Aplikace matematiky

Ivan Hlaváček

On a semi-variational method for parabolic equations. II

Aplikace matematiky, Vol. 18 (1973), No. 1, 43–64

Persistent URL: http://dml.cz/dmlcz/103447

SVAZEK 18 (1973) APLIKACE MATEMATIKY ČÍSLO 1

ON A SEMI-VARIATIONAL METHOD FOR PARABOLIC EQUATIONS II

IVAN HLAVÁČEK

(Received December 23, 1971)

INTRODUCTION

In Part I of this paper we presented a numerical procedure for approximate solution of abstract parabolic equations. It is the aim of the Part II to give further information on the proposed method and to show how to apply it to some problems, which are more general than those of Part I.

In Section 1 we prove the invariance of the n-th semi-variational approximation with respect to the polynomial bases and its coincidence with the Padé approximations in some sense. In Section 2 a parabolic equation with inhomogeneous mixed boundary conditions is solved by means of the semi-variational method. In Section 3 the exextension of the method to an abstract equation with two positive definite operators is discussed and the convergence and stability of the first and second approximations proved.

1. SOME PROPERTIES OF THE SEMI-VARIATIONAL APPROXIMATIONS

The semi-variational approximations have been constructed by means of the Lagrangian interpolation polynomials (cf. Section I.1). A question arises, whether this is the only possible way of derivation. We are going to show that, in case of homogeneous abstract equation (I.1.1), any polynomial bases lead to the same n-th approximation as a result. This assertion is formulated in the following Theorems II.1.1 and II.1.2 exactly. Let $\mathscr{P}_m$, $m = 0, 1, 2, \ldots$ denote the subspace in $L_2(0, \tau)$ of polynomials of degree m.

Theorem II.1.1. *Let* $\{S_k(t)_{k=0}^{k=n-1}$ *form a polynomial basis in* $\mathscr{P}_{n-1}$, $n \geqq 1$, *and let*

$$u^{(n-1)}(t) = \sum_{k=0}^{n-1} \sum_{i=1}^{N} S_k(t)\, a_i^{(k)} v_i \tag{1.1}$$

represent an approximation of a solution to the problem

$$\frac{\mathrm{d}u}{\mathrm{d}t} + Au = 0\,, \quad u(0) = \varphi_0\,, \quad 0 \leqq t \leqq \tau\,,$$

(*i.e., the Cauchy problem of Section* I.1 *with* $f = 0$ *and* $T = \tau$).

Then the approximation $u^{(n-1)}(t)$ *is determined uniquely by the variational conditions*

$$\int_0^\tau \Big(u^{(n-1)}(t) + \int_0^t A\, u^{(n-1)}(z)\,\mathrm{d}z - \varphi_0\,, S_j(t)\, v_m\Big)\,\mathrm{d}t = 0\,, \tag{1.2}$$

$$0 \leqq j \leqq n-1\,, \quad m = 1, 2, \ldots, N\,,$$

being independent of the choice of the polynomial basis.

Proof. Consider the Legendre polynomials $P_k(x)$, $k = 0, 1, 2, \ldots$, $-1 \leqq x \leqq 1$, and transform the interval $\langle -1, 1\rangle$ onto $\langle 0, \tau\rangle$, setting $x = 2t/\tau - 1$. Thus we obtain polynomials $P_k(2t/\tau - 1) = \bar{P}_k(t)$, which form an orthogonal sequence in $L_2(0, \tau)$. Substituting $S_k(t) = \bar{P}_k(t)$ into (1.1) and (1.2), we obtain

$$u^{(n-1)}(t) = \sum_{k=0}^{n-1} \sum_{i=1}^{N} \bar{P}_k(t)\, a_i^{(k)} v_i\,,$$

$$\sum_{k=0}^{n-1} \sum_{i=1}^{N} \int_0^\tau \Big\{ a_i^{(k)}\, \bar{P}_k(t)\, \bar{P}_j(t)\, (v_\iota, v_m) + \Big[\int_0^t a_i^{(k)}\, \bar{P}_k(z)\,\mathrm{d}z\, v_i, \bar{P}_j(t)\, v_m\Big]_A \Big\}\,\mathrm{d}t = \tag{1.3}$$

$$= \int_0^\tau (\varphi_0, v_m)\, \bar{P}_j(t)\,\mathrm{d}t\,,$$

$$0 \leqq j \leqq n-1\,, \quad 1 \leqq m \leqq N\,.$$

If we introduce matrices p and q with the terms

$$p_{jk} = \int_0^\tau \bar{P}_k(t)\, \bar{P}_j(t)\,\mathrm{d}t\,, \quad q_{jk} = \int_0^\tau \bar{P}_j(t) \int_0^t \bar{P}_k(z)\,\mathrm{d}z\,\mathrm{d}t\,,$$

the system (1.3) can be rewritten as follows

$$\mathscr{L}a = p_0 \omega_0\,, \tag{1.4}$$

where

$$a^T = \big((\boldsymbol{a}^{(0)})^T, (\boldsymbol{a}^{(1)})^T, \ldots, (\boldsymbol{a}^{(n-1)})^T\big)\,, \quad (\boldsymbol{a}^{(k)})^T = \big(a_1^{(k)}, a_2^{(k)}, \ldots, a_N^{(k)}\big),$$ [1]

$$\mathscr{L}_{jk} = p_{jk} G_+ q_{jk} \mathscr{A}\,, \quad p_0^T = (p_{00}, p_{10}, \ldots, p_{n-1,0})\,.$$

[1] M^T denotes the transpose of the matrix M.

Lemma II.1.1. *The system* (1.4) *possesses a unique solution for every* $n \geqq 1$, $N \geqq 1$ *and* $\tau > 0$.

Proof. We shall proceed by induction with respect to n. Let us derive eq. (1.4) explicitly. Using the orthogonality of Legendre polynomials and the formulas

$$\int_{-1}^{1} P_j^2(x)\,\mathrm{d}x = 2/(2j+1)\,, \quad P_j(-1) = (-1)^j\,, \quad P_j(1) = 1\,,$$

$$(2j+3)\int_{-1}^{x} P_{j+1}(\xi)\,\mathrm{d}\xi = P_{j+2}(x) - P_j(x)\,, \quad 0 \leqq j\,,$$

$$\int_{-1}^{x} P_0(\xi)\,\mathrm{d}\xi = P_0(x) + P_1(x)\,,$$

we obtain the system (1.4), in the following form

$$(1.5)_0 \qquad \tau\left(G + \tfrac{1}{2}\tau\mathscr{A}\right)\boldsymbol{a}^{(0)} - \tfrac{1}{6}\tau^2\mathscr{A}\boldsymbol{a}^{(1)} = \tau\omega_0\,,$$

$$(1.5)_1 \qquad \tfrac{1}{6}\tau^2\mathscr{A}\boldsymbol{a}^{(0)} + \tfrac{1}{3}\tau G\boldsymbol{a}^{(1)} - \frac{\tau^2}{30}\mathscr{A}\boldsymbol{a}^{(2)} = 0\,,$$

$$\vdots$$

$$(1.5)_j \qquad \frac{\tau^2}{2(2j+1)}\mathscr{A}\left(\frac{\boldsymbol{a}^{(j-1)}}{2j-1} - \frac{\boldsymbol{a}^{(j+1)}}{2j+3}\right) + \frac{\tau}{2j+1}G\boldsymbol{a}^{(j)} = 0\,,$$

$$j \geqq 2\,, \quad \left(a^{(n)} = 0\right).$$

If we multiply the equation $(1.5)_0$ by $2/\tau$, $(1.5)_1$ by $6/\tau$, $(1.5)_j$ by $2(2j+1)/\tau$, we are led to the equivalent system

$$(1.6) \qquad \mathscr{D}^{(n)}a = y\,,$$

where

$$(1.7) \qquad \mathscr{D}^{(n)} = \begin{bmatrix} 2G + \tau\mathscr{A}\,, & -\frac{\tau}{3}\mathscr{A}\,, & & & & \\ \tau\mathscr{A}\,, & 2G\,, & -\frac{\tau}{5}\mathscr{A}\,, & & & \\ & \frac{\tau}{3}\mathscr{A}\,, & 2G\,, & -\frac{\tau}{7}\mathscr{A}\,, & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & \frac{\tau}{2n-5}\mathscr{A}\,, & 2G\,, & -\frac{\tau}{2n-1}\mathscr{A} \\ & & & & \frac{\tau}{2n-3}\mathscr{A}\,, & 2G \end{bmatrix}, \quad y = \begin{bmatrix} 2\omega_0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Denote the matrix $(nN \times nN)$ of (1.7) by $D^{(n)}$. Suppose that (i) the inverse $(D^{(n)})^{-1}$ exists. Divide both $D^{(n)}$ and $(D^{(n)})^{-1}$ into block matrices $\mathscr{D}_{jk}^{(n)}$ and $(\mathscr{D}^{(n)})_{jk}^{-1}$ of the type $(N \times N)$, respectively, as has just been done in (1.7) for $D^{(n)}$. Moreover, suppose that (ii) the matrix $(\mathscr{D}^{(n)})_{n-1,n-1}^{-1}$ is positive definite. Then $(D^{(n+1)})^{-1}$ exists and the corresponding enty $(\mathscr{D}^{(n+1)})_{n,n}^{-1}$ is positive definite, as well. In fact, the method of inversion by partitioning yields

$$D^{(n+1)} = \begin{bmatrix} D^{(n)}, & U_n \\ V_n, & 2G \end{bmatrix},$$

$$V_n = \Bigl(\underbrace{0, 0, \ldots,}_{(n-1)\times} \frac{\tau}{2n-1}\mathscr{A}\Bigr), \quad U_n^T = \Bigl(\underbrace{0, 0, \ldots,}_{(n-1)\times} \frac{-\tau}{2n+1}\mathscr{A}\Bigr)$$

and the existence of $(D^{(n+1)})^{-1}$, if $(D^{(n)})^{-1}$ and ϑ^{-1} exist, where

$$\vartheta = 2G - V_n(D^{(n)})^{-1} U_n = 2G + \frac{\tau^2}{(2n+1)(2n-1)} \mathscr{A}(\mathscr{D}^{(n)})_{n-1,n-1}^{-1} \mathscr{A}\,.$$

From the induction assumptions (i) and (ii) both these conditions follow, because ϑ is positive definite $(N \times N)$ matrix for any τ, n, N. Moreover, by virtue of the relation

$$(\mathscr{D}^{(n+1)})_{n,n}^{-1} = \vartheta^{-1}\,,$$

the latter matrix is positive definite.

The assumptions (i) and (ii) hold for $n = 1$, when

$$D^{(1)} = 2G + \tau\mathscr{A}\,, \quad (\mathscr{D}^{(1)})_{0,0}^{-1} = (2G + \tau\mathscr{A})^{-1}\,.$$

Consequently, (i) and (ii) hold for every $n \geqq 1$, $N \geqq 1$ and $\tau > 0$.

Now we can continue the proof of Theorem II.1.1. Let $\{S_k(t)\}_0^{n-1}$ be any polynomial basis in $\mathscr{P}_{n-1}$. Then it holds

$$S_k(t) = \sum_{l=0}^{n-1} c_{kl}\,\bar{P}_l(t) = [C\,\bar{P}(t)]_k\,, \tag{1.8}$$

where C is a regular matrix. Consider the function

$$v^{(n-1)}(t) = \sum_{k=0}^{n-1} \sum_{i=1}^{N} S_k(t)\, a_i'^{(k)} v_i\,. \tag{1.9}$$

Denoting

$$s_{j0}' = \int_0^\tau S_j(t)\,\mathrm{d}t\,, \quad s_{jk} = \int_0^\tau S_k(t)\,S_j(t)\,\mathrm{d}t\,, \quad r_{jk} = \int_0^\tau S_j(t) \int_0^t S_k(z)\,\mathrm{d}z\,\mathrm{d}t\,,$$

and making use of (1.8), we obtain

$$s_{jk} = \sum_{l,r=0}^{n-1} c_{kl}c_{jr}p_{lr}\,, \quad r_{jk} = \sum_{l,r=0}^{n-1} c_{jl}c_{kr}q_{lr}\,, \quad s_{j0}' = \sum_{l=0}^{n-1} c_{jl}p_{l0}\,. \tag{1.10}$$

From (1.2) it follows

$$(1.11)\qquad \sum_{k=0}^{n-1}(s_{jk}G + r_{jk}\mathscr{A})\,\mathbf{a}'^{(k)} = s'_{j0}\omega_0\,.$$

Inserting (1.10) into (1.11), we may write

$$\sum_{l,r=0}^{n-1} c_{kr}c_{jl}\sum_{k=0}^{n-1}(p_{lr}G + q_{lr}\mathscr{A})\,\mathbf{a}'^{(k)} = \sum_{l=0}^{n-1} c_{jl}p_{l0}\Gamma_0\,,$$

which may be written in the matrix form

$$(1.12)\qquad C\mathscr{L}C^{T}a' = Cp_0\omega_0\,.$$

Multiplying (1.12) by C^{-1} and using Lemma II.1.1, we can conclude that

$$(1.13)\qquad C^{T}a' = a\,,\quad \text{i.e.,}\quad \sum_{k=0}^{n-1} c_{kr}\mathbf{a}'^{(k)} = \mathbf{a}^{(r)}\,.$$

Inserting (1.8) and (1.13) into (1.9) we obtain

$$v^{(n-1)}(t) = \sum_{l,k=0}^{n-1}\sum_{i=1}^{N} c_{kl}\,\bar{P}_l(t)\,a_i'^{(k)}v_i = \sum_{i=1}^{N}\sum_{l=0}^{n-1}\bar{P}_l(t)\,a_i^{(l)}v_i = u^{(n-1)}(t)\,.\quad \text{Q.E.D.}$$

Theorem II.1.2 *Let $\{S_k^{(n-1)}(t)\}_0^{n-1}$ and $\{S_k^{(n)}(t)\}_0^{n}$ form polynomial bases in $\mathscr{P}_{n-1}$ and $\mathscr{P}_n$, respectively, and let*

$$u^{(n-1)}(t) = \sum_{k=0}^{n-1}\sum_{i=1}^{N}\bar{P}_k(t)\,a_i^{(k)}v_i$$

be given. Then the approximation

$$u^{(n)}(t) = \sum_{k=0}^{n}\sum_{i=1}^{N} S_k^{(n)}(t)\,b_i^{(k)}v_i$$

is determined uniquely by the initial condition

$$(1.14)\qquad (u^{(n)}(0), v_m) = (\varphi_0, v_m)\,,\quad m = 1, 2, \ldots, N$$

and the projection condition

$$(1.15)\qquad \int_0^{\tau}(u^{(n)}(t) - u^{(n-1)}(t)\,,\ S_j^{(n-1)}(t)\,v_m)\,\mathrm{d}t = 0\,,$$

$$j = 0, 1, \ldots, n-1\,,\quad m = 1, 2, \ldots, N\,,$$

being independent of the choice of the polynomial bases in $\mathscr{P}_{n-1}$ and $\mathscr{P}_n$.

Proof. First let us substitute for both bases the Legendre polynomials, i.e., $S_k^{(n-1)} = \bar{P}_k$, $S_k^{(n)} = \bar{P}_k$. Then (1.15) yields

$$\sum_{k=0}^{n}\sum_{i=1}^{N} b_i^{(k)}p_{jk}\,G_{mi} - \sum_{k=0}^{n-1}\sum_{i=1}^{N} a_i^{(k)}p_{jk}G_{mi} = 0\,,$$

$$m = 1, 2, \ldots, N\,,\quad j = 0, 1, \ldots, n-1\,.$$

Hence we conclude that

$$pbG = \bar{p}aG$$

where $\bar{p}$ denotes the diagonal $(n \times n)$ matrix (p_{ij}) and p denotes the $(n \times (n+1))$ matrix, which is formed by adding one zero column to $\bar{p}$. Consequently,

$$\mathbf{b}^{(k)} = \mathbf{a}^{(k)}\,, \quad 0 \leqq k \leqq n-1\,. \tag{1.16}$$

The initial condition (1.14) yields

$$G\big(\sum_{k=0}^{n} \bar{P}_k(0)\,\mathbf{b}^{(k)}\big) = \omega_0$$

and therefore

$$\mathbf{b}^{(n)} = (-1)^n \big[G^{-1}\omega_0 - \sum_{k=0}^{n-1}(-1)^k\,\mathbf{a}^{(k)}\big] \tag{1.17}$$

Next let $\{S_k^{(n-1)}(t)\}_0^{n-1}$ and $\{S_k^{(n)}(t)\}_0^n$ be arbitrary bases and

$$v^{(n)}(t) = \sum_{k=0}^{n}\sum_{i=1}^{N} S_k^{(n)}(t)\, b_i'^{(k)} v_i\,. \tag{1.18}$$

It holds

$$\begin{aligned} S_k^{(n)}(t) &= \sum_{r=0}^{n} h_{kr}\,\bar{P}_r(t)\,, \quad 0 \leqq k \leqq n\,, \\ S_j^{(n-1)}(t) &= \sum_{l=0}^{n-1} c_{jl}\,\bar{P}_l(t)\,, \quad 0 \leqq j \leqq n-1\,. \end{aligned} \tag{1.19}$$

The projection condition (1.15) results in

$$\sum_{k=0}^{n}\sum_{i=1}^{N} b_i'^{(k)} t_{jk} G_{mi} = \sum_{k=0}^{n-1}\sum_{i=1}^{N} a_i^{(k)} z_{jk} G_{mi}\,, \tag{1.20}$$

$$m = 1, \ldots, N\,, \quad j = 0, 1, \ldots, n-1\,,$$

where

$$t_{jk} = \int_0^\tau S_j^{(n-1)} S_k^{(n)}\,\mathrm{d}t\,, \quad z_{jk} = \int_0^\tau S_j^{(n-1)}\,\bar{P}_k\,\mathrm{d}t\,.$$

We may write (1.20) in the matrix form

$$tb'G = zaG\,. \tag{1.21}$$

Substituting (1.19) into (1.21), we obtain

$$t = Cph^T\,, \quad z = C\bar{p}\,,$$

$$Cph^Tb' = C\bar{p}a\,,$$

consequently, (cf. (1.16)),

$$(h^T b')^{(k)} = \boldsymbol{a}^{(k)} , \quad 0 \leqq k \leqq n - 1 . \tag{1.22}$$

The initial condition (1.14) yields, if we use also (1.19),

$$G(\sum_{k,r=0}^{n} h_{kr} \bar{P}_r(0) \, b'^{(k)}) = \omega_0 . \tag{1.23}$$

Hence we obtain from (1.22) and (1.23)

$$(h^T b')^{(n)} = (-1)^n \left[G^{-1} \omega_0 - \sum_{k=0}^{n-1} (-1)^k \, \boldsymbol{a}^{(k)} \right] . \tag{1.24}$$

By comparison of (1.22) and (1.24) with (1.16) and (1.17), respectively, we conclude that

$$h^T b' = b , \quad \text{i.e.,} \quad \sum_{k=0}^{n} h_{kr} \boldsymbol{b}'^{(k)} = \boldsymbol{b}^{(r)}$$

and by virtue of (1.19), we have

$$v^{(n)}(t) = \sum_k \sum_i S_k^{(n)} b_i'^{(k)} v_i = \sum_i \sum_{k,r} \bar{P}_r(t) \, h_{kr} \, b_i'^{(k)} v_i = \sum_i \sum_r \bar{P}_r(t) \, b_i^{(r)} v_i = u^{(n)}(t) . \quad \text{Q.E.D.}$$

Theorem II.1.3. *Let* $u^{(n)}(t)$, $n \geqq 1$, *be the* n-th *semi-variational approximation of the solution to the initial-value problem for the ordinary differential equation*

$$\frac{\mathrm{d}u}{\mathrm{d}t} + Au = 0 , \quad 0 < t \leqq \tau , \tag{1.25}$$

$$0 < A = \text{const.} , \quad u(0) = \varphi_0 .$$

Then

$$u^{(n)}(\tau) = \varphi_0 \, Q_n(-\alpha)/Q_n(\alpha) , \tag{1.26}$$

where

$$Q_n(\alpha) = \sum_{k=0}^{n} \frac{(2n - k)! \, n!}{(2n)! \, k! \, (n - k)!} \alpha^k , \quad \alpha = A\tau . \tag{1.27}$$

Remark II.1.1. The rational function $Q_n(-\alpha)/Q_n(\alpha)$ coincides with the well-known Padé approximation $R_{nn}(\alpha)$ of $\exp(-\alpha)$, consequently, the error of $u^{(n)}(\tau)$ is $0(\alpha^{2n+1})$ (cf. [2]).

Proof of Th.II.1.3. We can interpret eq. (1.25) as a particular case of the abstract eq. (I.1.1) if we set $H = R$ (the space of real numbers), $\mathscr{M} = \mathscr{V} = R, N = 1, v_1 = 1, G = 1, \mathscr{A} = A$. Then the Theorems II.1.1 and 1.2 imply that any polynomial basis can be employed instead of the Lagrangian interpolation polynomials. Let us choose the Legendre polynomials $\bar{P}_k(t)$. Then the system for $\mathbf{q}^{(k)}$ follows immediately from (1.7). Here we have

$$\mathscr{D}^{(n)} = \left[\begin{array}{cccccc} 2 + \alpha, & -\frac{\alpha}{3}, & & & & \\ \alpha, & 2, & -\frac{\alpha}{5}, & & & \\ & \frac{\alpha}{3}, & 2, & -\frac{\alpha}{7}, & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & \frac{\alpha}{2n-5}, & 2, & & -\frac{\alpha}{2n-1} \\ & & & \frac{\alpha}{2n-3}, & & 2 \end{array}\right] \tag{1.28}$$

and $\omega_0 = \varphi_0$. Denote

$$\det \mathscr{D}^{(n)} = D_n(\alpha)\,. \tag{1.29}$$

First we shall prove the following

Lemma II.1.2. *Let $D_n(\alpha)$ and $Q_n(\alpha)$ be defined by (1.29) and (1.27), respectively. Then*

$$D_n(\alpha) = 2^n\, Q_n(\alpha) \tag{1.30}$$

holds for every $n = 1, 2, \ldots$

Proof. Expanding D_n with respect to the last column, we obtain

$$D_n = 2D_{n-1} + \frac{\alpha}{2n-1} X_{n-1} \quad \text{for} \quad n \geqq 2\,, \tag{1.31}$$

where X_{n-1} is the corresponding subdeterminant. Consider D_{n+1}, multiply its $(n+1)$-th row by $\alpha/[2(2n+1)]$ and add the result to the n-th row. Then expanding the modified D_{n+1} with respect to the last column, we obtain

$$D_{n+1} = 2\left(A_n D_{n-1} + \frac{\alpha}{2n-1} X_{n-1}\right), \tag{1.32}$$

where

$$A_n = 2 + \frac{\alpha^2}{2(2n+1)(2n-1)}\,.$$

Elimination of X_{n-1} from (1.31) and (1.32) leads to the recurrence formula

$$(1.33) \qquad D_{n+1} = 2D_n + \frac{\alpha^2}{(2n+1)(2n-1)} D_{n-1}\,, \quad n \geqq 2\,,$$

which enables us to prove (1.30) by induction. It is easy to verify (1.30) for $n = 1, 2$, by direct calculation. Suppose that (1.30) holds for $n = 1, 2, \ldots, m$ and calculate D_{m+1}, making use of (1.33). Thus we have

$$D_{m+1} = 2^{m+1} \sum_{k=0}^{m} \frac{(2m-k)!\, m!}{(2m)!\, k!\, (m-k)!} \alpha^k +$$
$$+ \frac{\alpha^2}{(2m+1)(2m-1)} 2^{m-1} \sum_{k=0}^{m-1} \frac{(2m-2-k)!\,(m-1)!}{(2m-2)\, k!\, (m-1-k)!} \alpha^k\,.$$

The linear part of D_{m+1} is

$$2^{m+1}(1 + \tfrac{1}{2}\alpha)$$

and the coefficient by α^j, $2 \leqq j \leqq m+1$, may be shown equal to

$$2^{m+1} \frac{(2m+2-j)!\,(m+1)!}{(2m+2)!\, j!\, (m+1-j)!}\,.$$

Hence the formula (1.30) holds for D_{m+1}, consequently it holds for all $n \geqq 1$. Q.E.D.

Now recall the proof of Theorem II.1.2. Using (1.16) and (1.17) we derive

$$u^{(n)}(\tau) = \varphi_0 \frac{\bar{P}_n(\tau)}{\bar{P}_n(0)} + \sum_{k=0}^{n-1} a^{(k)} \left[\bar{P}_k(\tau) - \frac{\bar{P}_n(\tau)}{\bar{P}_n(0)} \bar{P}_k(0) \right].$$

From there we deduce

$$(1.34) \qquad \frac{u^{(n)}(\tau)}{\varphi_0} = 1 + \frac{2}{\varphi_0} \sum_{i=0}^{m-1} a^{(2i+1)} \quad \text{for} \quad n = 2m \quad \text{even}\,,$$

$$(1.35) \qquad \frac{u^{(n)}(\tau)}{\varphi_0} = -1 + \frac{2}{\varphi_0} \sum_{i=0}^{m} a^{(2i)} \quad \text{for} \quad n = 2m + 1 \text{ odd}\,.$$

Let $n = 2m$. Adding all even equations of (1.6) (i.e., those for $j = 1, 3, \ldots$) we obtain

$$\alpha a^{(0)} + 2 \sum_{i=0}^{m-1} a^{(2i+1)} = 0\,,$$

consequently, from (1.34) it follows that

$$(1.36) \qquad \frac{u^{(n)}(\tau)}{\varphi_0} = 1 - \frac{\alpha}{\varphi_0} a^{(0)}\,.$$

If $n = 2m + 1$, we add all odd equations of (1.6) (i.e., for $j = 0,2, \ldots$) to obtain

$$\alpha a^{(0)} + 2 \sum_{i=0}^{m} a^{(2i)} = 2\varphi_0 .$$

Then from (1.35) we come again to the formula (1.36).

It is easy to show, using (1.36), (1.28) and the Cramer's rule, that

$$\frac{u^{(n)}(\tau)}{\varphi_0} = \frac{E_n(\alpha)}{D_n(\alpha)} ,$$

where $E_n(\alpha)$ differs from $D_n(\alpha)$ only in the first entry $(E_n(\alpha))_{11}$, which is equal to $2 - \alpha$. Multiplying every even column and row of $E_n(\alpha)$ by (-1), we can immediately conclude that

$$E_n(\alpha) = D_n(-\alpha) .$$

Consequently, with the use of Lemma II.1.2, it holds

$$\frac{u^{(n)}(\tau)}{\varphi_0} = \frac{D_n(-\alpha)}{D_n(\alpha)} = \frac{Q_n(-\alpha)}{Q_n(\alpha)} , \quad \text{Q.E.D.}$$

Remark II. 1.2. For an abstract homogeneous equation (I.1.1) we can derive an analogous relation (cf. [6])

$$\boldsymbol{w}_1 = Q_n(-\alpha) \left[Q_n(\alpha)\right]^{-1} \boldsymbol{w}_0 = \left[Q_n(\alpha)\right]^{-1} Q_n(-\alpha) \boldsymbol{w}_0 ,$$

where $\boldsymbol{w}_1$ represents the vector of coefficients w_{1i}, $(i = 1, 2, \ldots, N)$ in the expansion

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

In order to prove this relation, we multiply every matrix equation (row) of (1.6) by G^{-1} and introduce a regular matrix

$$\alpha = \tau G^{-1} \mathscr{A} .$$

Thus we are led again to the matrix (1.28), where 2 is replaced by $2I_N$ (I_N being the unit matrix) in the diagonal entries.

The set of all polynomials $R(\alpha)$ with the matrix argument α generates a commutative linear algebra. Therefore the determinants with matrix entries can be defined precisely in the same way as the usual determinants. Using these generalized determinants, we can introduce (1.29) and prove Lemma II.1.2. Then (1.27) and (1.30) imply that $D_n(\alpha)$ is a regular matrix. We deduce again (cf. (1.36))

$$\boldsymbol{w}_1 = \boldsymbol{w}_0 - \alpha \boldsymbol{a}^{(0)}$$

and

$$\boldsymbol{a}^{(0)} = 2\left[D_n(\alpha)\right]^{-1} S_{11}^{(n)}(\alpha) \boldsymbol{w}_0 ,$$

where $S_{11}^{(n)}(\alpha)$ is the complement of $(2I_N + \alpha)$ in the determinant $D_n(\alpha)$. From there it follows that

$$\boldsymbol{w}_1 = [D_n(\alpha)]^{-1} D_n(-\alpha) \boldsymbol{w}_0 .$$

Finally, we use Lemma II.1.2 and the relation

$$[R(\alpha)]^{-1} S(\alpha) = S(\alpha) [R(\alpha)]^{-1}$$

holding for any pair of polynomials $R(\alpha)$, $S(\alpha)$ if $R(\alpha)$ is regular.

Remark II.1.3. Theorems II.1.1 and II.1.2 indicate that the orthogonal system of Legendre polynomials may be used to develop the semi-variational approximations even in the case of abstract parabolic homogeneous equation (I.1.1). The advantage of this particular version becomes evident for $n \geqq 3$, when the zero (matrix) entries appear in the systems for the unknowns $\boldsymbol{a}^{(k)}$ and the relative number of zeros increases with n (cf. (1.7)). The coefficients $\boldsymbol{a}^{(k)}$ can be calculated from (1.6) and then

$$u^{(n)}(t) = \sum_{i=1}^{N} [\sum_{k=0}^{n-1} a_i^{(k)} \bar{P}_k(t) + b_i^{(n)} \bar{P}_n(t)] v_i ,$$

where $b^{(n)}$ is given by (1.17).

2. PARABOLIC EQUATION WITH INHOMOGENEOUS BOUNDARY CONDITIONS

Let us consider the parabolic equation

$$\frac{\partial u}{\partial t} - \sum_{i,j=1}^{N} \frac{\partial}{\partial x_i} \left[a_{ij}(X) \frac{\partial u}{\partial x_j} \right] = f(X, t), \quad 0 < t \leqq T, \tag{2.1}$$

$$(x_1, \ldots, x_N) = X \in \Omega \subset E_N ,$$

with the initial condition

$$u(X, 0) = \varphi_0(X) \tag{2.2}$$

and the mixed boundary conditions of the following type

$$u = g \quad \text{on} \quad \Gamma_u \times (0, T\rangle , \tag{2.3}$$

$$\sum_{i,j=1}^{N} a_{ij}(X) \, \nu_i \frac{\partial u}{\partial x_j} = P(X, t) \quad \text{on} \quad \Gamma_h \times (0, T\rangle , \tag{2.4}$$

$$\alpha(X) \, u + \sum_{i,j=1}^{N} a_{ij}(X) \, \nu_i \frac{\partial u}{\partial x_j} = P(X, t) \quad \text{on} \quad \Gamma_v \times (0, T\rangle . \tag{2.5}$$

We assume that Ω is a Lipschitz region[1]), its boundary Γ is divided into four mutually disjoint parts

$$\Gamma = \Gamma_u \cup \Gamma_h \cup \Gamma_v \cup \Gamma_0 \,,$$

where $mes\ \Gamma_0 = 0$ and each of Γ_u, Γ_h, Γ_v is either open in Γ[2]) or empty.

Moreover, assume that $f(\cdot, t)$, $g(\cdot, t) \in W_2^{(1)}(\Omega)$ (the Sobolev space of square-integrable functions which possess the first derivatives in the generalized sense in $L_2(\Omega)$) and $P(\cdot, t) \in L_2(\Gamma_v \cup \Gamma_h)$ for each $t \in \langle 0, T \rangle$, $(\partial g/\partial t) \in \mathscr{C}(I, L_2(\Omega))$, i.e., continuous mapping of $I = \langle 0, T \rangle$ into $L_2(\Omega)$, $a_{ij}(X)$ are measurable functions of $X \in \bar{\Omega}$ such that the matrix $(a_{ij}(X))$ is symmetric and positive definite with its spectrum bounded above and below by positive numbers C_0 and η, respectively, which are independent of the argument X[3]). The function $\alpha(X)$ is measurable and it holds

$$0 < \alpha_0 \leqq \alpha(X) \leqq \alpha_1 \,, \quad X \in \Gamma_v \,, \tag{2.6}$$

ν_i are the components of the unit outward noımal to Γ and $\varphi_0 \in L_2(\Omega)$.

Assume that the solution of the problem (2.1)–(2.5) is such that

$$w = (u - g) \in L_2(I, \mathscr{V}) \,, \quad \frac{\partial u}{\partial t} \in \mathscr{C}(I, L_2(\Omega)) \,, \tag{2.7}$$

$$\left(\frac{\partial u}{\partial t}, v\right) + [u, v]_A = (f, v) + (P, v)_\Gamma \,, \quad 0 < t \leqq T, \quad v \in \mathscr{V} \,, \tag{2.8}$$

$$(u(\cdot, 0), v) = (\varphi_0, v) \,, \quad v \in \mathscr{V} \,, \tag{2.9}$$

where

$$\mathscr{V} = \{v \in W_2^{(1)}(\Omega),\ v = 0 \text{ for } X \in \Gamma_u\} \quad \text{and} \tag{2.10}$$

$$\mathscr{V} = \left\{v \in W_2^{(1)}(\Omega),\ \int_\Omega v \,\mathrm{d}X = 0\ ^{2)}\right\} \quad \text{if} \quad \Gamma_u = \Gamma_v = \emptyset \,, \quad \text{respectively,}$$

$$(u, v) = \int_\Omega uv \,\mathrm{d}X \,, \tag{2.11}$$

[1]) A bounded region $\Omega \subset E_N$ is called Lipschitz, if its boundary has the following properties: a) to each point $X \in \Gamma$ an open hypersphere S_X about X exists, such that the intersection $S_X \cap \Gamma$ may be described by means of a Lipschitz function and b) $S_X \cap \Gamma$ divides S_X into exterior and interior parts with respect to Ω.

[2]) A set $G \subset \Gamma$ will be called open in Γ, if for any point $X_0 \in G$ there exists an $\eta > 0$ such that each $X \in \Gamma$, satisfying dist $(X - X_0) < \eta$, belongs to G.

[3]) From there it follows, with the use of Schwartz theorem, that all the functions $a_{ij}(X)$ are bounded on $\bar{\Omega}$.

[2]) Or an equivalent condition — see e.g. [3].

$$[u, v]_A = \int_\Omega \sum_{i,j} a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} \, \mathrm{d}X + \int_{\Gamma_v} \alpha u v \, \mathrm{d}\Gamma \,, \tag{2.12}$$

$$(P, v)_\Gamma = \int_{\Gamma_h \cup \Gamma_v} P v \, \mathrm{d}\Gamma \,. \tag{2.13}$$

Moreover, we suppose that

$$\lim_{t \to 0+} \| u(\cdot, t) = u(\cdot, 0) \|_{W_2^{(1)}(\Omega)} = 0 \,. \tag{2.14}$$

Note that the formula (2.8) can be obtained, if we multiply (2.1) by v, integrate over Ω by parts and make use of the boundary conditions (2.4), (2.5) and the definition (2.10) of $\mathscr{V}$.

Setting $u = g + w$, the formulation (2.7), (2.8), (2.9) can be rewritten as follows

$$w \in L_2(I, \mathscr{V}) \,, \quad \frac{\partial w}{\partial t} \in \mathscr{C}(I, L_2(\Omega)) \,, \tag{2.15}$$

$$\left(\frac{\partial w}{\partial t}, v \right) + [w, v]_A = \langle \bar{f}, v \rangle \,, \quad 0 < t \leqq T \,, \quad v \in \mathscr{V} \,. \tag{2.16}$$

$$(w(\cdot, 0), v) = (\psi_0, v) \,, \quad v \in \mathscr{V} \,, \tag{2.17}$$

where

$$\langle \bar{f}, v \rangle = \left(f - \frac{\partial g}{\partial t}, v \right) - [g, v]_A + (P, v)_\Gamma \,, \tag{2.18}$$

$$\psi_0 = \varphi_0 - g(\cdot, 0) \,. \tag{2.19}$$

The problem (2.15)–(2.19) can be interpreted in the form similar to (I.1.1.), (I.1.2) in the sense of functionals on $\mathscr{V}$, i.e.,

$$\frac{\mathrm{d}w}{\mathrm{d}t} = A \, w(t) = \bar{f}(t) \,, \quad 0 < t \leqq T \,, \tag{2.20}$$

$$w(0) = \psi_0 \,,$$

where A is the second order differential operator of (2.1) with

$D(A) = \{ w \in C^{(2)}(\mathrm{w})$, w satisfies the homogeneous boundary conditions (2.3), (2.4), (2.5)$\}$,

$\bar{f}(t)$ is a linear functional on $\mathscr{V}$ for each $t \in I$, defined through (2.18) and ψ_0 is defined in (2.19).

Using the assumptions on f, $\partial g / \partial t$, a_{ij}, α, g, P, the Cauchy-Buniakowski inequality and the embedding theorems, we can deduce easily that $\bar{f}(t)$ is continuous on $\mathscr{V}$. Let us compare the notation of the present problem with that of Section I.1. If we set

$H = L_2(\Omega)$, define $\mathscr{V}$ by (2.10) (with $\|u\| = \|u\|_{W_2^{(1)}(\Omega)}$) and $[u, v]_A$ by (2.12), then the assumptions on $[u, v]_A$, (I.1.3) and (I.1.4) hold. In fact, from the boundedness of the spectrum of a_{ij}, (2.6) and the embedding theorem we obtain that

$$(2.21) \qquad [u, v]_A \leqq C\|u\| \, \|v\| \,, \quad u, v \in W_2^{(1)}(\Omega) \,,$$

i.e., the bilinear form is continuous on $\mathscr{V} \times \mathscr{V}$. The inequality (I.1.3) is evident. The first part of (I.1.4) is an immediate consequence of the integration by parts. In order to prove the inequality

$$(2.22) \qquad c_0\|u\|^2_{W_2^{(1)}} \leqq [u, u]_A \,,$$

we may deduce

$$[u, u]_A \geqq \eta \int_\Omega \sum_{i=1}^N \left(\frac{\partial u}{\partial x_i}\right)^2 \mathrm{d}X + \alpha_0 \int_{\Gamma_v} u^2 \, \mathrm{d}\Gamma \geqq$$

$$\geqq \min(\eta, \alpha_0) \left[\int_{\Gamma_v} u^2 \, \mathrm{d}\Gamma + \int_\Omega \sum_{i=1}^N \left(\frac{\partial u}{\partial x_i}\right)^2 \mathrm{d}X\right]$$

and consider the following cases separately:

a) $\Gamma_v \neq \emptyset$. Then the square root of the expression in brackets defines a norm equivalent with $\|u\|_{W_2^{(1)}}$ (see e.g. [4] Th. 1.1.9), consequently (2.22) holds.

b) $\Gamma_v = \emptyset$, $\Gamma_u \neq \emptyset$. Then the norms $\|u\|_{W_2^{(1)}}$ and

$$\left[\int_{\Gamma_u} u^2 \, \mathrm{d}\Gamma + \int_\Omega \sum_i \left(\frac{\partial u}{\partial x_i}\right)^2 \mathrm{d}X\right]^{1/2}$$

are equivalent (according to the same theorem in [4]), consequently (2.22) holds again, because $u \in \mathscr{V}$ vanishes on Γ_u.

c) $\Gamma_v = \emptyset$, $\Gamma_u = \emptyset$. Using Poincaré's inequality (or Theorem 2.3 and Remark 3 of [3]) we obtain the inequality

$$\sum_i \int_\Omega \left(\frac{\partial u}{\partial x_i}\right)^2 \mathrm{d}X \geqq c\|u\|^2_{W_2^{(1)}} \,,$$

which yields (2.22).

Hence all the derivation of the semi-variational approximations of Section I.1 can be applied, with the following minor changes: the functions $f(t)$ have to be replaced by the functionals $\bar{f}(t)$, all the products of the form $(f(t), v)$ by the expressions $\langle \bar{f}(t), v\rangle$ and φ_0 by ψ_0. For example, we come to the Crank-Nicholson-Galerkin approximation (I.1.17), (I.1.22), where now

$$\omega_{0j} = (\psi_0, v_j) \,,$$

$$(2.23a) \qquad F^0_{mj} = \tfrac{1}{2}\langle \bar{f}(m\tau) + \bar{f}(m\tau + \tau), v_j\rangle \quad \text{or}$$

$$(2.23b) \qquad F^0_{mj} = \langle \bar{f}(m\tau + \tfrac{1}{2}\tau), v_j\rangle \,.$$

Douglas and Dupont proved in [5] some a priori estimates for the latter procedure applied to non-linear equations of the type (2.1) with $a_{ij}\,(u,\, \partial u/\partial x_k,\, X,\, t)$, the boundary conditions (2.3), (2.4) ($\Gamma_v = \emptyset$) and for its linearization by means of the predictor-corrector method. The proofs of Theorems 7.2, 7.3 and Lemma 7.1 of [5] may be extended also to the mixed boundary conditions of the type (2.3)–(2.5) and consequently, the estimates from [5] hold for the linear C.N.G. approximation (I.1.22), (2.23b), as well. For the procedure (I.1.22), (2.23a), the estimate of Theorem 7.2 [5] can be proved easily with the norms in $H_0^{(1)}$ replaced by norms in $W_2^{(1)}(\Omega)$, if we suppose also (I.2.5). Replacing f by $\bar{f}$ and the products $(f(t), v)$ by $\langle \bar{f}(t), v \rangle$ also in Section I.2, the proofs of Theorems I.2.1 and I.2.3 remain without any other change. In Theorem I.2.2 and its proof we have to replace only $\mathring{W}_2^{(1)}$ by $W_2^{(1)}$.

3. A CLASS OF MORE GENERAL EQUATIONS

The method of semi-variational approximations may be easily applied to a class of abstract problems of more general type, namely to the following initial value problem

$$B\,\frac{\mathrm{d}u}{\mathrm{d}t} + Au = f\,, \quad 0 < t \leqq T\,, \tag{3.1}$$

$$u(0) = \varphi_0\,,$$

where A and B are linear symmetric and positive definite operators in a real Hilbert space H, which do not depend on t. We assume that two Hilbert spaces $\mathscr{V}_0$, $\mathscr{V}_1$ with the norms $\|u\|_0$ and $\|u\|_1$, respectively, a bilinear form $[u, v]_A$, continuous and symmetric on $\mathscr{V}_0 \times \mathscr{V}_0$, a bilinear form $[u, v]_B$, continuous and symmetric on $\mathscr{V}_1 \times \mathscr{V}_1$ and positive constants c, α, β exist, such that $\mathscr{V}_0, \mathscr{V}_1 \subset H$, the domains $D(A) \subset \mathscr{V}_0$, $D(B) \subset \mathscr{V}_1$,

$$(Au, v) = [u, v]_A\,, \quad u, v \in D(A)\,, \tag{3.2}$$

$$(Bu, v) = [u, v]_B, \quad u, v \in D(B)\,,$$

$$\alpha\|u\|_0^2 \leqq [u, u]_A\,, \quad u \in \mathscr{V}_0\,, \tag{3.3}$$

$$\beta\|u\|_1^2 \leqq [u, u]_B\,, \quad u \in \mathscr{V}_1\,,$$

$$\mathscr{V}_0 \subset \mathscr{V}_1\,, \quad \|u\|_1 \leqq c\|u\|_0\,, \quad u \in \mathscr{V}_0\,. \tag{3.4}$$

Furthermore, $f(t)$ is assumed to be a linear continuous functional on $\mathscr{V} = \mathscr{V}_0 \cap \mathscr{V}_1$ for each $t \in I = \langle 0, T \rangle$ and $\varphi_0 \in \mathscr{V}_1$.

Obviously, if B is the identity operator, $\mathscr{V}_1 = H$ and $\mathscr{V}_0 = \mathscr{V}$, then the present problem reduces to that of the abstract parabolic equation (I.1.1.) and all the assumptions (I.1.4) are satisfied.

The derivation of the first semi-variational approximation goes through similarly to that of Section I.1, with the following minor changes:

a) all the scalar products of the type (u, v) in (I.1.9)–(I. 1.16) should be replaced by $[u, v]_B$, consequently G_{ij} by $\mathscr{B}_{ij} = [v_i, v_j]_B$ and the matrix G by the matrix $\mathscr{B}$,

b) all products concerning the right-hand side of the form $(f(t), v)$ should be substituted by $\langle f(t), v\rangle$.

The projection condition (I.1.18) remains unchanged. Thus we obtain the system

$$(3.5) \qquad \left(\mathscr{B} + \frac{\tau}{2}\mathscr{A}\right)\mathbf{a}_m = \mathscr{B}\mathbf{w}_m + \frac{\tau}{4}\left[\mathbf{F}(m\tau) + \mathbf{F}(m\tau + \tau)\right],$$
$$\mathscr{B}\mathbf{w}_0 = \omega_0\,, \quad \mathbf{w}_{m+1} = 2\mathbf{a}_m - \mathbf{w}_m\,,$$

which is equivalent to the Crank-Nicholson-Galerkin scheme

$$(3.6) \qquad \frac{1}{\tau}\left[U_{m+1} - U_m, V\right]_B + \tfrac{1}{2}\left[U_{m+1} + U_m, V\right]_A = \tfrac{1}{2}\langle f(0) + f(\tau), V\rangle\,,$$
$$0 \leqq m \leqq T/\tau - 1\,, \quad V = v_j\,, \quad j = 1, \ldots, N\,,$$
$$[U_0, V]_B = [\varphi_0, V]_B\,.$$

As the second approximation is concerned, similar modifications lead to the following system

$$(3.7) \qquad \mathscr{B}\mathbf{c}_m - \left(\frac{\tau}{12}\mathscr{A} + \frac{1}{2}\mathscr{B}\right)\mathbf{b}_m = \mathscr{B}\mathbf{w}_m + \frac{\tau}{12}\left[\mathbf{F}(m\tau) - \mathbf{F}(m\tau + \tau)\right],$$
$$\mathscr{A}\mathbf{c}_m + \frac{1}{\tau}\mathscr{B}\mathbf{b}_m = \frac{1}{6}\left[\mathbf{F}(m\tau) + 4\,\mathbf{F}\left(m\tau + \frac{\tau}{2}\right) + \mathbf{F}(m\tau + \tau)\right]$$

together with (I.1.42) and (I.1.43). Remark I.1.2 remains in force. Modifications of the third approximation are analogous.

Remark II.3.1. It is easy to see that the matrix $\mathscr{B}$ is also positive definite. Hence (3.5) has a unique solution for every m and any τ. The solvability of (3.7) at each time step can be proved like in Remark I.2.1, replacing only G by $\mathscr{B}$.

Let us investigate the convergence of the first and second approximations. Assume that the mapping $f(t)$ is continuous on I and the solution u of the problem (3.1) is such that (cf. Section I.2)

$$(3.8) \qquad u \in L_2(I, \mathscr{V})\,, \quad \mathrm{d}u/\mathrm{d}t \in \mathscr{C}(I, \mathscr{V}_1)\,,$$

$$(3.9) \qquad [\mathrm{d}u/\mathrm{d}t, v]_B + [u, v]_A = \langle f, v\rangle\,, \quad 0 < t \leqq T\,, \quad v \in \mathscr{V}\,,$$

$$(3.10) \qquad [u(0), v]_B = [\varphi_0, v]_B\,, \quad v \in \mathscr{V}\,,$$

$$(3.11) \qquad \lim_{t \to 0+} \|u(t) - u(0)\|_0 = 0\,.$$

Using the basic ideas of the proof of Theorem I.2.1, we obtain

Theorem II.3.1. Suppose that the solution u of (3.1) satisfies (3.8)–(3.11), possesses continuous derivatives in $\mathscr{V}_1$ up to the second order and the norms $\|\mathrm{d}^3u/\mathrm{d}t^3\|_1$ are bounded for $0 < t < T$. Denote $z_m = u_m - U_m$, where U_m is the solution of (3.6), $u_m = u(m\tau)$, $\tilde{u}$ any function of the form $\tilde{u}(t) = \sum_{i=1}^{N} \alpha_i(t)\, v_i$, $s_{m+1/2} = \frac{1}{2}(s_m + s_{m+1})$, δ_{ik} the Kronecker's delta.

Then there exist positive constants C and τ_0, independent of τ, such that for $\tau \leqq \tau_0$ and any k, $1 \leqq k \leqq T/\tau$, it holds

$$\|z_k\|_1^2 + \sum_{m=0}^{k-1} \tau \|z_{m+1/2}\|_0^2 \leqq \tag{3.12}$$

$$\leqq C \left\{ \sum_{m=0}^{k-1} \tau \left[\|(u - \tilde{u})_{m+1/2}\|_0^2 + (1 - \delta_{1k}) \left\| \frac{1}{\tau}\, \delta(u - \tilde{u})_{m-1/2} \right\|_1^2 \right] + \right.$$

$$\left. + \|(u - \tilde{u})_0\|_1^2 + \|(u - \tilde{u})_{1/2}\|_1^2 + \|(u - \tilde{u})_{k-1/2}\|_1^2 + \tau^4 \right\}.$$

Remark II. 3.2. In case that (3.4) fails to hold, Theorem II.3.1 remains in force, if the term $\|(u - \tilde{u})_{m+1/2}\|_1^2$ is added in the square bracket of the right-hand side of (3.12)

The system (3.7) of the second approximation is equivalent to the following finite difference scheme (cf. (I.2.2)–(I.2.3))

$$\frac{4}{\tau}\left[U_m - 2U_{m+1/2} + U_{m+1}, v_j\right]_B + \left[U_{m+1} - U_m, v_j\right]_A = \langle f_{m+1} - f_m, v_j \rangle\,, \tag{3.13}$$

$$\frac{1}{\tau}\left[U_{m+1} - U_m, v_j\right]_B + \tfrac{1}{6}\left[U_m + 4U_{m+1/2} + U_{m+1}, v_j\right]_A = \tfrac{1}{6}\langle f_m + 4f_{m+1/2} + f_{m+1}, v_j \rangle\,,$$

$$\left[U_0, v_j\right]_B = \left[\varphi_0, v_j\right]_B\,,$$

$$0 \leqq m \leqq T/\tau - 1\,, \quad j = 1, \ldots, N\,.$$

Note that here $U_{m+1/2} = u^{(2)}(m\tau + \frac{1}{2}\tau)$, $f_{m+1/2} = f(m\tau + \frac{1}{2}\tau)$.

Theorem II.3.2. *Suppose that the solution u of (3.1) satisfies (3.8)–(3.11), possesses continuous derivatives in $\mathscr{V}_1$ up to the fourth order on $\langle 0, T\rangle$ and the norms $\|\mathrm{d}^5u/\mathrm{d}t^5\|_1$ are bounded for $0 < t < T$. Denote $z_m = u_m - U_m$ where U_m is the solution of (3.13), $u_m = u(m\tau)$, $\tilde{u}$ as in the Theorem II.3.1, $s_m^{\wedge} = \frac{1}{6}(s_m + 4s_{m+1/2} + s_{m+1})$, $\delta s_m = s_{m+1} - s_m$.*

Then there exist positive constants C and τ_0, independent of τ, such that for $\tau \leqq \tau_0$ and any k, $1 < k \leqq T/\tau$, it holds

$$(3.14)\qquad \|z_k\|_1^2 + \sum_{m=0}^{k-1} \tau(\|z_m^{\wedge}\|_0^2 + \|\delta z_m\|_0^2 + \|z_m\|_1^2) \leqq$$

$$\leqq C \left\{ \sum_{m=0}^{k-1} \tau \left[\|(u - \tilde{u})_m^{\wedge}\|_0^2 + \left\| \frac{1}{\tau} \delta(u - \tilde{u})_m \right\|_1^2 + \|\delta(u - \tilde{u})_m\|_0^2 \right] + \right.$$

$$\left. + \sum_{m=0}^{k-2} \tau \left\| \frac{1}{\tau} \delta(u - \tilde{u})_{m+1/2} \right\|_1^2 + \|(u - \tilde{u})_0\|_1^2 + (\|u - \tilde{u})_0^{\wedge}\|_1^2 + \|(u - \tilde{u})_{k-1}^{\wedge}\|_1^2 + \tau^8 \right\}.$$

Moreover, for $k = 1$ we have the estimate

$$(3.15)\qquad \|z_1\|_1^2 + \tau(\|z_0^{\wedge}\|_0^2 + \|\delta z_0\|_0^2) \leqq$$

$$\leqq C \left\{ \tau \left[\|(u - \tilde{u})_0^{\wedge}\|_0^2 + \left\| \frac{1}{\tau} \delta(u - \tilde{u})_0 \right\|_1^2 + \|\delta(u - \tilde{u})_0\|_0^2 \right] + \|(u - \tilde{u})_0\|_1^2 + \tau^8 \right\}.$$

Proof is nearly the same as that of Theorem I.2.1 with some obvious minor changes. Namely, instead of (u, v) and $|u|$ we employ the bilinear form $[u, v]_B$ and $[u, u]_B^{1/2} = \|u\|_B$, respectively and use the inequalities (3.3). Note that, in contrast with Remark II.3.2, we are not able to modify the proof of Theorem I.2.1 so simply, unless (3.4) holds.

The estimates (3.12) or (3.14)–(3.15) can be used to get rates of convergence. To this end, let us consider the following mixed problem

$$(3.16)\qquad -\Delta \frac{\partial u}{\partial t} + \Delta\Delta u = f\,, \quad 0 < t \leqq T\,,$$

$$u(\cdot, 0) = \varphi_0\,,$$

$$u = 0\,, \quad \partial u/\partial \nu = 0\,, \quad (x_1, x_2) \in \partial\Omega\,,$$

where $\Omega = (0,1) \times (0,1)$, Δ is the Laplace operator, ν denotes the normal to the boundary $\partial\Omega$, $\varphi_0 \in \mathring{W}_2^{(1)}(\Omega)$ and f a linear continuous functional on $\mathring{W}_2^{(2)}(\Omega)$. The problem (3.16) represents a particular case of (3.1). In fact, if we set

$$H = L_2(\Omega)\,, \quad B = -\Delta\,, \quad A = \Delta\Delta\,,$$

$$D(B) = \{u \in C^{(2)}(\Omega),\ u = 0 \text{ on } \partial\Omega\}\,,$$

$$D(A) = \{u \in C^{(4)}(\Omega),\ u = \partial u/\partial \nu = 0 \text{ on } \partial\Omega\}\,, \quad \mathscr{V}_0 = \mathring{W}_2^{(2)}(\Omega)\,,$$

$$\mathscr{V}_1 = \mathring{W}_2^{(1)}(\Omega)\,, \quad \mathscr{V} = \mathscr{V}_0 \cap \mathscr{V}_1 = \mathring{W}_2^{(2)}(\Omega)\,,$$

$$[u, v]_A = \int_\Omega \Delta u\, \Delta v \,\mathrm{d}X\,, \quad [u, v]_B = \int_\Omega \sum_{i=1}^{2} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} \,\mathrm{d}X\,,$$

then the assumptions on the bilinear forms and domains, (3.2), (3.3) and (3.4) can be verified easily.

Let us employ the Hermite interpolation theory in the (x_1, x_2)-plane, as in Section I.2. Then we come to the following

Theorem II.3.3. *Let $\mathscr{M} = H_h^{(n)} \cap \overset{\circ}{W}{}_2^{(2)}(\Omega)$, $n \geqq 2$. Let u, U_m, z_m, $z_{m+1/2}$ be as in Theorem* II.3.1, *related to the problem* (3.16). *Suppose that for each $t \in \langle 0, T\rangle$ u and $\partial u/\partial t$ satisfy the hypotheses of Lemma* I.2.1 *and that*

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

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

where C' is independent of t, $\chi \in L_2(0, T)$, D^α denotes spatial derivatives only.

Then there exist constants C, τ_0, independent of h, τ such that for $\tau \leqq \tau_0$ and any k, $1 \leqq k \leqq T/\tau$ it holds

$$\|z_k\|^2_{{}^{\circ}W_2{}^{(1)}} + \sum_{m=0}^{k-1} \tau \|z_{m+1/2}\|^2_{{}^{\circ}W_2{}^{(2)}} \leqq C\left(h^{2(2n-2)} + \tau^4\right).$$

Theorem II.3.4. Let $\mathscr{M}$, u, $\partial u/\partial t$ be as in Theorem II.3.3 and U_m, z_m, $z_m^\wedge$, δz_m as in Theorem II.3.2, related to the problem (3.16).

Then there exist constants C, τ_0, independent of h, τ such that for $\tau \leqq \tau_0$ and any k, $1 < k \leqq T/\tau$ it holds

$$\|z_k\|^2_{{}^{\circ}W_2{}^{(1)}} + \sum_{m=0}^{k-1} \tau\left(\|z_m^\wedge\|^2_{{}^{\circ}W_2{}^{(2)}} + \|\delta z_m\|^2_{{}^{\circ}W_2{}^{(2)}} + \|z_m\|^2_{{}^{\circ}W_2{}^{(1)}}\right) \leqq$$
$$\leqq C\left(h^{2(2n-2)} + \tau^8\right).$$

Proofs of both Theorems are analogous to that of Theorem I.2.2.

Theorem II.3.5. Let $f = 0$ in (3.6). Then

$$(3.17) \qquad \beta^{1/2}\|U_{m+1}\|_1 \leqq \|U_{m+1}\|_B \leqq \|U_m\|_B \leqq \|\varphi_0\|_B \leqq C\|\varphi_0\|_1$$

holds for every $m = 0, 1, \ldots, T/\tau - 1$.

Proof follows directly from (3.6), if we insert $V = U_m + U_{m+1}$, use (3.3) and the Schwartz's inequality.

Theorem II.3.6. Let $f = 0$ in (3.13). Then (3.17) and

$$\|U_m^\wedge\|_B \leqq \|U_m\|_B , \quad \|U_{m+1/2}\|_B \leqq 2\|U_m\|_B$$

hold for every $m = 0, 1, \ldots, T/\tau - 1$.

Proof is analogous to that of Theorem I.2.3.

APPENDIX

There exists an alternative approach to derive the semi-variational approximations. Let us consider again the case of homogeneous equation (I.1.1).

Theorem II.3.7. Denote $\{S_k^{(m)}\}_{k=0}^m$ a basis in $\mathscr{P}_m$ (cf. Section 1). Let us set

$$\text{(A.1)} \qquad u^{(n)} = \sum_{k=0}^{n} \sum_{i=1}^{N} b_i^{(k)} S_k^{(n)}(t)\, v_i\,,$$

$$\text{(A.2)} \qquad \int_0^\tau \left(\frac{\mathrm{d}u^{(n)}}{\mathrm{d}t} + Au^{(n)}, S_j^{(n-1)}(t)\, v_m\right) \mathrm{d}t = 0\,, \quad j = 0, 1, \ldots, n-1\,,$$

$$\text{(A.3)} \qquad (v_m, u^{(n)}(0)) = (v_m, \varphi_0)\,, \quad m = 1, 2, \ldots, N\,.$$

Then the n-th semi-variational approximation is determined uniquely by the conditions (A.2), (A.3), being independent of the choice of the polynomial bases.

Proof. First let us consider the Legendre polynomials $\bar{P}_k = S_k^{(n)} = S_k^{(n-1)}$. Using the formula

$$\frac{\mathrm{d}}{\mathrm{d}x} P_k(x) = \sum_{\substack{2s+1 \leqq k \\ s \geqq 0}} (2k - 4s - 1)\, P_{k-2s-1}(x)\,,$$

the equations (A.1)–(A.3) with $b_i^{(k)} = \beta_i^{(k)}$ lead to the following system

$$\text{(A.4)} \qquad \sum_{k=0}^{n} \sum_{i=1}^{N} \beta_i^{(k)} (G_{im} R_{jk} + \mathscr{A}_{im} p_{jk}) = 0\,, \quad 0 \leqq j \leqq n-1\,,$$

$$\sum_{i=1}^{N} G_{mi} \Big(\sum_{k=0}^{n} \beta_i^{(k)} \bar{P}_k(0)\Big) = \omega_{0m}\,, \quad m = 1, \ldots, N\,,$$

where

$$p_{jk} = \frac{\tau}{2j+1}\, \delta_{jk}$$

was introduced in the proof of Theorem II.1.1 and

$$\text{(A.5)} \qquad R_{jk} = \int_0^\tau \bar{P}_j \frac{\mathrm{d}}{\mathrm{d}t} \bar{P}_k \,\mathrm{d}t = \begin{cases} 2 & \text{if } j = k - 2s - 1,\ s \geqq 0\,, \\ 0 & \text{otherwise.} \end{cases}$$

The system (A.4) may be rewritten as follows

$$\text{(A.6)} \qquad B\beta = g\,,$$

$$\text{(A.7)} \qquad B = \begin{bmatrix} G, & -G, & G, & \ldots, & (-1)^n G \\ \tau\mathscr{A}, & 2G, & 0, & 2G, & \ldots \\ & \frac{\tau}{3}\mathscr{A}, & 2G, & 0, & \ldots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & \frac{\tau}{2n-3}\mathscr{A}, & 2G, & 0 & \\ & & \frac{\tau}{2n-1}\mathscr{A}, & 2G & \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta^{(0)} \\ \beta^{(1)} \\ \vdots \\ \vdots \\ \beta^{(n)} \end{bmatrix}, \quad g = \begin{bmatrix} \omega_0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}.$$

Recalling the system (1.14) (or (1.17)), (1.7) and comparing it with (A.6), (A.7), we observe that, in the block matrix form,

$$D = \begin{bmatrix} G\, P^T(0) \\ D' \end{bmatrix} = LB\,, \tag{A.8}$$

where

$$D' = [\mathscr{D}^{(n)}, 0]\,,$$

$$\begin{aligned} L_{1j} &= \delta_{1j} \\ L_{2j} &= 2\delta_{1j} + (\delta_{2j} - \delta_{3j}) \quad \text{for} \quad n \geqq 2 \quad \text{or} \quad L_{2j} = 2\delta_{1j} + \delta_{2j} \quad \text{for} \quad n = 1\,, \\ L_{kj} &= \delta_{k-1,j} - \delta_{k+1,j} \quad \text{for} \quad 3 \leqq k \leqq n\,, \\ L_{n+1,j} &= \delta_{nj}\,, \quad \text{for} \quad n \geqq 2\,. \end{aligned}$$

We can see easily that det $|L| = 1$. As D is regular [cf. Lemma II.1.1], by virtue of (A.8) the matrix B is also regular. Moreover

$$(Lg)^T = (\omega_0, 2\omega_0, 0, 0, \ldots, 0)$$

and consequently, the system (A.6) is equivalent to the system (1.17), (1.7) of Section 1.

In order to prove the independence of the choice of bases, let us recall (1.19) with regular matrices h and C and set

$$v^{(n)}(t) = \sum_{k=0}^{n} \sum_{i=1}^{N} b_i^{(k)}\, S_k^{(n)}(t)\, v_i\,.$$

We obtain from (A.2), (A.3) and (1.19)

$$\begin{aligned} C(GR + \mathscr{A}p)\, h^T b &= 0\,, \\ G\, P^T(0)\, h^T b &= \omega_0\,. \end{aligned} \tag{A.9}$$

By comparison of (A.9) with (A.4) we conclude that

$$h^T b = \beta$$

and using the latter result, we deduce

$$v^{(n)}(t) = \sum_{k=0}^{n} \sum_{i=1}^{N} S_n^{(k)}(t)\, b_i^{(k)} v_i = \sum_{k,r} \sum_{i} \bar{P}_r(t)\, h_{kr} b_i^{(k)} v_i = \sum_{r,i} \beta_i^{(r)}\, \bar{P}_r(t)\, v_i = u^{(n)}(t)\,, \quad \text{Q.E.D.}$$

References

[1] *I. Hlaváček:* On a semi-variational method for parabolic equations I. Aplikace matematiky 17 (1972), 5, 327—351.

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

[3] *I. Hlaváček, J. Nečas:* On inequalities of Korn's type. Archive for Rational Mechanics and Analysis, 36, 4, 1970, 305—344.

[4] *J. Nečas:* Les méthodes directes en théorie des équations elliptiques. Prague, Academia 1967.

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

[6] *R. S. Varga:* Matrix Iterative Analysis, Prentice-Hall, 1962.

Souhrn

O JEDNÉ POLOVARIAČNÍ METODĚ PRO PARABOLICKÉ ROVNICE II

IVAN HLAVÁČEK

V druhé části práce jsou dokázány další vlastnosti n-té polovariační aproximace řešení daného problému s homogenní rovnicí: nezávislost na volbě polynomiální báze v t a úzká souvislost s Padéovou aproximací. Dále je metoda aplikována na počáteční úlohu pro parabolickou rovnici s nehomogenními okrajovými podmínkami a na obecnější abstraktní problém se dvěma positivně definitními operátory.

Author's address: Ing. *Ivan Hlaváček*, CSc., Matematický ústav ČSAV v Praze, Žitná 25, 115 67 Praha 1.