspace of quadratic functions.

In order to determine the coefficients a_{0i}, a_{1i} and a_{2i}, we employ the variational condition

(1.48)
$$\int_0^\tau \left(\left(u^{(2)}(t) + \int_0^t A\,u^{(2)}(z)\,\mathrm{d}z - \varphi_0 - \int_0^t f^{(2)}(z)\,\mathrm{d}z\,,\quad \delta u^{(2)}(\tau - t)\right)\mathrm{d}t = 0\,,$$

where the term with the operator A will be replaced, as previously, by the corresponding bilinear form.

We have

(1.49)
$$\delta u^{(2)}(t) = N_k^{(2)}(t)\,v_j\,; \quad k = 0, \tfrac{1}{2}, 1\,; \quad j = 1, 2, \ldots, N\,,$$

$$N_0^{(2)}(\tau - t) = N_1^{(2)}(t)\,, \quad N_1^{(2)}(\tau - t) = N_0^{(2)}(t)\,,$$

$$N_{1/2}^{(2)}(\tau - t) = N_{1/2}^{(2)}(t)\,.$$

Inserting (1.44), (1.45), (1.47) and (1.49) into (1.48) and integrating we obtain the following system of equations

(1.50)
$$(48G + 5\mathscr{A}\tau)\,\boldsymbol{a}_0 + 4(6G - \mathscr{A}\tau)\,\boldsymbol{a}_1 - (12G + \mathscr{A}\tau)\,\boldsymbol{a}_2 =$$

$$= 60\omega_0 + \frac{3}{10}\tau\left[13\,\mathbf{F}(0) + 6\,\mathbf{F}\left(\frac{\tau}{3}\right) - 21\,\mathbf{F}\left(\frac{2}{3}\tau\right) + 2\,\mathbf{F}(\tau)\right],$$

$$(6G + 11\mathscr{A}\tau)\,\boldsymbol{a}_0 + (48G + 20\mathscr{A}\tau)\,\boldsymbol{a}_1 + (6G - \mathscr{A}\tau)\,\boldsymbol{a}_2 =$$

$$= 60\omega_0 + \frac{3}{20}\tau\left[47\,\boldsymbol{F}(0) + 129\,\boldsymbol{F}\left(\frac{\tau}{3}\right) + 21\,\boldsymbol{F}\left(\frac{2}{3}\tau\right) + 3\,\boldsymbol{F}(\tau)\right],$$

$$(-12G + 11\mathscr{A}\tau)\,\boldsymbol{a}_0 + 4(6G + 11\mathscr{A}\tau)\,\boldsymbol{a}_1 + (48G + 5\mathscr{A}\tau)\,\boldsymbol{a}_2 =$$

$$= 60\omega_0 + \frac{3}{10}\tau\left[23\,\boldsymbol{F}(0) + 96\,\boldsymbol{F}\left(\frac{\tau}{3}\right) + 69\,\boldsymbol{F}\left(\frac{2}{3}\tau\right) + 12\,\boldsymbol{F}(\tau)\right],$$

where G, $\mathscr{A}$, ω_0 and $\boldsymbol{F}(t)$ were introduced in (1.12).

Let us replace $u^{(2)}(t)$ by a cubic approximation

$$u^{(3)}(t) = U_0 + N_{1/3}^{(3)}(t)\,\bar{U}_{1/3} + N_{2/3}^{(3)}(t)\,\bar{U}_{2/3} + N_1^{(3)}(t)\,\bar{U}_1\,,$$

where U_0 and $\bar{U}_1$ were defined by (1.15), (1.16), $N_{1/3}^{(3)}(t)$, $N_{2/3}^{(3)}(t)$, $N_1^{(3)}(t)$ by (1.47) and

$$\bar{U}_{1/3} = \sum_{i=1}^{N} \bar{u}_{1/3i} v_i\,, \quad \bar{U}_{2/3} = \sum_{i=1}^{N} \bar{u}_{2/3i} v_i\,.$$

The coefficients $\bar{u}_{1/3i}$, $\bar{u}_{2/3i}$ and $\bar{u}_{1i}$ will be determined from the projection condition

$$\int_0^\tau (u^{(3)} - u^{(2)}, \delta u^{(2)})\,\mathrm{d}t = 0\,,$$

which yields the formulas

(1.51)
$$\begin{aligned}
\bar{\boldsymbol{u}}_1 &= \boldsymbol{a}_0 + \boldsymbol{a}_2 - 2\boldsymbol{w}_0\,,\\
\bar{\boldsymbol{u}}_{2/3} &= \tfrac{2}{27}(-7\boldsymbol{a}_0 + 12\boldsymbol{a}_1 + 3\boldsymbol{a}_2 - 8\boldsymbol{w}_0)\,,\\
\bar{\boldsymbol{u}}_{1/3} &= \tfrac{1}{27}(17\boldsymbol{a}_0 + 24\boldsymbol{a}_1 - 3\boldsymbol{a}_2 - 38\boldsymbol{w}_0)\,.
\end{aligned}$$

If the coefficients $\boldsymbol{w}_m$ of the expansion for $u^{(3)}(m\tau)$ are known, the next step is to solve the system of the type (1.50) for $\boldsymbol{a}_0^{(m)}$, $\boldsymbol{a}_1^{(m)}$, $\boldsymbol{a}_2^{(m)}$, where $G\boldsymbol{w}_m$ is substituted for ω_0 and to all arguments of $\boldsymbol{F}\,m\tau$ is added. Then the vectors of coefficients in the expansion

$$u^{(3)}(t) = \sum_{i=1}^{N} [w_{mi} + N_{1/3}^{(3)}(t)\,\bar{u}_{1/3i} + N_{2/3}^{(3)}(t)\,\bar{u}_{2/3i} + N_1^{(3)}(t)\,\bar{u}_{1i}]\,v_i\,,$$

$$m\tau \leqq t \leqq m\tau + \tau$$

are given by the formulas (1.51), where $\boldsymbol{a}_k$ $(k = 0, 1, 2)$ and $\boldsymbol{w}_0$ are replaced by $\boldsymbol{a}_k^{(m)}$ and $\boldsymbol{w}_m$, respectively.

2. CONVERGENCE OF THE SECOND APPROXIMATION

The first approximation has been studied by J. Douglas, Jr. and T. Dupont [1] in detail even for non-linear parabolic equations. These authors proved several a priori estimates not only for the Crank-Nicholson-Galerkin approximation but also for its linearization by means of the predictor-corrector procedure or by the extrapolation.

In the present section we shall derive an error bound for the second approximation. Then the error bound will be used with approximation theory for Hermite interpolation in two variables to give rates of convergence. The fundamental line of thought is similar to that of Douglas and Dupont [1].

From (1.42) and (1.43) we conclude that

$$\mathbf{w}_{m+1/2} - \mathbf{w}_m = \tfrac{3}{2}\mathbf{c}_m - \tfrac{1}{4}(\mathbf{w}_{m+1} - \mathbf{w}_m) - \tfrac{3}{2}\mathbf{w}_m\,,$$

where

$$u^{(2)}\left(m\tau + \frac{\tau}{2}\right) = \sum_{i=1}^{N} w_{m+1/2,i}v_i\,.$$

Consequently, we have

$$c_m = \tfrac{1}{6}(\mathbf{w}_m + 4\mathbf{w}_{m+1/2} + \mathbf{w}_{m+1})\,. \tag{2.1}$$

Inserting (2.1) and (1.42) into (1.41) and returning to the scalar products, we derive that (1.41) is equivalent to the following system of equations

$$\frac{4}{\tau}(U_m - 2U_{m+1/2} + U_{m+1}, V) + [U_{m+1} - U_m, V]_A = (f_{m+1} - f_m, V)\,, \tag{2.2}$$

$$\left(\frac{U_{m+1} - U_m}{\tau}, V\right) + \tfrac{1}{6}[U_m + 4U_{m+1/2} + U_{m+1}, V]_A = \tag{2.3}$$

$$= \tfrac{1}{6}(f_m + 4f_{m+1/2} + f_{m+1}, V)\,,$$

where

$$U_m = u^{(2)}(m\tau)\,, \quad U_{m+1/2} = u^{(2)}\left(m\tau + \frac{\tau}{2}\right),$$

$$f_m = f(m\tau)\,, \quad f_{m+1/2} = f\left(m\tau + \frac{\tau}{2}\right), \quad m = 0, 1, \ldots$$

and V is any element of $\mathscr{M}$.

Let $L_2(I, \mathscr{V})$ denote the space of functions $u(t)$, mapping the interval $I = \langle 0, T\rangle$ into $\mathscr{V}$ and such that

$$\int_0^T \|u(t)\|^2 \,\mathrm{d}t < \infty\,.$$

$\mathscr{C}(I, H)$ denotes the space of continuous mappings of I into H.

Let $f \in \mathscr{C}(I, H)$. Assume that the solution u of the problem (1.1), (1.2) is such that $u \in L_2(I, \mathscr{V})$, $\mathrm{d}u/\mathrm{d}t \in \mathscr{C}(I, H)$ and

$$(2.4) \qquad \left(\frac{\mathrm{d}u}{\mathrm{d}t}, v\right) + [u, v]_A = (f, v), \quad 0 < t \leqq T, \quad v \in \mathscr{V},$$
$$(u(0), v) = (\varphi_0, v), \quad v \in \mathscr{V}$$

(for the concepts of weak solutions, see e.g. [5], chpt. IV.) Moreover, suppose that

$$\lim_{t \to 0+} u(t) = u(0) \quad \text{in} \quad \mathscr{V}$$

exists, consequently

$$(2.5) \qquad \left(\frac{\mathrm{d}u}{\mathrm{d}t}(0+), v\right) + [u(0), v]_A = (f(0), v), \quad v \in \mathscr{V}.$$

Remark I.2.1. We can prove easily, that the system (2.2), (2.3) possesses a unique solution at each time step. In fact, (2.2), (2.3) is equivalent to (1.41). Note that $\mathscr{A}$ and G are positive definite symmetric matrices, therefore $\mathscr{A}^{-1}$ exists and is positive definite as well. From the second eq. (1.41) we obtain

$$\mathbf{c}_m = -\frac{1}{\tau}\mathscr{A}^{-1}G\mathbf{b}_m + \frac{1}{6}\mathscr{A}^{-1}\left[\mathbf{F}(m\tau) + 4\,\mathbf{F}\left(m\tau + \frac{\tau}{2}\right) + \mathbf{F}(m\tau + \tau)\right]$$

and substituting this expression into the first equation, we are led to the equation

$$\left(\frac{1}{\tau}G\mathscr{A}^{-1}G + \frac{1}{2}G + \frac{\tau}{12}\mathscr{A}\right)\mathbf{b}_m = -G\mathbf{w}_m + \frac{\tau}{12}\left[\mathbf{F}(m\tau + \tau) - \mathbf{F}(m\tau)\right].$$

The matrix in brackets on the left-hand side is positive definite and the system possesses a unique solution $\mathbf{b}_m$. Then the uniqueness of $\mathbf{c}_m$ is evident.

Henceforth we shall use the following notation

$$(2.6) \qquad \delta u_m = u_{m+1} - u_m, \quad \delta u_{m+1/2} = u_{m+3/2} - u_{m+1/2}$$
$$\Delta^2 u_m = u_m - 2u_{m+1/2} + u_{m+1}$$
$$\hat{u}_m = \tfrac{1}{6}(u_m + 4u_{m+1/2} + u_{m+1}),$$
$$u'_m = \frac{\mathrm{d}u}{\mathrm{d}t}(m\tau), \quad u'_0 = \frac{\mathrm{d}u}{\mathrm{d}t}(0+), \quad u'_M = \frac{\mathrm{d}u}{\mathrm{d}t}(T-), \quad \tau = \frac{T}{M}, \quad M \text{ integer},$$

$$(2.7) \qquad \tilde{u}(t) = \sum_{i=1}^{N} \alpha_i(t)\, v_i,$$

$\alpha_i(t)$ real functions, C a generic constant, which is not necessarily the same at each occurrence.

We are going to deduce an a priori estimate, which implies that the second approximation is fourth order correct in time.

Theorem I.2.1. *Suppose that the solution $u(t)$ of (2.4) possesses continuous derivatives in H up to the fourth order on $\langle 0, T\rangle$ and the norms $|\mathrm{d}^5u/\mathrm{d}t^5|$ are bounded uniformly for $0 < t < T$. Denote $z_m = u_m - U_m$, where U_m is the solution of (2.2), (2.3) with the initial condition (1.16).*

Then there exist positive constants γ, C and τ_0, independent of τ, such that for $\tau \leqq \tau_0$ and for any function $\tilde{u}$ of the form (2.7) the following inequality holds

$$(2.8)\qquad |z_M|^2 + \gamma \sum_{m=0}^{M-1} \tau(\|\delta z_m\|^2 + \|\hat{z}_m\|^2) \leqq$$

$$\leqq C\left\{\sum_{m=0}^{M-1} \tau\left(\|(u-\tilde{u})_m^\wedge\|^2 + \left\|\frac{1}{\tau}\,\delta(u-\tilde{u})_m\right\|^2\right) + \sum_{m=0}^{M-2} \tau \left\|\frac{1}{\tau}\,\delta(u-\tilde{u})_{m+1/2}\right\|^2 + \right.$$

$$\left. + |(u-\tilde{u})_0|^2 + |(u-\tilde{u}_0^\wedge)|^2 + |(u-\tilde{u})_{M-1}^\wedge|^2 + \tau^8\right\}.$$

Proof. Making use of (2.4), (2.5) and (2.6), we obtain

$$(2.9)\qquad [u_k, v]_A = (f_k - u_k', v)\,, \quad k = 0, \tfrac{1}{2}, 1, \ldots, M\,; \quad v \in \mathscr{V}\,,$$

consequently

$$\left(\frac{1}{\tau}\,\delta u_m, v\right) + [\hat{u}_m, v]_A - (\hat{f}_m, v) = \left(\frac{1}{\tau}\,\delta u_m - u_m'^\wedge, v\right) = (\varrho_m, v)\,, \quad v \in \mathscr{V}\,,$$

$$m = 0, 1, 2, \ldots, M-1\,,$$

where

$$(2.10)\qquad |\varrho_m| < C\tau^4$$

(with C independent of m) can be deduced whenever $|\mathrm{d}^5u/\mathrm{d}t^5| < C$ holds for $0 < t < T$. Theferore we have

$$(2.11)\qquad \left(\frac{1}{\tau}\,\delta u_m, v\right) + [\hat{u}_m, v]_A = (\hat{f}_m + \varrho_m, v)\,, \quad v \in \mathscr{V}\,,$$

$$0 \leqq m \leqq M-1\,.$$

Using (2.9), we can see that

$$[\delta u_m, v]_A = (\delta f_m - \delta u_m', v), \quad v \in \mathscr{V}\,, \quad m = 0, 1, 2, \ldots, M-1\,,$$

consequently

$$(2.12)\qquad \frac{4}{\tau}(\Delta^2 u_m, v) + [\delta u_m, v]_A - (\delta f_m, v) = \left(\frac{4}{\tau}\,\Delta^2 u_m - \delta u_m', v\right) = (\zeta_m, v)\,,$$

where

$$|\zeta_m| < C\tau^3 \tag{2.13}$$

(with C independent of m). Therefore we may write

$$\left(\frac{4}{\tau}\Delta^2 u_m, v\right) + [\delta u_m, v]_A = (\delta f_m + \zeta_m, v), \quad v \in \mathscr{V}, \quad 0 \leqq m \leqq M-1. \tag{2.14}$$

If we subtract (2.3) from (2.11), with $v = V = (\tilde{u} - U)^{\wedge}_m$, we obtain

$$\left(\frac{1}{\tau}\delta z_m, (\tilde{u} - U)^{\wedge}_m\right) + [\hat{z}_m, (\tilde{u} - U)^{\wedge}_m]_A = (\varrho_m, (\tilde{u} - U)^{\wedge}_m). \tag{2.15}$$

Subtracting eq. (2.2) from (2.14) with $v = V = \delta(\tilde{u} - U)_m$, we obtain

$$\left(\frac{4}{\tau}\Delta^2 z_m, \delta(\tilde{u} - U)_m\right) + [\delta z_m, \delta(\tilde{u} - U)_m]_A = (\zeta_m, \delta(\tilde{u} - U)_m). \tag{2.16}$$

Let us consider the following identity, where

$$z_m = u_m - \tilde{u}_m + (\tilde{u}_m - U_m)$$

is used several times:

$$\begin{aligned} &\frac{4}{\tau}(\Delta^2 z_m, \delta z_m) + [\delta z_m, \delta z_m]_A + \frac{12}{\tau}(\delta z_m, \hat{z}_m) + 12[\hat{z}_m, \hat{z}_m]_A = \\ &= \frac{4}{\tau}(\Delta^2 z_m, \delta(u - \tilde{u})_m) + [\delta z_m, \delta(u - \tilde{u})_m]_A + \frac{12}{\tau}(\delta z_m, (u - \tilde{u})^{\wedge}_m) + \\ &+ 12[\hat{z}_m, (u - \tilde{u})^{\wedge}_m]_A + \left\{\frac{4}{\tau}(\Delta^2 z_m, \delta(\tilde{u} - U)_m) + [\delta z_m, \delta(\tilde{u} - U)_m]_A + \right. \\ &\left. + \frac{12}{\tau}(\delta z_m, (\tilde{u} - U)^{\wedge}_m) + 12[\hat{z}_m, (\tilde{u} - U)^{\wedge}_m]_A - 12(\varrho_m, (\tilde{u} - U)^{\wedge}_m) - (\zeta_m, \delta(\tilde{u} - U)_m)\right\} + \\ &+ 12(\varrho_m, (\tilde{u} - U)^{\wedge}_m) + (\zeta_m, \delta(\tilde{u} - U)_m). \end{aligned} \tag{2.17}$$

The expression in brackets in the right-hand side vanishes because of (2.15) and (2.16). The right-hand side of (2.17) is bounded by

$$\begin{aligned} &\left(4\Delta^2 z_m, \frac{1}{\tau}\delta(u - \tilde{u})_m\right) + \frac{12}{\tau}(\delta z_m, (u - \tilde{u})^{\wedge}_m) + \left(\tau\zeta_m, \frac{1}{\tau}\delta(\tilde{u} - U)_m\right) + \\ &+ C_0\|\delta z_m\| \, \|\delta(u - \tilde{u})_m\| + 12C_0\|\hat{z}_m\| \, \|(u - \tilde{u})^{\wedge}_m\| + 12|\varrho_m| \, |(\tilde{u} - U)^{\wedge}_m|. \end{aligned} \tag{2.18}$$

Using the relations

$$\begin{aligned} (\tilde{u} - U)_m &= (\tilde{u} - u)_m + z_m, \\ \Delta^2 z_m &= -3\hat{z}_m + \tfrac{3}{2}(z_m + z_{m+1}), \end{aligned} \tag{2.19}$$

the scalar products in (2.18) may be bounded above by

$$(2.20)\qquad 12|\hat{z}_m| \left|\frac{1}{\tau}\,\delta(u - \hat{u})_m\right| + 6|z_m + z_{m+1}| \left|\frac{1}{\tau}\,\delta(u - \tilde{u})_m\right| +$$
$$+ \left(\frac{1}{\tau}\,\delta z_m,\ 12(u - \tilde{u})_m^{\wedge} + \tau\zeta_m\right) + |\tau\zeta_m| \left|\frac{1}{\tau}\,\delta(\tilde{u} - u)_m\right| \leqq$$
$$\leqq C_1\,\varepsilon\|\hat{z}_m\|^2 + C\left[|z_m|^2 + |z_{m+1}|^2 + \left\|\frac{1}{\tau}\,\delta(u - \tilde{u})_m\right\|^2 + |\tau\zeta_m|^2\right] +$$
$$+ \left(\frac{1}{\tau}\,\delta z_m,\ 12(u - \tilde{u})_m + \tau\zeta_m\right).$$

Here we have used (1.3) and the well-known inequality

$$|ab| \leqq \varepsilon a^2 + b^2/4\varepsilon\,,$$

so that $\varepsilon > 0$ is arbitrary, C_1 does not depend on ε, $C = C(\varepsilon)$. Proceeding similarly with the other terms in (2.18), we obtain the bound for the whole right-hand side in the form

$$(2.21)\qquad C_0\varepsilon\|\delta z_m\|^2 + C_2\varepsilon\|\hat{z}_m\|^2 + \psi\,,$$

where

$$\psi = C\left[\left\|\frac{1}{\tau}\,\delta(u - \tilde{u})_m\right\|^2 + \|(u - \tilde{u})_m^{\wedge}\|^2 + |\varrho_m|^2 + |\tau\zeta_m|^2 + |z_m|^2 + |z_{m+1}|^2\right] +$$
$$+ \left(\frac{1}{\tau}\,\delta z_m,\ 12(u - \tilde{u})_m^{\wedge} + \tau\zeta_m\right)$$

and the constants C_0, C_2 do not depend on ε, τ, m.

The left-hand side of (2.17) can be bounded below, using (2.19) and (1.4), by

$$\frac{6}{\tau}\left(|z_{m+1}|^2 - |z_m|^2\right) + \alpha\|\delta z_m\|^2 + 12\alpha\|\hat{z}_m\|^2\,.$$

Altogether we have the inequality

$$(2.22)\qquad \frac{6}{\tau}\left(|z_{m+1}|^2 - |z_m|^2\right) + (\alpha - C_0\varepsilon)\,\|\delta z_m\|^2 + (12\alpha - C_2\varepsilon)\,\|\hat{z}_m\|^2 \leqq \psi\,.$$

Multiplying (2.22) by $\tau/6$ and increasing the constant C, we conclude that

$$(2.23)\qquad (1 - C\tau)\,|z_{m+1}|^2 - (1 + C\tau)\,|z_m|^2 + \gamma\tau\left(\|\delta z_m\|^2 + \|\hat{z}_m\|^2\right) \leqq$$
$$\leqq C\tau\left[\left\|\frac{1}{\tau}\,\delta(u - \tilde{u})_m\right\|^2 + \|(u - \tilde{u})_m^{\wedge}\|^2 + |\varrho_m|^2 + |\tau\zeta_m|^2\right] - \tau|z_m|^2 +$$
$$+ \tau\left(\frac{1}{\tau}\,\delta z_m,\ 2(u - \tilde{u})_m^{\wedge} + \frac{\tau}{6}\,\zeta_m\right)$$

holds for sufficiently small ε and with positive constants γ and C, independent of τ, m.

Let us introduce a function

$$k(\tau) = (1 - C\tau)/(1 + C\tau)$$

and multiply (2.23) by $k(\tau)^m(1 + C\tau)^{-1}$. One can derive easily, that such constants k_0, k_1 and τ_0 exist for which

$$0 < k_0 \leqq k(\tau)^m \leqq k_1 \tag{2.24}$$

holds for all $1 \leqq m \leqq M = T/\tau$ and $\tau \leqq \tau_0$. Using (2.24) we can see that

$$\begin{aligned} (2.25) \qquad & k(\tau)^{m+1}|z_{m+1}|^2 - k(\tau)^m |z_m|^2 + \gamma_1\tau(\|\delta z_m\|^2 + \|\hat{z}_m\|^2) \leqq \\ & \leqq C\tau\left[\left\|\frac{1}{\tau}\,\delta(u - \tilde{u})_m\right\|^2 + \|(u - \tilde{u})_m^\wedge\|^2 + |\varrho_m|^2 + |\tau\zeta_m|^2\right] - \gamma_2\tau|z_m|^2 + \\ & + \tau\left(\frac{1}{\tau}\,\delta z_m,\ k(\tau)^m\, W_m/(1 + C\tau)\right), \\ & W_m = 2(u - \tilde{u})_m^\wedge + \tfrac{1}{6}\tau\zeta_m \end{aligned}$$

holds for $\tau \leqq \tau_0$, $0 \leqq m \leqq M - 1$.

Let us sum (2.25) on $m = 0, 1, \ldots, M - 1$, making use of the following estimates

$$\begin{aligned} & \sum_{m=0}^{M-1} \tau\left(\frac{1}{\tau}\,\delta z_m,\ k(\tau)^m\, W_m/(1 + C\tau)\right) = \\ & = \frac{1}{1 + C\tau}\left\{-\,(z_0, W_0) + (z_M, k(\tau)^{M-1}\, W_{M-1}) + \right. \\ & \left. + \sum_{m=1}^{M-1} \tau\left(z_m,\ \frac{1}{\tau}\,(k(\tau)^{m-1}\, W_{m-1} - k(\tau)^m\, W_m)\right)\right\} \leqq \\ & \leqq |z_0|\,|W_0| + k_1|z_M|\,|W_{M-1}| + \sum_{m=1}^{M-1} \tau k_1 |z_M| \left|\frac{1}{\tau}\,(W_{m-1} - k(\tau)\, W_m)\right|, \end{aligned}$$

$$\begin{aligned} & \left|\frac{1}{\tau}\,(W_{m-1} - k(\tau)\, W_m)\right| = \left|\frac{1}{\tau}\,(W_{m-1} - W_m + (1 - k(\tau))\, W_m)\right| \leqq \\ & \leqq \left|\frac{1}{\tau}\,(W_m - W_{m-1})\right| + 2C|W_m| \leqq C\left[\left\|\frac{1}{\tau}\,\delta(u - \tilde{u})_m\right\| + \left\|\delta(u - \tilde{u})_{m-1}\,\frac{1}{\tau}\right\| + \right. \\ & \left. + \left\|\frac{1}{\tau}\,\delta(u - \tilde{u})_{m-1/2}\right\| + \|(u - \tilde{u})_m^\wedge\| + |\tau\zeta_m| + |\delta\zeta_{m-1}|\right]. \end{aligned}$$

Thus we may write

$$\sum_{m=0}^{M-1} \tau\left(\frac{1}{\tau}\,\delta z_m,\ k(\tau)^m\, W_m/(1 + C\tau)\right) \leqq$$

$$\leqq \tfrac{1}{2}|z_0|^2 + k_1\varepsilon|z_M|^2 + C[|(u - \tilde{u})_0^\wedge|^2 + |\tau\zeta_0|^2 + |(u - \tilde{u})_{M-1}^\wedge|^2 + |\tau\zeta_{M-1}|^2] +$$

$$+\sum_{m=1}^{M-1}\tau\left[k_1\varepsilon|z_m|^2 + C\left(\left\|\frac{1}{\tau}\,\delta(u-\tilde{u})_{m-1}\right\|^2 + \left\|\frac{1}{\tau}\,\delta(u-\tilde{u})_m\right\|^2 + \right.\right.$$

$$\left.\left. + \left\|\frac{1}{\tau}\,\delta(u-\tilde{u})_{m-1/2}\right\|^2 + \|(u-\tilde{u})^{\wedge}_m\|^2 + |\tau\zeta_m|^2 + |\delta\zeta_{m-1}|^2\right)\right].$$

The sum of (2.25) yields the following inequality

$$(2.26)\qquad (k(\tau)^M - k_1\varepsilon)\,|z_M|^2 - |z_0|^2 + \sum_{m=0}^{M-1}\gamma_1\tau(\|\delta z_m\|^2 + \|\hat{z}_m\|^2) \leqq$$

$$\leqq C\tau\sum_{m=0}^{M-1}\left[\left\|\frac{1}{\tau}\,\delta(u-\tilde{u})_m\right\|^2 + \|(u-\tilde{u})^{\wedge}_m\|^2 + |\varrho_m|^2 + |\tau\zeta_m|^2\right] +$$

$$+ \tau\sum_{m=0}^{M-1}(k_1\varepsilon - \gamma_2)\,|z_m|^2 + \tfrac{1}{2}|z_0|^2 + C\left[|(u-\tilde{u})^{\wedge}_0|^2 + |(u-\tilde{u})^{\wedge}_{M-1}|^2 + \right.$$

$$\left. + |\tau\zeta_0|^2 + |\tau\zeta_{M-1}|^2 + \tau\sum_{m=0}^{M-2}\left(|\delta\zeta_m|^2 + \left\|\frac{1}{\tau}\,\delta(u-\tilde{u})_{m+1/2}\right\|^2\right)\right].$$

From (2.4) and (1.16) we conclude that

$$V \in \mathcal{M} \Rightarrow ((u-U)_0, V) = 0\,,$$

consequently

$$|z_0|^2 = (z_0, (u-\tilde{u})_0 + (\tilde{u}-U)_0) = (z_0, (u-\tilde{u})_0) \leqq |z_0|\,|(u-\tilde{u})_0|\,,$$

hence

$$(2.27)\qquad |z_0| \leqq |(u-\tilde{u})_0|\,.$$

If ε is sufficiently small, we deduce from (2.26), (2.24) and (2.27) that

$$|z_M|^2 + \sum_{m=0}^{M-1}\gamma\tau(\|\delta z_m\|^2 + \|\hat{z}_m\|^2) \leqq C\left\{|(u-\tilde{u})_0|^2 + |(u-\tilde{u})^{\wedge}_0|^2 + |(u-\tilde{u})^{\wedge}_{M-1}|^2 + \right.$$

$$+ |\tau\zeta_0|^2 + |\tau\zeta_{M-1}|^2 + \sum_{m=0}^{M-1}\tau\left(\left\|\frac{1}{\tau}\,\delta(u-\tilde{u})_m\right\|^2 + \|(u-\tilde{u})^{\wedge}_m\|^2 + |\varrho_m|^2 + |\tau\zeta_m|^2\right) +$$

$$\left. + \sum_{m=0}^{M-2}\tau\left(\left\|\frac{1}{\tau}\,\delta(u-\tilde{u})_{m+1/2}\right\|^2 + |\delta\zeta_m|^2\right)\right\}.$$

The terms $|\varrho_m|^2$ and $|\tau\zeta_m|^2$, $m = 0, 1, \ldots, M-1$ are $0(\tau^8)$. By virtue of the boundedness of $|d^5u/dt^5|$, we can prove easily, that also

$$|\delta\zeta_m|^2 \leqq C\tau^8\,,$$

where C does not depend on m, τ. Hence (2.8) follows and the proof is complete.

We shall demonstrate how the estimate (2.8) can be used to get rates of convergence, on the same example as in [1]. Let us consider the parabolic equation on $\Omega = (0,1) \times (0,1)$ of the following form

$$\frac{\partial u}{\partial t} - \sum_{i,j=1}^{2} \frac{\partial}{\partial x_i}\left[a_{ij}(x)\frac{\partial u}{\partial x_j}\right] = f(x,t),$$

the initial condition

$$u(\cdot,0) = \varphi_0 \in \mathring{W}_2^{(1)}(\Omega)$$

and the boundary condition

$$u(x,t) = 0, \quad x \in \partial\Omega.$$

We shall employ the Hermite interpolation theory in the (x_1, x_2)-plane [6]. We have $H = L_2(\Omega)$, $V = \mathring{W}_2^{(1)}(\Omega)$ (the well-known Sobolev space of functions vanishing on the boundary). Let us denote

$$h = J^{-1}, \quad J \text{ a positive integer,}$$

$$D^\alpha f(x) = \frac{\partial^{|\alpha|} f(x)}{\partial x_1^{\alpha_1}\,\partial x_2^{\alpha_2}}, \quad |\alpha| = \alpha_1 + \alpha_2\,; \quad \alpha_1, \alpha_2 \text{ non-negative integers.}$$

Let $H_h^{(n)}$ denote the set of real-valued functions g such that, for all $0 \leqq \alpha_1, \alpha_2 \leqq n - 1$, $D^\alpha g$ is continuous on $\bar{\Omega}$ and such that on each square $(kh, kh + h) \times (lh, lh + h)$, where k and l are integers satisfying $0 \leqq k, l \leqq J - 1$, g is a polynomial of the form

$$\sum_{i,j=0}^{2n-1} c_{ij} x_1^i x_2^j.$$

Suppose that f is a function on $\bar{\Omega}$ such that $D^\alpha f$ is continuous on $\bar{\Omega}$ for $0 \leqq \alpha_1, \alpha_2 \leqq n - 1$. We say that "*$f_{n,h}$ is the $H_h^{(n)}$-interpolate of f*" if $f_{n,h} \in H_h^{(n)}$ and

$$D^\alpha(f - f_{n,h})(kh, lh) = 0$$

for all k, l, α_1 and α_2 integers such that $0 \leqq k, l \leqq J$ and $0 \leqq \alpha_1, \alpha_2 \leqq n - 1$.

In the following we shall need a special case of Theorem 5 of [6], namely

Lemma I.2.1. *Let $D^\alpha f$ be continuous on $\bar{\Omega}$ for $|\alpha| < 2n$ and $D^\alpha f \in L_2(\Omega)$ for $|\alpha| = 2n$. Let $f_{n,h}$ be the $H_h^{(n)}$-interpolate of f. Then there exists a constant Q, which is independent of h and such that*

$$\|D^\alpha(f - f_{n,h})\|_{L_2} \leqq Q h^{2n-|\alpha|},$$

where $|\alpha| \leqq 2n - 1$, $0 \leqq \alpha_1, \alpha_2 \leqq n$. Further,

$$Q = Q' \sum_{|\alpha|=2n} \|D^\alpha f\|_{L_2}$$

where Q' does not depend on f and h.

Theorem I.2.2. *Let* $\mathscr{M} = H_h^{(n)} \cap \mathring{W}_2^{(1)}(\Omega)$. *Let* u, U *and* z *be as in Theorem I.2.1. Suppose that for every* $t \in \langle 0, T \rangle$ u *and* $\partial u/\partial t$ *satisfy the hypotheses of Lemma I.2.1 and that*

$$\sum_{\alpha=2n} \|D^\alpha u(\cdot, t)\|_{L_2} \leqq C' , \tag{2.28}$$

$$\sum_{\alpha=2n} \left\| D^\alpha \frac{\partial u}{\partial t}(\cdot, t) \right\|_{L_2} \leqq \chi(t) ,$$

where C' *is independent of* t, $\chi \in L_2(0, T)$ *and* D^α *denotes spatial derivatives only. Then there exist constants* C, τ_0 *independent of* h, τ *such that* for $\tau \leqq \tau_0$

$$\|z_M\|_{L_2}^2 + \sum_{k=0}^{M-1} \tau\left(\|z_{k+1} - z_k\|_{\mathring{W}_2(1)}^2 + \|\tfrac{1}{6}(z_k + 4z_{k+1/2} + z_{k+1})\|_{\mathring{W}_2(1)}^2\right) \leqq \tag{2.29}$$
$$\leqq C(h^{2(2n-1)} + \tau^8) .$$

Proof. Let $\tilde{u}(x, t)$ be the $H_h^{(n)}$-interpolate of $u(x, t)$, $t \in \langle 0, T \rangle$. Using Lemma I.2.1 we see that for $k = 0, \frac{1}{2}, 1, \ldots, M$

$$\|(u - \tilde{u})_k\|_{L_2} \leqq Q'C'h^{2n} , \quad \|(u - \tilde{u})_k^\wedge\|_{L_2} \leqq Q'C'h^{2n} , \tag{2.30}$$
$$\|(u - \tilde{u})_k^\wedge\|_{\mathring{W}_2(1)} \leqq Ch^{2n-1} .$$

Note that $\partial \tilde{u}/\partial t$ is just the $H_h^{(n)}$-interpolate of $\partial u/\partial t$, because of the coincidence of $D^\alpha u$ and $D^\alpha \tilde{u}$ along the straight lines (kh, lh, t) for $0 \leqq \alpha_1, \alpha_2 \leqq n - 1$. We may therefore write for $0 \leqq k \leqq M - 1$, using also (2.28), the following estimate

$$\left\|\frac{1}{\tau}\left[(u - \tilde{u})_{k+1} - (u - \tilde{u})_k\right]\right\|_{\mathring{W}_2(1)}^2 \leqq \frac{1}{\tau^2}\left\|\int_{k\tau}^{(k+1)\tau} \frac{\partial(u - \tilde{u})}{\partial t}\, \mathrm{d}t\right\|_{\mathring{W}_2(1)}^2 \leqq \tag{2.31}$$
$$\leqq \frac{1}{\tau^2}\left(\int_{k\tau}^{(k+1)\tau} \left\|\frac{\partial(u - \tilde{u})}{\partial t}\right\|_{\mathring{W}_2(1)} \mathrm{d}t\right)^2 \leqq \frac{1}{\tau^2}\left(\int_{k\tau}^{(k+1)\tau} \sqrt{(2)}\, h^{2n-1}\, Q'\chi(t)\, \mathrm{d}t\right)^2 \leqq$$
$$\leqq 2(Q')^2 \frac{1}{\tau} h^{2(2n-1)} \int_{k\tau}^{(k+1)\tau} \chi^2(t)\, \mathrm{d}t .$$

Inserting (2.30) and (2.31) into (2.8) we obtain the estimate (2.29).

Finally, we shall briefly discuss the stability of the process (2.2), (2.3), (1.16) with respect to the initial condition.

Theorem I.2.3. *Let* $f = 0$ *in* (2.2), (2.3). *Then*

$$|U_{m+1}| \leqq |U_m| \leqq |\varphi_0| , \tag{2.32}$$

$$|\hat{U}_m| \leqq |U_m| , \quad |U_{m+1/2}| \leqq 2|U_m| \tag{2.33}$$

hold for every $m = 0, 1, 2, \ldots, M - 1$.

Proof. If we insert $V = \partial U_m$ into (2.2) and $V = \hat{U}_m$ into (2.3), we can see that

$$\frac{4}{\tau}(|U_{m+1}|^2 - |U_m|^2) - \frac{8}{\tau}(U_{m+1/2}, \delta U_m) + [\delta U_m, \delta U_m]_A = 0\,,$$

$$\frac{1}{6\tau}(|U_{m+1}|^2 - |U_m|^2) + \frac{2}{3\tau}(\delta U_m, U_{m+1/2}) + [\hat{U}_m, \hat{U}_m]_A \quad = 0\,.$$

Consequently

$$\frac{6}{\tau}(|U_{m+1}|^2 - |U_m|^2) + [\delta U_m, \delta U_m]_A + 12[\hat{U}_m, \hat{U}_m]_A = 0 \tag{2.34}$$

holds. Further, we have

$$|U_0|^2 = (\varphi_0, U) \leqq |\varphi_0| \cdot |U_0|\,. \tag{2.35}$$

Then (2.32) results from (2.34) and (2.35).

In order to derive (2.33), let us insert $V = \hat{U}_m$ into (2.2) and $V = \delta U_m$ in (2.3). Using also (2.19), we derive by subtraction that

$$(\delta U_m, \delta U_m) = 4(-3\hat{U}_m) + \tfrac{3}{2}(U_m + U_{m+1}), \hat{U}_m)\,,$$

consequently

$$|U_{m+1} - U_m)|^2 + 12|\hat{U}_m|^2 = 6(U_{m+1} + U_m, \hat{U}_m) \leqq \tag{2.36}$$
$$\leqq 6|U_m + U_{m+1}|\,|\hat{U}_m|\,.$$

Then (2.33) follows from (2.36), (2.32) and (2.6).

Remark I.2.2. If the base-functions $v_i(x)$ correspond with Lagrange interpolation polynomials, then the coefficients $(w_m + \bar{u}_{1/2})_i$ and $(w_m + \bar{u}_1)_i = (w_m + b_m)_i$ coincide with the nodal values of the second approximation. Then the inequality (2.32) means that the corresponding two-steps difference scheme is unconditionally stable with respect to the initial condition. The same result can be obtained on the base of certains theorem of A. A. Samarskij [7], transforming (1.41) to the canonical two-steps form. See also Remark II. 1.2, which yields somewhat stronger relation.

3. THE COMPARISON OF THE THREE APPROXIMATIONS FOR ORDINARY DIFFERENTIAL EQUATIONS

The case of one ordinary differential equation

$$\frac{dy}{dt} + Ay = f,\ y(0) = y_0\,, \quad t \in \langle 0, \tau \rangle\,, \quad A > 0 \tag{3.1}$$

may serve as an interesting example to compare the efficiency of the three semi-variational approximations, introduced in Section 1.

Assuming that $f(t)$ is sufficiently regular, the solution of (3.1) at $t = \tau$

$$y(\tau) = y_0 e^{-A\tau} + \int_0^\tau f(z)\, e^{-A(\tau - z)}\, dz$$

may be written in terms of power series as follows

$$(3.2) \qquad y(\tau) = y_0 \left(1 - A\tau + \frac{1}{2} A^2\tau^2 - \frac{1}{6} A^3\tau^3 + \frac{1}{24} A^4\tau^4 - \frac{1}{120} A^5\tau^5 + \right.$$

$$\left. + \frac{1}{720} A^6\tau^6 - \frac{1}{5040} A^7\tau^7 + \ldots \right) +$$

$$+ \left[\tau f - \frac{1}{2} A\tau^2 f + \frac{1}{24} \tau^3(4A^2 f + 2Af' + f'') - \frac{1}{48} \tau^4(2A^3 f + 2A^2 f' + Af'') + \right.$$

$$\left. + \frac{\tau^5}{1920} (16A^4 f + 24A^3 f' + 16A^2 f'' + 4Af''' + f^{\mathrm{IV}}) - \ldots \right]\Bigg|_{t=\tau/2} .$$

The first approximation at $t = \tau$ is

$$(3.3) \qquad y^{(1)}(\tau) = y_0 \frac{1 - \frac{1}{2}A\tau}{1 + \frac{1}{2}A\tau} + \frac{\tau}{2} \frac{f_0 + f_1}{1 + \frac{1}{2}A\tau} =$$

$$= y_0(1 - A\tau + \tfrac{1}{2}A^2\tau^2 - \tfrac{1}{4}A^3\tau^3 + \ldots) + [\tau f - \tfrac{1}{2}A\tau^2 f + \tau^3(\tfrac{1}{4}A^2 f + \tfrac{1}{8}f'') - \ldots]\big|_{t=\tau/2} .$$

Hence the error $y^{(1)}(\tau) - y(\tau) = 0(\tau^3)$.

The second approximation is equal to

$$(3.4) \quad y^{(2)}(\tau) = y_0 \frac{1 - \frac{1}{2}A\tau + \frac{1}{12}A^2\tau^2}{1 + \frac{1}{2}A\tau + \frac{1}{12}A^2\tau^2} + \frac{\frac{1}{6}\tau(f_0 + 4f_{1/2} + f_1) + \frac{1}{12}A\tau^2(f_1 - f_0)}{1 + \frac{1}{2}A\tau + \frac{1}{12}A^2\tau^2} =$$

$$= y_0(1 - A\tau + \tfrac{1}{2}A^2\tau^2 - \tfrac{1}{6}A^3\tau^3 + \tfrac{1}{24}A^4\tau^4 - \tfrac{1}{144}A^5\tau^5 + \ldots) +$$

$$+ [\tau f - \tfrac{1}{2}A\tau^2 f + \tfrac{1}{24}\tau^3(4A^2 f + 2Af' + f'') - \tfrac{1}{48}\tau^4(2A^3 f + 2A^2 f' + Af'') + \ldots]\big|_{t=\tau/2} ,$$

consequently $y^{(2)}(\tau) - y(\tau) = 0(\tau^5)$.

The third approximation, applied to homogeneous equation (3.1) only, is equal to

$$(3.5) \qquad y^{(3)}(\tau) = y_0 \frac{1 - \frac{1}{2}A\tau + \frac{1}{10}A^2\tau^2 - \frac{1}{120}A^3\tau^3}{1 + \frac{1}{2}A\tau + \frac{1}{10}A^2\tau^2 + \frac{1}{120}A^3\tau^3} =$$

$$= y_0[1 - A\tau + A^2 \tfrac{1}{2}\tau^2 - \tfrac{1}{6}A^3\tau^3 + \tfrac{1}{24}A^4\tau^4 - \tfrac{1}{120}A^5\tau^5 + \tfrac{1}{720}A^6\tau^6 - \tfrac{1}{4800}A^7\tau^7 + \ldots] ,$$

consequently $y^{(3)}(\tau) - y(\tau) = 0(\tau^7)$.

The rational functions standing by y_0 in (3.3), (3.4), (3.5) agree with Padé approximations of $e^{-A\tau}$ (see e.g. [8]). In Part II, we shall prove the latter coincidence for the general n-th approximation.

4. NUMERICAL EXAMPLES — COMPARISON OF THE TWO APPROXIMATIONS FOR A PARABOLIC EQUATION

Obviously, the mixed problem for parabolic equations represents the most important application of the method of semi-variational approximations. We consider the following problem in $\Omega \times \langle 0, \infty)$, $\Omega = (0, 1)$:

$$\frac{\partial u}{\partial t} - 6\frac{\partial^2 u}{\partial x^2} = e^{-t}(1 - t)\, x^2(1 - x)^2 - 12te^{-t}(1 - 6x + 6x^2)\,,$$

$$u = 0 \quad \text{for} \quad x = 0\,, \quad x = 1\,,$$

$$u = 0 \quad \text{for} \quad t = 0\,,$$

the solution of which is

$$u = te^{-t}\, x^2(1 - x)^2\,.$$

According to the formulas (1.21) and (1.41), the first (Crank-Nicholson-Galerkin) and the second approximation, respectively, were calculated with piecewise cubic base functions $v_i(x)$ (Hermite interpolation polynomials for $n = 2$). The relative errors

$$e^{(j)}(t) = |u^{(j)}(x, t) - u(x, t)|/u(x, t)\,, \quad (j = 1, 2)$$

for $x = 1/2$ and $0 < t \leqq 1.6$ are presented in the following

Table 1

t	$h = 1/12$		$h = 1/8$	
	$\tau = 0{\cdot}1$	$\tau = 0{\cdot}2$	$\tau = 0{\cdot}2$	$\tau = 0{\cdot}4$
	$10^4\, e^{(1)}(t)$	$10^4\, e^{(2)}(t)$	$10^4\, e^{(1)}(t)$	$10^4\, e^{(2)}(t)$
0·1	6·188	0·497		
0·2	1·472	0·089	15·319	2·302
0·2	1·657	0·401		
0·4	1·000	0·014	1·702	0·561
0·5	0·957	0·358		
0·6	0·749	0·078	5·102	2·172
0·7	0·701	0·332		
0·8	0·602	0·121	1·582	0·378
0·9	0·570	0·315		
1·0	0·509	0·149	3·143	2·078
1·1	0·488	0·303		
1·2	0·446	0·170	1·480	0·209
1·3	0·433	0·295		
1·4	0·400	0·185	2·350	1·997
1·5	0·392	0·288		
1·6	0·365	0·196	1·393	0·053

Note that even the four-times greater time increment $\tau = 0.4$ gives approximately twice better results in the semi-variational method than the increment $\tau = 0.1$ in Crank-Nicholson-Galerkin method, when the errors at the basic time instants $t = 0.4; 0.8; 1.2; 1.6$ are compared.

References

[1] *J. Douglas, Jr., T. Dupont:* Galerkin methods for parabolic equations. SIAM J. Numer. Anal. 7 (1970), 4, 575—626.

[2] *I. Hlaváček:* Variational formulation of the Cauchy problem for equations with operator coefficients. Aplikace matematiky 16 (1971), 1, 46—63.

[3] *I. Hlaváček:* Variational principles for parabolic equations. Aplikace matematiky 14 (1969), 4, 278—297.

[4] *E. L. Wilson, R. E. Nickell:* Application of the finite element method to heat conduction analysis. Nucl. Eng. and Design 4 (1966), 276—286.

[5] *J. L. Lions:* Equations differentielles operationelles et problèmes aux limites. Grundlehren Math. Wiss. Bd. 111, Springer 1961.

[6] *G. Birkhoff, M. H. Schultz, R. S. Varga:* Piecewise Hermite interpolation in one and two variables with application to partial differential equations. Numerische Math. 11 (1968), 232—256.

[7] *А. А. Самарский:* Некоторые вопросы общей теории разностных схем. Сб. „Дифференциалные уравления с частными производными.“ Издат. Наука, Москва 1970.

[8] *A. Ralston:* A first course in numerical analysis. Mc Graw-Hill, 1965.

Souhrn

O JEDNÉ POLOVARIAČNÍ METODĚ PRO PARABOLICKÉ ROVNICE

IVAN HLAVÁČEK

Cílem této práce je další rozvoj metody konečných prvků, aplikované na smíšené úlohy pro parabolickou rovnici. Hodně již byla prostudována metoda Galerkinova typu řádu τ^2, (kde τ je časový krok), obdobná známé Crank-Nicholsonově metodě sítí [1], [4]. Zde se navrhuje posloupnost aproximací s rostoucí přesností vzhledem k času. Prvá aproximace se ztotožňuje s uvedenou metodou Crank-Nicholson-Galerkinovou. Pro druhou aproximaci se dokazuje rychlost konvergence řádu τ^4 a stabilita vůči počáteční podmínce. Účinnost metody je doložena numerickými příklady.

Author's address: Ing. *Ivan Hlaváček*, CSc., Matematický ústav ČSAV v Praze, Žitná 25, Praha 1.