Discrete Time Galerkin Methods for a Parabolic Boundary Value Problem (*) (**).

JAMES H. BRAMBLE (Ithaca, N.Y., USA) - VIDAR THOMÉE (Göteborg, Sweden)

Summary. – *Single step discrete time Galerkin methods for the mixed initial-boundary value problem for the heat equation are studied. Two general theories leading to error estimates are developed. Among the examples analyzed in the application of these theories are methods in which the related quadratic form is required to be definite only on the subspace of approximating functions and two classes containing methods of arbitrary given order of accuracy, one requiring satisfaction of certain boundary conditions by the elements of the subspace, the other making no such requirements.*

1. – Introduction.

The purpose of this paper is to continue the line of investigation of our previous paper [3]. There we considered the initial boundary value problem for the heat equation in a cylinder under homogeneous boundary conditions. The methods studied consist in discretizing with respect to time and solving approximately the resulting elliptic problem for fixed time by least squares methods similar to those of BRAMBLE and SCHATZ [2]. The approximating functions in the least squares method were not required to satisfy prescribed homogeneous boundary conditions so that the methods were applicable to domains of general shape.

Here we abstract the essential features of [3] into a general theorem (Theorem 1) which can be applied to extend results of [3] to a class of single time step methods which include methods of arbitrary given order of accuracy.

In PRICE and VARGA [7] and DOUGLAS and DUPONT [4] the initial-boundary value problem is approximated instead by first projecting into a finite dimensional space of approximating functions in the space variables, keeping the differentiation with respect to time. In order to estimate the error, an auxiliary approximation to a related elliptic problem is utilized. This technique makes it possible to make less stringent approximation assumptions of the approximating spaces than in Theorem 1.

Our second theorem (Theorem 2) takes advantage of the features of this method. In doing so we notice that a certain related quadratic form in this case need be definite only on the subspace of approximating functions. This allows us to include among our examples an extension to the parabolic case of a method for treating Dirichlet's problem due to NITSCHE [6].

(*) Entrata in Redazione il 19 settembre 1972.

(**) Sponsored by the United States Army under Contract No.: DA-31-124-ARO-D-462.

The aim of each of our two theorems is to provide simple sufficient conditions on single step discrete time Galerkin methods which lead to error estimates. Some of the examples which we then study satisfy the conditions of Theorem 1, some satisfy those of Theorem 2 and some fall into both categories. Neither of the two theorems is stronger than the other.

An outline of the paper is as follows. In the next section the problem to be studied is defined precisely and certain related technical lemmas are given. In Sections 3 and 4 the two general theorems are presented and proved. Section 5 contains some technical estimates concerning the consistency of the Galerkin equations with the initial boundary value problem, which are needed in the applications. Sections 6, 7 and 8 present the various examples. In Section 6 we study methods which require that elements of the subspace in which we seek the approximations satisfy certain homogeneous boundary conditions. Utilizing such subspaces, methods of arbitrarily high order of accuracy are constructed. In Section 7 some methods in which the subspaces need not satisfy the above mentioned boundary requirements are shown to fit into the theories of Sections 3 and/or 4. Here methods are described for which the related quadratic forms are definite only on the approximating subspaces. The final section is devoted to least squares methods. The purely implicit method of [3] is contained here as a special case and Theorem 2 is seen to yield some new error estimates for that method. Again methods are constructed with arbitrarily given order of accuracy but without requiring prescribed boundary conditions to be satisfied by the approximating functions.

Throughout this paper, C and c will denote positive constants, not necessarily the same at different occurrences.

2. – Preliminaries.

Let Ω be a bounded domain in Euclidean N-space R^N with C^∞ boundary $\partial\Omega$. We shall use the following notation for inner products and norms in the real function spaces $L^2(\Omega)$ and $L^2(\partial\Omega)$, respectively, namely

$$(v, w) = \int_\Omega v(x)w(x)\,dx\,, \qquad \|v\| = (v, v)^{\frac{1}{2}}\,,$$

$$\langle v, w\rangle = \int_{\partial\Omega} v(x)w(x)\,ds\,, \qquad |v| = \langle v, v\rangle^{\frac{1}{2}}\,.$$

Other norms will be distinguished by use of subscripts. In particular we shall use the norm in $H^s = W_2^s(\Omega)$ for s a positive integer,

$$\|v\|_{H^s} = \Big(\sum_{|\alpha|\leqslant s} \|D^\alpha v\|^2\Big)^{\frac{1}{2}}\,.$$

We shall also frequently use the Dirichlet form

$$D(v, w) = \int_\Omega \sum_{j=1}^N \frac{\partial v}{\partial x_j}\frac{\partial w}{\partial x_j}\,dx\,.$$

We shall consider the approximate solution of the following mixed initial-boundary value problem for $u = u(x, t)$, namely

$$\frac{\partial u}{\partial t} = \Delta u \equiv \sum_{j=1}^{N} \frac{\partial^2 u}{\partial x_j^2} \text{ in } \Omega \times (0, T], \qquad T > 0, \tag{2.1}$$

$$u = 0 \text{ on } \partial\Omega \times [0, T], \tag{2.2}$$

$$u(x, 0) = v(x) \text{ in } \Omega. \tag{2.3}$$

We associate with this problem the eigenvalue problem

$$\Delta\varphi + \lambda\varphi = 0 \text{ in } \Omega, \qquad \varphi = 0 \text{ on } \partial\Omega, \tag{2.4}$$

about which the following is well-known.

LEMMA 2.1. – *The eigenvalue problem* (2.4) *admits a nondecreasing sequence* $\{\lambda_m\}_1^\infty$ *of positive eigenvalues (which tend to* $+\infty$ *with* m*) and a corresponding sequence* $\{\varphi_m\}_1^\infty$ *of eigenfunctions which constitute an orthonormal basis in* $L^2(\Omega)$; *every* $v \in L^2(\Omega)$ *may be represented as*

$$v(x) = \sum_{m=1}^{\infty} v_m \varphi_m(x), \qquad v_m = (v, \varphi_m), \tag{2.5}$$

and Parseval's relation

$$(v, w) = \sum_{m=1}^{\infty} v_m w_m, \qquad v_m = (v, \varphi_m), \qquad w_m = (w, \varphi_m),$$

holds.

In the sequel we shall often work with the eigenfunction expansions of functions $v \in L^2(\Omega)$; v_m will then without explicit mention, as in (2.5), denote (v, φ_m), and $\sum_m$ will denote $\sum_{m=1}^{\infty}$.

For s non-negative, let $\dot{H}^s$ be the subspace of $L^2(\Omega)$ defined by the norm

$$\|v\|_{\dot{H}^s} = \Big(\sum_m \lambda_m^s v_m^2\Big)^{\frac{1}{2}},$$

and let $\dot{H}^\infty = \bigcap_{s>0} \dot{H}^s$. Notice in particular that for each m, $\varphi_m \in \dot{H}^\infty$.

The spaces $\dot{H}^s$ for s a non-negative integer can also be characterized as follows:

LEMMA 2.2. – *For* s *a non-negative integer,*

$$\dot{H}^s = \{v;\, v \in H^s,\, \Delta^j v = 0 \text{ on } \partial\Omega \text{ for } j < \tfrac{1}{2}s\}.$$

In particular,

$$\dot{H}^\infty = \{v;\, v \in C^\infty(\bar{\Omega}),\ \Delta^j v = 0 \text{ on } \partial\Omega \text{ for all } j\}\,.$$

PROOF. – We shall first prove that if $v \in H^s$ and $\Delta^j v = 0$ on $\partial\Omega$ for $j < \frac{1}{2}s$ then $v \in \dot{H}^s$. For $s = 1$ we have easily for $v \in C_0^\infty(\Omega)$,

$$\lambda_m v_m = (v, \lambda_m \varphi_m) = -(v, \Delta\varphi_m) = -(\Delta v, \varphi_m)\,,$$

and hence

$$\|v\|^2_{\dot{H}^1} = \sum_m \lambda_m v_m^2 = -\sum_m (v, \varphi_m)(\Delta v, \varphi_m) = -(v, \Delta v) = D(v, v)\,,$$

and since $C_0^\infty(\Omega)$ is dense in the subspace $\{v;\ v \in H^1,\ v = 0 \text{ on } \partial\Omega\}$ of H^1, this proves the result for $s = 1$. For $s = 2p + 1$ we have by the result for $s = 1$,

$$\begin{aligned}\|v\|^2_{\dot{H}^{2p+1}} &= \sum_m \lambda_m^{2p+1} v_m^2 = \sum_m \lambda_m (v, \lambda_m^p \varphi_m)^2 \\ &= \sum_m \lambda_m \big((-\Delta)^p v, \varphi_m\big)^2 = D(\Delta^p v, \Delta^p v) < \infty\,.\end{aligned}$$

For $s = 2p$ finally we have

$$\|v\|^2_{\dot{H}^{2p}} = \sum_m \lambda_m^{2p} v_m^2 = \sum_m (v, \lambda_m^p \varphi_m)^2 = \sum_m \big((-\Delta)^p v, \varphi_m\big)^2 = \|\Delta^p v\|^2 < \infty\,.$$

We now prove the opposite inclusion. Consider the case $s = 2p$ and let $\tilde{v}$ be any linear combination of finitely many of the eigenfunctions φ_m. Then by the above computation,

$$\|\Delta^p \tilde{v}\| = \|\tilde{v}\|_{\dot{H}^{2p}}\,.$$

On the other hand, by a well-known a priori inequality for the elliptic operator Δ^p, we have, since $\Delta^j \tilde{v} = 0$ on $\partial\Omega$ for $j < p$,

$$\|\tilde{v}\|_{H^{2p}} \leqslant C\|\Delta^p \tilde{v}\|\,.$$

Since the $\tilde{v}$ are dense in $\dot{H}^{2p}$ we conclude

$$\|v\|_{H^{2p}} \leqslant C\|v\|_{\dot{H}^{2p}}\,, \qquad v \in \dot{H}^{2p}\,,$$

and hence $\dot{H}^{2p} \subset H^{2p}$. Since

$$|\Delta^j v| \leqslant C\|v\|_{H^{2p}}\,, \qquad j < p\,,$$

and the $\Delta^j \tilde{v}$ vanish on $\partial\Omega$ we conclude that this holds for $\Delta^j v$ also. The proof for odd s is similar.

For the mixed initial-boundary value problem we have the following well-known result:

LEMMA 2.3. – For $v \in L^2(\Omega)$ the problem (2.1), (2.2), (2.3) admits a unique solution in $\dot{H}^\infty$ for $t > 0$ which can be represented as

$$u(x, t) = \big(E(t)v\big)(x) = \sum_m \exp[-t\lambda_m] v_m \varphi_m(x) . \tag{2.6}$$

The solution operator is bounded in $\dot{H}^s$ for any $s \geqslant 0$,

$$\|E(t)v\|_{\dot{H}^s} \leqslant \|v\|_{\dot{H}^s}, \qquad v \in \dot{H}^s, \tag{2.7}$$

and for $0 \leqslant s \leqslant l$ there is a constant C such that

$$\|E(t)v\|_{\dot{H}^l} \leqslant Ct^{-\frac{1}{2}(l-s)} \|v\|_{\dot{H}^s}, \qquad v \in \dot{H}^s. \tag{2.8}$$

PROOF. – Obviously, (2.6) defines for $v \in L^2(\Omega)$ a solution which is in $\dot{H}^\infty$ for $t > 0$. The inequalities (2.7) and (2.8) follow from

$$\|E(t)v\|_{\dot{H}^l} = \Big(\sum_m \lambda_m^l \exp[-2\lambda_m t] v_m^2\Big)^{\frac{1}{2}} \leqslant Ct^{-\frac{1}{2}(l-s)} \Big(\sum_m \lambda_m^s v_m^2\Big)^{\frac{1}{2}} = Ct^{-\frac{1}{2}(l-s)} \|v\|_{\dot{H}^s},$$

where

$$C = \sup_{\tau > 0} \tau^{\frac{1}{2}(l-s)} \exp(-\tau) .$$

The uniqueness follows by the standard energy identity

$$\frac{d}{dt} \|u\|^2 = -2D(u, u) .$$

We note for later use the following:

LEMMA 2.4. – For any $s \geqslant 0$,

$$\|\big(E(t) - I\big) v\|_{\dot{H}^s} \leqslant t \|v\|_{\dot{H}^{s+2}}, \qquad v \in \dot{H}^{s+2} .$$

PROOF. – We have

$$\|\big(E(t) - I\big) v\|_{\dot{H}^s} = \Big(\sum_m \lambda_m^s (\exp[-t\lambda_m] - 1)^2 v_m^2\Big)^{\frac{1}{2}} \leqslant \Big(\sum_m \lambda_m^s (t\lambda_m)^2 v_m^2\Big)^{\frac{1}{2}} = t\|v\|_{\dot{H}^{s+2}},$$

which proves the lemma.

As will be described in the subsequent sections, the approximate solution of (2.1), (2.2), (2.3) will be obtained by first discretizing in some way the equation (2.1) in time using a time step k. In each instance this will introduce an elliptic problem depending on the parameter k. An approximate solution of this elliptic problem will then be sought in a finite dimensional space S_h depending on a small parameter h which can be though of as an analogue of the mesh-width in the finite difference theory. We shall introduce the «mesh-ratio» $\lambda = k/h^2$ and always assume below that it is kept constant as h and k tend to zero.

In addition to the norms defined above in H^s and $\dot{H}^s$ we shall use the following norms in which the derivatives are weighted depending on their order, namely

$$\|v\|_{H_h^s} = \Big(\sum_{|\alpha|\leqslant s} h^{2|\alpha|}\|D^\alpha v\|^2\Big)^{\frac{1}{2}},$$

and

$$\|v\|_{\dot{H}_h^s} = \Big(\sum_m (1 + k\lambda_m)^s v_m^2\Big)^{\frac{1}{2}}.$$

In the same way as above, since k/h^2 is constant, these norms are equivalent, uniformly in h, for s an integer and $v \in \dot{H}^s$. The latter norm is again defined for s not necessarily an integer. For different h the corresponding spaces H_h^s contain the same elements but their Hilbert space structures are different. The same holds for $\dot{H}_h^s$.

We shall need the following interpolation lemma.

LEMMA 2.5. – Let $0 \leqslant s_1 \leqslant s \leqslant s_2$. Then there is a constant C such that if $\mathcal{A}$ is a bounded linear mapping from $\dot{H}^{s_1}$ into a normed linear space $\mathcal{N}$ with

$$\|\mathcal{A}v\|_{\mathcal{N}} \leqslant A \min(\|v\|_{\dot{H}_h^{s_1}}, h^{s_2}\|v\|_{\dot{H}^{s_2}}),$$

then

$$\|\mathcal{A}v\|_{\mathcal{N}} \leqslant CAh^s\|v\|_{\dot{H}^s}.$$

PROOF. – Let M be such that $k\lambda_M \leqslant 1 \leqslant k\lambda_{M+1}$. Setting

$$\tilde{v}(x) = \sum_{m=1}^{M} v_m \varphi_m(x),$$

we obtain

$$\|\mathcal{A}v\|_{\mathcal{N}} \leqslant \|\mathcal{A}\tilde{v}\|_{\mathcal{N}} + \|\mathcal{A}(v-\tilde{v})\|_{\mathcal{N}} \leqslant Ah^{s_2}\|\tilde{v}\|_{\dot{H}^{s_2}} + A\|v-\tilde{v}\|_{\dot{H}_h^{s_1}} \leqslant$$

$$\leqslant CA\Big\{\sum_{m=1}^{M}(k\lambda_m)^{s_2}v_m^2 + \sum_{m=M+1}^{\infty}(1+k\lambda_m)^{s_1}v_m^2\Big\}^{\frac{1}{2}}.$$

Since

$$(1+\tau)^{s_1} \leqslant 2^{s_1}\tau^s, \qquad \text{for } \tau \geqslant 1,$$

$$\tau^{s_2} \leqslant \tau^s, \qquad \text{for } \tau \leqslant 1,$$

it follows that

$$\|Av\|_{\mathcal{N}} \leqslant CA\Big(\sum_{m=1}^{\infty}(k\lambda_m)^s v_m^2\Big)^{\frac{1}{2}} \leqslant CAh^s\|v\|_{\dot{H}^s},$$

which proves the lemma.

We shall now introduce our assumptions on the finite dimensional spaces S_h which we will use. Let $0 \leqslant \sigma \leqslant \nu$ and let $h_0 > 0$. We say that the family $\{S_h\} = \{S_h; 0 < h \leqslant h_0\}$, of finite dimensional subspaces of $L^2(\Omega)$ belongs to $\mathcal{S}_{\sigma,\nu}$ if for each h, $S_h \subset H^\sigma$ and if there is a constant C such that for $v \in \dot{H}^\nu$,

$$\inf_{\chi\in S_h}\|v-\chi\|_{H_h^\sigma} \leqslant Ch^\nu\|v\|_{\dot{H}^\nu}. \tag{2.9}$$

The property (2.9) is shared by many of the families of piecewise polynomial spaces which have recently been employed in Galerkin or finite element investigations.

We now prove a lemma which affirms that the estimate (2.9) generalizes to large ranges of the parameters: in previous papers this stronger condition has often been made an assumption.

LEMMA 2.6. – If $\{S_h\} \in \mathcal{S}_{\sigma,\nu}$ then for $0 \leqslant \tau \leqslant \varrho \leqslant \nu$ and $\tau \leqslant \sigma$ there is a constant C such that for $v \in \dot{H}^\varrho$,

$$\inf_{\chi\in S_h}\|v-\chi\|_{H_h^\tau} \leqslant Ch^\varrho\|v\|_{\dot{H}^\varrho}.$$

PROOF. – Let $P_{\tau,h}$ be the orthogonal projection in H_h^τ onto S_h. Then by assumption

$$\|(I-P_{\tau,h})v\|_{H_h^\tau} = \inf_{\chi\in S_h}\|v-\chi\|_{H_h^\tau} \leqslant \inf_{\chi\in S_h}\|v-\chi\|_{H_h^\sigma} \leqslant Ch^\nu\|v\|_{\dot{H}^\nu}.$$

On the other hand, since $I-P_{\tau,h}$ has norm 1 in H_h^τ,

$$\|(I-P_{\tau,h})v\|_{H_h^\tau} \leqslant \|v\|_{H_h^\tau} \leqslant C\|v\|_{\dot{H}_h^\tau}.$$

The result now follows by Lemma 2.5.

Let $\dot{\mathcal{S}}_{\sigma,\nu}$ be the subclass of $\mathcal{S}_{\sigma,\nu}$ such that for $\{S_h\} \in \dot{\mathcal{S}}_{\sigma,\nu}$, $S_h \in \dot{H}^\sigma$ for all h. By Lemma 2.2 this means that for $v \in S_h$, we have $\Delta^j v = 0$ on $\partial\Omega$ for $j < \frac{1}{2}\sigma$.

We recall the following trace inequality (cf. e.g. Lemma 4.1 of [3]).

LEMMA 2.7. – There is a positive constant C such that for any $\varepsilon > 0$ and $v \in H^1$,

$$|v| \leqslant \varepsilon\|v\|_{H^1} + C\varepsilon^{-1}\|v\|.$$

A consequence which will be frequently used below is the inequality

$$h^{|\alpha|+\frac{1}{2}}|D^\alpha v| \leqslant C\|v\|_{H_h^{|\alpha|+1}}, \qquad v \in H^{|\alpha|+1}.$$

3. – Basic convergence theory for Galerkin methods.

We shall now introduce the general form of the Galerkin equations which we shall treat. Let h, k be small positive numbers and assume as before that they are tied together by the relation $kh^{-2} = \text{constant}$. Although in the sequel k is completely determined by h and conversely, we shall find it suggestive to keep both these parameters, with k denoting the time step and h the mesh width in space.

For each h, k let there be given two bilinear forms $A_k(\varphi, \psi)$, $B_k(\varphi, \psi)$ and a finite dimensional subspace S_h of $L^2(\Omega)$. We shall consider approximations $U_n(x) \in S_h$ of $u(x, nk) = E(nk)v$, at times $t = nk$, $n = 1, 2, \ldots$, defined by

$$\begin{aligned} &U_0 = v\,, \\ &A_k(U_{n+1}, \chi) = B_k(U_n, \chi)\,, \qquad \chi \in S_h\,. \end{aligned} \tag{3.1}$$

If $\{\omega_j\}_1^{N_h}$ is a basis in the finite dimensional space S_h, the problem of finding w for given v such that

$$A_k(w, \chi) = B_k(v, \chi)\,, \qquad \chi \in S_h\,, \tag{3.2}$$

can also be formulated as the problem of finding $w = \sum\limits_{j=1}^{N_h} \alpha_j \omega_j$ such that $(\alpha_1, \ldots, \alpha_{N_h})$ is the solution of the finite linear system of equations

$$\sum_{j=1}^{N_h} \alpha_j A_k(\omega_j, \omega_l) = B_k(v, \omega_l)\,, \qquad l = 1, \ldots, N_h\,. \tag{3.3}$$

We shall now make a number of assumptions about the bilinear forms and the subspaces which will make it possible to affirm that the procedure (3.1) defines a uniquely determined sequence $U_n(x)$, $n = 0, 1, 2, \ldots$. These assumptions will relate the Galerkin equations to the original mixed initial-boundary value problem and be such that $U_n(x)$ and $u(x, nk)$ may be compared.

Assumptions about A_k, B_k and S_h. There exist non-negative integers a, b, μ and ν such that

(i) $A_k(\varphi, \psi)$ is the inner product in a Hilbert space $\mathcal{H}_k$ in $L^2(\Omega)$ and containing $\dot{H}^a$ and S_h, and there exists a constant C such that for $\varphi \in \mathcal{H}_k \cap H^a$,

$$a_k(\varphi) \leqslant C \|\varphi\|_{H_h^a}\,,$$

where $a_k(\varphi) = A_k(\varphi, \varphi)^{\frac{1}{2}}$.

(ii) $B_k(\varphi, \psi)$ is defined on $(\mathcal{H}_k \cup \dot{H}^b) \times \mathcal{H}_k$ and for φ, $\psi \in \mathcal{H}_k$,

$$|B_k(\varphi, \psi)| \leqslant a_k(\varphi) a_k(\psi)\,.$$

(iii) The Galerkin equations (3.2) are satisfied by the exact solution of the differential equation with accuracy μ in the sense that for $b \leqslant s \leqslant 2\mu + 2$ there is a constant C such that for $v \in \dot{H}^b$ and $\psi \in \mathcal{H}_k$,

$$|A_k(E(k)v, \psi) - B_k(v, \psi)| \leqslant Ch^s \|v\|_{\dot{H}^s} a_k(\psi).$$

(iv) The family $\{S_h\}$ belongs to $\mathcal{S}_{a,\nu}$.

The inequality in (ii) can be considered as a stability property and (iii) expresses the consistency of the discretization in time. In applications A_k will be a differential form and the parameter a corresponds to its order. The presence of the parameter b is related to the final estimates when the initial data possess minimal regularity properties. The parameters μ and ν describe the maximum accuracy obtainable for smooth initial data.

We shall now estimate the error in the $\mathcal{H}_k$-norm. To this end we start with some simple consequences of the above assumptions.

PROPOSITION 3.1. – *The Galerkin equations* (3.2) *admit for given* $v \in \dot{H}^b \cup \mathcal{H}_k$ *a unique solution* $w = E_{kh}v \in S_h$. *For* $v \in \mathcal{H}_k$,

$$a_k(E_{kh}v) \leqslant a_k(v).$$

PROOF. – The existence and uniqueness follow at once from the fact that the matrix $(A_k(\omega_j, \omega_l))$ in (3.3) is positive definite. With $\chi = w = E_{kh}v$ we obtain by (3.2),

$$a_k(w)^2 = B_k(v, w) \leqslant a_k(v) a_k(w), \qquad v \in \mathcal{H}_k,$$

which implies the inequality.

With this notation, the approximate solution of our problem at time $t = nk$ is $U_n = E^n_{kh} v$.

PROPOSITION 3.2. – *For* $v \in \dot{H}^b \cup \mathcal{H}_k$ *the Galerkin equations over* $\mathcal{H}_k$,

$$A_k(w, \psi) = B_k(v, \psi), \qquad \psi \in \mathcal{H}_k, \tag{3.4}$$

admit a unique solution $w = E_k v \in \mathcal{H}_k$.

PROOF. – For $v \in \mathcal{H}_k$ (or $\dot{H}^b$) it follows from (ii) (or (iii)) that $B_k(v, \psi)$ is a bounded linear functional on $\mathcal{H}_k$. Hence by the Riesz representation theorem there is a $w = E_k v \in \mathcal{H}_k$ such that

$$A_k(w, \psi) = B_k(v, \psi), \qquad \psi \in \mathcal{H}_k.$$

We shall think of E_k as the exact solution operator of the time discrete problem. It is related to E_{kh} in the following way:

PROPOSITION 3.3. – Let P_h be the orthogonal projection in $\mathcal{H}_k$ onto S_h. Then $E_{kh} = P_h E_k$.

PROOF. – This follows at once by the fact that by (3.2) and (3.4),

$$A_k(E_{kh}v - E_k v, \chi) = 0 , \qquad \chi \in S_h .$$

PROPOSITION 3.4. – There is a constant C such that

$$a_k\big((I - P_h)v\big) \leqslant Ch^\nu \|v\|_{\dot{H}^\nu} , \qquad v \in \dot{H}^\nu .$$

PROOF. – By assumptions (i) and (iv) we have

$$a_k\big((I - P_h)v\big) = \inf_{\chi \in S_h} a_k(v - \chi) \leqslant C \inf_{\chi \in S_h} \|v - \chi\|_{H_h^a} \leqslant Ch^\nu \|v\|_{\dot{H}^\nu} .$$

PROPOSITION 3.5. – For $b \leqslant s \leqslant \min(2\mu + 2, \nu)$ there is a constant C such that

$$a_k\big(E_{kh}v - E(k)v\big) \leqslant Ch^s \|v\|_{\dot{H}^s} , \qquad v \in \dot{H}^s .$$

PROOF. – By the triangle inequality, using Proposition 3.3 and the fact that P_h has norm 1 in $\mathcal{H}_k$ we obtain,

$$a_k\big(E_{kh}v - E(k)v\big) \leqslant a_k\big((I - P_h)E(k)v\big) + a_k\Big(P_h\big(E(k) - E_k\big)v\Big) \leqslant$$
$$\leqslant a_k\big((I - P_h)E(k)v\big) + a_k\Big(\big(E(k) - E_k\big)v\Big) .$$

For the first term we have by Proposition 3.4 and Lemma 2.3,

$$a_k\big((I - P_h)E(k)v\big) \leqslant Ch^\nu \|E(k)v\|_{\dot{H}^\nu} \leqslant Ch^s \|v\|_{\dot{H}^s} , \qquad 0 \leqslant s \leqslant \nu ,$$

and the consistency condition (ii) implies for the second term

$$a_k\Big(\big(E(k) - E_k\big)v\Big) \leqslant Ch^s \|v\|_{\dot{H}^s} , \qquad b \leqslant s \leqslant 2\mu + 2 .$$

Hence the result follows.

We can now state and prove the basic error estimate.

THEOREM 1. – Assume that the conditions (i), (ii), (iii) and (iv) are satisfied and let $s \geqslant b$. Then there is a constant C such that with $\varrho = \min(2\mu, \nu - 2)$,

$$a_k\big(E_{kh}^n v - E(nk)v\big) \leqslant C\left(\log\frac{1}{h}\right)^{\delta_{s,\varrho}} h^{\min(s,\varrho)} \|v\|_{\dot{H}^s} , \qquad v \in \dot{H}^s ,$$

where $\delta_{s,\varrho}$ is the Kronecker delta.

PROOF. – We have using the stability of E_{kh} in $\mathcal{K}_k$ (Proposition 3.1),

$$a_k\big(E_{kh}^n v - E(nk)v\big) \leqslant \sum_{j=0}^{n-1} a_k\big(E_{kh}^{n-1-j}(E_{kh} - E(k))\,E(jk)v\big) \leqslant \sum_{j=0}^{n-1} a_k\Big(\big(E_{kh} - E(k)\big)E(jk)v\Big)\,.$$

For the term with $j = 0$ we have by Proposition 3.5,

$$a_k\Big(\big(E_{kh} - E(k)\big)v\Big) \in Ch^{\min(s,\varrho)}\|v\|_{\dot{H}^s}, \qquad s \geqslant b\,.$$

For the terms with $j > 0$ we have using Proposition 3.5 (with $s = \varrho + 2$) and Lemma 2.3,

$$\sum_{j=1}^{n-1} a_k\Big(\big(E_{kh} - E(k)\big)E(jk)v\Big) \leqslant Ckh^\varrho \sum_{j=1}^{n-1} \|E(jk)v\|_{\dot{H}^{\varrho+2}} \leqslant$$
$$\leqslant Ch^\varrho\Big\{k\sum_{j=1}^{n-1}(jk)^{-(\varrho+2-s)/2}\Big\}\|v\|_{\dot{H}^s}, \qquad 0 \leqslant s \leqslant \varrho + 2\,.$$

The result now follows since

$$k\sum_{j=1}^{n-1}(jk)^{-(\varrho+2-s)/2} \leqslant \begin{cases} C\,, & s > \varrho\,,\\ C\log n\,, & s = \varrho\,,\\ Ck^{-(\varrho-s)/2}\,, & 0 \leqslant s < \varrho\,. \end{cases}$$

4. – The stationary projection method.

The result in Section 3 is in a certain sense non-optimal with respect to the approximation; in order to obtain a ϱ-th order result we have to employ a family of subspaces in some $\mathcal{S}_{a,\nu}$ with $\nu \geqslant \varrho + 2$. This loss of $O(h^{-2}) = O(k^{-1}) = O(n)$ stems from the summation with respect to j in the proof above. We shall present below an alternative treatment which, when applicable, avoids this loss.

The main point in this more refined analysis is to add the following fifth condition in which we introduce

$$G_k(v, \chi) = A_k(v, \chi) - B_k(v, \chi)\,.$$

The condition is then the following:

(v) For given $v \in \dot{H}^\nu$ the equations

$$G_k(w - v, \chi) = 0\,, \qquad \chi \in S_h\,, \tag{4.1}$$

have a unique solution $w = Q_h v \in S_h$ and there are positive constants ν_0 and C such that the linear operator Q_h thus defined satisfies

$$\|(I-Q_h)v\|_{\mathcal{H}_k} \leqslant Ch^{\nu-\nu_0}\|v\|_{\dot{H}^\nu}.$$

Notice that, by the stability requirement, $G_k(v, v) \geqslant 0$. In the case that $G_k(\varphi, \psi)$ is symmetric and positive definite, Q_h is the orthogonal projection onto S_h in the Hilbert space $\mathcal{G}_k$ defined by the inner product $G_k(\varphi, \psi)$. In this case (v) means that the orthogonal projection with respect to $\mathcal{G}_k$ has a specific approximation property also in $\mathcal{H}_k$. In most examples below Q_h will be optimal with respect to $\mathcal{H}_k$ which will mean $\nu_0 = 0$. In one case (Section 8) we are only able to establish (v) with $\nu_0 = \frac{1}{2}$. The equation in (v) is related to the stationary version of the Galerkin equation.

In the treatment below it turns out that as we add condition (v) some of the other conditions may be relaxed in that some of the estimates only need to be valid on the subspaces S_h rather than on the whole Hilbert space $\mathcal{H}_k$. There will in fact be examples in what follows in which this is important so that the basic theory does not apply but the present does.

We now present the alternative conditions.

(i′) There is a Hilbert space $\mathcal{H}_k$ in $L^2(\Omega)$ and containing $\dot{H}^a$ and S_h such that $A_k(\varphi, \psi)$ is defined on $\mathcal{H}_k \times S_h$, is symmetric on $S_h \times S_h$, and there is a constant C such that for $\varphi \in S_h$,

$$\|\varphi\|_{\mathcal{H}_k} \leqslant Ca_k(\varphi),$$

where again $a_k(\varphi) = A_k(\varphi, \varphi)^{\frac{1}{2}}$, and such that for $\varphi \in \mathcal{H}_k \cap H^a$,

$$\|\varphi\|_{\mathcal{H}_k} \leqslant C\|\varphi\|_{H_h^a}.$$

(ii′) $B_k(\varphi, \psi)$ is defined on $(\mathcal{H}_k \cup \dot{H}^b) \times S_h$ and there is a constant C such that

$$|B_k(\varphi, \psi)| \leqslant \begin{cases} C\|\varphi\|_{\mathcal{H}_k} a_k(\psi), & \varphi \in \mathcal{H}_k,\ \psi \in S_h, \\ a_k(\varphi)a_k(\psi), & \varphi, \psi \in S_h. \end{cases}$$

(iii′) For $b \leqslant s \leqslant 2\mu + 2$ there is a constant C such that for $v \in \dot{H}^b$, $\psi \in S_h$,

$$|A_k(E(k)v, \psi) - B_k(v, \psi)| \leqslant Ch^s\|v\|_{\dot{H}^s} a_k(\psi).$$

Clearly, if (i), (ii), (iii) are satisfied then (i′), (ii′), (iii′) hold with the same $\mathcal{H}_k$. We notice that with the same proof as before we have the following:

Proposition 4.1. – Under the new assumptions, the Galerkin equations (3.2) define for given $v \in \dot{H}^b \cup \mathcal{H}_k$ a uniquely determined $w = E_{kh}v \in S_h$ where again E_{kh} is a linear operator, and for $v \in S_h$,

$$a_k(E_{kh}v) \leqslant a_k(v).$$

This inequality expresses the stability in the subspace of the discrete solution operator but will not be explicitely used in the error analysis. We notice that the exact solution operator E_k of the time discrete problem has no analogue in the new theory.

We can now state and prove the main result of this section.

THEOREM 2. – Assume that the conditions (i′), (ii′), (iii′), (iv) and (v) are satisfied and let $s \geqslant b$. Then there is a constant C such that for $v \in \dot{H}^s$,

$$\|E_{kh}^n v - E(nk)v\|_{\mathcal{H}_k} \leqslant C\left\{\left(\log\frac{1}{h}\right)^{\delta_{s,\nu}} h^{\min(s,\nu)-\nu_0} + \left(\log\frac{1}{h}\right)^{\delta_{s,2\mu}} h^{\min(s,2\mu)}\right\} \|v\|_{\dot{H}^s}$$

PROOF. – Set

$$u_n = E(nk)v\,, \qquad U_n = E_{kh}^n v \quad \text{and} \quad e_n = u_n - U_n\,.$$

We shall write e_n in the form

$$e_n = \xi_n + \eta_n\,, \qquad \text{with} \qquad \xi_n = (I - Q_h)u_n\,,$$

so that

$$\eta_n = e_n - \xi_n = Q_h u_n - U_n\,.$$

Notice that $\eta_n \in S_h$ and that ξ_n satisfies

$$G_k(\xi_n, \chi) = 0 \qquad \text{for} \qquad \chi \in S_h\,. \tag{4.2}$$

By the consistency condition (iii′) and the Galerkin equations we have for $b \leqslant s_1 \leqslant 2\mu + 2$, $\chi \in S_h$,

$$|A_k(e_{j+1}, \chi) - B_k(e_j, \chi)| \leqslant Ch^{s_1}\|u_j\|_{\dot{H}^{s_1}} a_k(\chi)\,. \tag{4.3}$$

Using (4.2) it follows in particular that

$$|A_k(\eta_{j+1}, \chi) - B_k(\eta_j, \chi) + B_k(\xi_{j+1} - \xi_j, \chi)| \leqslant Ch^{2\mu+2}\|u_j\|_{\dot{H}^{2\mu+2}} a_k(\chi)\,,$$

and hence with $\chi = \eta_{j+1}$ using the first inequality in (ii′),

$$a_k(\eta_{j+1})^2 \leqslant B_k(\eta_j, \eta_{j+1}) + C\{\|\xi_{j+1} - \xi_j\|_{\mathcal{H}_k} + kh^{2\mu}\|u_j\|_{\dot{H}^{2\mu+2}}\} a_k(\eta_{j+1})\,.$$

By the stability assumption in (ii′) and (v) this implies after cancellation of $a_k(\eta_{j+1})$,

$$\begin{aligned} a_k(\eta_{j+1}) &\leqslant a_k(\eta_j) + C\{\|(I - Q_h)(u_{j+1} - u_j)\|_{\mathcal{H}_k} + kh^{2\mu}\|u_j\|_{\dot{H}^{2\mu+2}}\} \leqslant \\ &\leqslant a_k(\eta_j) + C\{h^{\nu-\nu_0}\|(E(k) - I)\,u_j\|_{\dot{H}^\nu} + kh^{2\mu}\|u_j\|_{\dot{H}^{2\mu+2}}\}\,. \end{aligned} \tag{4.4}$$

Using Lemmas 2.4 and 2.3 we obtain for $j>0$, $0\leqslant s\leqslant \nu+2$,

$$\|(E(k)-I)\,u_j\|_{\dot H^\nu}\leqslant k\|E(jk)v\|_{\dot H^{\nu+2}}\leqslant Ck(jk)^{-(\nu+2-s)/2}\|v\|_{\dot H^s}.$$

Applying the same argument to the last term in (4.4) and summing over j we obtain for $n\geqslant 1$, $s\geqslant 0$,

$$\begin{aligned}a_k(\eta_n)&\leqslant a_k(\eta_1)+C\left\{h^{\nu-\nu_0}k\sum_{j=1}^{n-1}(jk)^{-(\nu+2-s)/2}+h^{2\mu}k\sum_{j=1}^{n-1}(jk)^{-(2\mu+2-s)/2}\right\}\|v\|_{\dot H^s}\leqslant\\&\leqslant a_k(\eta_1)+C\left\{\left(\log\frac1h\right)^{\delta_{s,\nu}}h^{\min(s,\nu)-\nu_0}+\left(\log\frac1h\right)^{\delta_{s,2\mu}}h^{\min(s,2\mu)}\right\}\|v\|_{\dot H^s}.\end{aligned}$$

To estimate $a_k(\eta_1)$ consider (4.3) with $j=0$. Since $e_0=0$ we obtain easily for $b\leqslant s_1\leqslant 2\mu+2$,

$$a_k(\eta_1)\leqslant C\{\|\xi_1\|_{\mathcal{H}_k}+h^{s_1}\|v\|_{\dot H^{s_1}}\}=C\{\|(I-Q_h)u_1\|_{\mathcal{H}_k}+h^{s_1}\|v\|_{\dot H^{s_1}}\}.$$

By assumption (v) this implies for $b\leqslant s_1\leqslant 2\mu+2$, $0\leqslant s_2\leqslant\nu$,

$$\begin{aligned}a_k(\eta_1)&\leqslant C\{h^{\nu-\nu_0}\|E(k)v\|_{\dot H^\nu}+h^{s_1}\|v\|_{\dot H^{s_1}}\}\\&\leqslant C\{h^{s_2-\nu_0}\|v\|_{\dot H^{s_2}}+h^{s_1}\|v\|_{\dot H^{s_1}}\}.\end{aligned}$$

Altogether we obtain for $s\geqslant b$,

$$\|\eta_n\|_{\mathcal{H}_k}\leqslant Ca_k(\eta_n)\leqslant C\left\{\left(\log\frac1h\right)^{\delta_{s,\nu}}h^{\min(s,\nu)-\nu_0}+\left(\log\frac1h\right)^{\delta_{s,2\mu}}h^{\min(s,2\mu)}\right\}\|v\|_{\dot H^s}.$$

On the other hand, by (v) we have for $n\geqslant 1$, $0\leqslant s\leqslant\nu$,

$$\begin{aligned}\|\xi_n\|_{\mathcal{H}_k}&=\|(I-Q_h)u_n\|_{\mathcal{H}_k}\leqslant Ch^{\nu-\nu_0}\|E(nk)v\|_{\dot H^\nu}\leqslant\\&\leqslant Ch^{\nu-\nu_0}(nk)^{-(\nu-s)/2}\|v\|_{\dot H^s}\leqslant Ch^{s-\nu_0}\|v\|_{\dot H^s},\end{aligned}$$

and hence finally for $s\geqslant b$, $n\geqslant 1$,

$$\|e_n\|_{\mathcal{H}_k}\leqslant\|\xi_n\|_{\mathcal{H}_k}+\|\eta_n\|_{\mathcal{H}_k}\leqslant C\left\{\left(\log\frac1h\right)^{\delta_{s,\nu}}h^{\min(s,\nu)-\nu_0}+\left(\log\frac1h\right)^{\delta_{s,2\mu}}h^{\min(s,2\mu)}\right\}\|v\|_{\dot H^s},$$

which completes the proof.

5. – Some consistency estimates.

We shall prove here a lemma which will be used in establishing the consistency estimates in all the examples below.

LEMMA 5.1. – Let $r(\tau) = b(\tau)/a(\tau)$ be a rational function with

$$a(\tau) = \sum_{j=0}^{\alpha} a_j \tau^j, \qquad b(\tau) = \sum_{j=0}^{\beta} b_j \tau^j, \qquad a_\alpha \neq 0,\ b_\beta \neq 0,\ a_0 = b_0 = 1,$$

such that

$$\text{(i)} \qquad a(\tau) > 0, \qquad |r(\tau)| \leqslant 1 \quad \text{for} \quad \tau > 0,$$

and for some $\mu \geqslant 1$ with $\mu + 1 \geqslant \beta$,

$$\text{(ii)} \qquad r(\tau) = \exp[-\tau] + O(\tau^{\mu+1}) \quad \text{as} \quad \tau \to 0.$$

Then for $v, w \in \dot{H}^\infty$,

$$\|a(-k\Delta)E(k)v - b(-k\Delta)v\| \leqslant Ch^s \|v\|_{\dot{H}^s}, \qquad 2\beta \leqslant s \leqslant 2\mu + 2, \tag{5.1}$$

$$\big|\big(a(-k\Delta)E(k)v - b(-k\Delta)v, w\big)\big| \leqslant Ch^s \|v\|_{\dot{H}^s}\big(a(-k\Delta)w, w\big)^{\frac12}, \tag{5.2}$$
$$\max(2\beta - \alpha, 0) \leqslant s \leqslant 2\mu + 2.$$

PROOF. – We have, with the notation of Section 2,

$$\|a(-k\Delta)E(k)v - b(-k\Delta)v\|^2 = \sum_m \big(a(k\lambda_m)\exp[-k\lambda_m] - b(k\lambda_m)\big)^2 v_m^2,$$

$$\big(a(-k\Delta)E(k)v - b(-k\Delta)v, w\big) = \sum_m \big(a(k\lambda_m)\exp[-k\lambda_m] - b(k\lambda_m)\big) v_m w_m,$$

and

$$\big(a(-k\Delta)w, w\big) = \sum_m a(k\lambda_m) w_m^2.$$

The first result therefore follows from the inequality

$$|a(\tau)e^{-\tau} - b(\tau)| \leqslant |a(\tau)\,(e^{-\tau} - r(\tau))| \leqslant C\tau^\sigma, \qquad \tau > 0,\ \beta \leqslant \sigma \leqslant \mu + 1,$$

since hence with $\sigma = s/2$,

$$\|a(-k\Delta)E(k)v - b(-k\Delta)v\| \leqslant C\Big(\sum_m (k\lambda_m)^s v_m^2\Big)^{\frac12} \leqslant Ch^s \|v\|_{\dot{H}^s}.$$

The second result follows similarly from

$$|a(\tau)^{\frac{1}{2}}(e^{-\tau}-r(\tau))|\leqslant C\tau^{\sigma}\,, \qquad \max(\beta-\tfrac{1}{2}\alpha,\,0)\leqslant\sigma\leqslant\mu+1\,.$$

A special set of rational functions satisfying the assumptions of Lemma 5.1 is formed by the diagonal and subdiagonal Padé approximations of $e^{-\tau}$ (cf. [8]). These are defined by

$$r_{\alpha\beta}(\tau)=\frac{b_{\alpha\beta}(\tau)}{a_{\alpha\beta}(\tau)}\,,$$

where

$$a_{\alpha\beta}(\tau)=\sum_{j=0}^{\alpha}\frac{(\alpha+\beta-j)!\,\alpha!}{(\alpha+\beta)!\,j!\,(\alpha-j)!}\,\tau^{j}=\sum_{j=0}^{\alpha}a_{\alpha\beta j}\,\tau^{j}\,,$$

$$b_{\alpha\beta}(\tau)=\sum_{j=0}^{\beta}\frac{(\alpha+\beta-j)!\,\beta!}{(\alpha+\beta)!\,j!\,(\beta-j)!}\,(-\tau)^{j}=\sum_{j=0}^{\beta}b_{\alpha\beta j}\,\tau^{j}\,.$$

The Padé approximations are the most accurate approximations of $e^{-\tau}$ near the origin with given degrees of a and b. The assumptions of Lemma 5.1 are satisfied for $\beta\leqslant\alpha$ with $\mu=\alpha+\beta$. Further,

$$|b_{\alpha\beta j}|<a_{\alpha\beta j}\,, \qquad 1\leqslant j\leqslant\alpha\,,\ \beta<\alpha\,, \tag{5.3}$$

and

$$b_{\alpha\alpha j}=(-1)^{j}a_{\alpha\alpha j}\neq 0\,, \qquad 0\leqslant j\leqslant\alpha\,,\ \beta=\alpha\,. \tag{5.4}$$

6. – Some methods with subspaces satisfying prescribed boundary conditions.

We shall introduce here a class of methods which illustrates the results in Sections 3 and 4. The class will contain methods of arbitrary order of accuracy but the high order of accuracy will be achieved in the examples of this section only for subspaces which satisfy quite restrictive boundary conditions.

We begin with a particular case, namely the methods analyzed by DOUGLAS and DUPONT [4]. Set

$$A_k(\varphi,\psi)=(\varphi,\psi)+\varkappa kD(\varphi,\psi)\,,$$

$$B_k(\varphi,\psi)=(\varphi,\psi)-(1-\varkappa)kD(\varphi,\psi)\,,$$

where $\varkappa\geqslant\frac{1}{2}$ and let $\mathcal{K}_k$ be the completion of $C_0^\infty(\Omega)$ with respect to the norm $a_k(\varphi)=A_k(\varphi,\varphi)^{\frac{1}{2}}$. Obviously

$$c\|\varphi\|_{\dot H_h^1}\leqslant a_k(\varphi)\leqslant C\|\varphi\|_{\dot H_h^1}\,.$$

The Hilbert space $\mathcal{H}_k$ contains the same functions as $\dot{H}^1$ and we therefore assume $\{S_h\} \in \dot{\mathrm{S}}_{1,\nu}$ so that in particular the elements of S_h vanish on $\partial\Omega$. Hence conditions (i) and (iv) are satisfied with $a = 1$. By Cauchy's inequality we have, since $\varkappa \geq \frac{1}{2}$,

$$|B_k(\varphi, \psi)| \leq \|\varphi\| \cdot \|\psi\| + \varkappa k D(\varphi, \varphi)^{\frac{1}{2}} D(\psi, \psi)^{\frac{1}{2}} \leq a_k(\varphi) a_k(\psi),$$

so that (ii) holds with $b = 1$ for $\varkappa \neq 1$, and $b = 0$ for $\varkappa = 1$.

To see that Theorem 1 applies it remains only to consider the consistency condition (iii). For $v, \psi \in \dot{H}^\infty$ we have by integration by parts, since the functions vanish on $\partial\Omega$,

$$\begin{aligned} A_k\big(E(k)v, \psi\big) - B_k(v, \psi) &= \\ &= \big((E(k) - I)v, \psi\big) + kD\big((\varkappa E(k) + (1-\varkappa)I)v, \psi\big) = \\ &= \big((E(k) - I)v, \psi\big) - k\big(\Delta(\varkappa E(k) + (1-\varkappa)I)v, \psi\big) = \\ &= \big((I - \varkappa k\Delta)E(k)v - (I + (1-\varkappa)k\Delta)v, \psi\big). \end{aligned}$$

Hence applying Lemma 5.1 with

$$r(\tau) = \frac{1 - (1-\varkappa)\tau}{1 + \varkappa\tau}, \tag{6.1}$$

we obtain

$$|A_k\big(E(k)v, \psi\big) - B_k(v, \psi)| \leq Ch^s \|v\|_{\dot{H}^s} a_k(\psi), \qquad b \leq s \leq 2\mu + 2,$$

where $\mu = 2$ if $\varkappa = \frac{1}{2}$ and $\mu = 1$ otherwise. In particular for $\varkappa = \frac{1}{2}$ and $\nu = 6$, Theorem 1 gives

$$\|E_{kh}^n v - E(nk)v\|_{\dot{H}_h^1} \leq \begin{cases} Ch^4 \|v\|_{\dot{H}^s}, & s > 4, \\ Ch^4 \log \dfrac{1}{h} \|v\|_{\dot{H}^4}, & \\ Ch^s \|v\|_{\dot{H}^s}, & 1 \leq s < 4. \end{cases}$$

In order to apply Theorem 2 we choose the same $\mathcal{H}_k$ as above. The conditions (i′), (ii′), (iii′), (iv) are then satisfied as before with the appropriate parameters. We now turn to condition (v). The bilinear form $G_k(\varphi, \psi) = kD(\varphi, \psi)$ is here positive definite so that the equations (4.1) have a unique solution $w = Q_h v \in S_h$. We have with $g_k(\varphi) = G_k(\varphi, \varphi)^{\frac{1}{2}}$,

$$g_k\big((I - Q_h)v\big) \leq C \inf_{\chi \in S_h} \|v - \chi\|_{\dot{H}_h^1} \leq Ch^s \|v\|_{\dot{H}^s}, \qquad 1 \leq s \leq \nu. \tag{6.2}$$

In order to estimate $\tilde{v} = (I - Q_h)v$ in $\mathcal{H}_k$ it remains to estimate $\tilde{v}$ in $L^2(\Omega)$. For this purpose we use a technique due to NITSCHE [5]. Let $w \in \dot{H}^2$ be defined as the solution of the Dirichlet problem

$$-\Delta w = \tilde{v} \text{ in } \Omega, \quad w = 0 \text{ on } \partial\Omega. \tag{6.3}$$

Then

$$\|\tilde{v}\|^2 = -(\tilde{v}, \Delta w) = D(\tilde{v}, w).$$

By the definition of $\tilde{v}$ we have

$$D(\tilde{v}, \chi) = k^{-1}G_k(\tilde{v}, \chi) = 0, \quad \chi \in S_h,$$

so that with $\chi = Q_h w$ and $\tilde{w} = (I - Q_h)w$,

$$\|\tilde{v}\|^2 = D(\tilde{v}, (I - Q_h)w) \leqslant Ch^{-2} g_k(\tilde{v}) g_k(\tilde{w}). \tag{6.4}$$

Now by (6.2) and a standard a priori estimate for the solution of (6.3),

$$g_k(\tilde{w}) \leqslant Ch^2 \|w\|_{\dot{H}^2} \leqslant Ch^2 \|\tilde{v}\|.$$

Hence by (6.4),

$$\|\tilde{v}\| \leqslant C g_k(\tilde{v}) \leqslant Ch^\nu \|v\|_{\dot{H}^\nu}.$$

This proves that (v) is satisfied with $\nu_0 = 0$.

In particular, Theorem 2 with $\varkappa = \frac{1}{2}$, $\nu = 4$ now gives

$$\|E_{kh}^n v - E(nk)v\|_{\dot{H}_h^1} \leqslant \begin{cases} Ch^4 \|v\|_{\dot{H}^s}, & s > 4, \\ Ch^4 \log \dfrac{1}{h} \|v\|_{\dot{H}^4}, & \\ Ch^s \|v\|_{\dot{H}^s}, & 1 \leqslant s < 4. \end{cases}$$

Notice the reduction from $\nu = 6$ to $\nu = 4$ in the approximability assumptions for the subspaces. Hence in this case Theorem 2 is stronger than Theorem 1.

The preceding example, as mentioned above, corresponds to the rational function (6.1). For the purpose of including examples of higher order accuracy we shall now construct Galerkin methods based on more general rational functions.

To this end let $r(\tau) = b(\tau)/a(\tau)$ be a rational function satisfying the assumptions of Lemma 5.1. Define for $v, w \in \dot{H}^\infty$,

$$A_k(v, w) = \big(a(-k\Delta)v, w\big) = \sum_{j=0}^{\alpha} a_j(-k)^j (\Delta^j v, w).$$

By integrating by parts we find because of the boundary conditions

$$(\Delta^{2j} v, w) = (\Delta^j v, \Delta^j w), \quad -(\Delta^{2j+1} v, w) = D(\Delta^j v, \Delta^j w),$$

so that A_k is symmetric. In terms of the coefficients of the eigenfunction expansions of v and w we have

$$A_k(v, w) = \sum_m a(k\lambda_m) v_m w_m ,$$

and it follows that $a_k(\varphi)$ is a norm equivalent (uniformly in k) to that in $\dot{H}_h^\alpha$. We now choose $\mathcal{H}_k$ to be the completion of $\dot{H}^\infty$ with respect to $a_k(\varphi)$. Assuming $\{S_h\} \in \dot{\mathcal{S}}_{\alpha,\nu}$, the conditions (i) and (iv) are then satisfied with $a = \alpha$.

Similarly, for $v, w \in \dot{H}^\infty$ we define

$$B_k(v, w) = \big(b(-k\Delta)v, w\big) = \sum_m b(k\lambda_m) v_m w_m ,$$

and make the obvious extension to $\dot{H}^b \times \dot{H}^\alpha$ with $b = \max(2\beta - \alpha, 0)$. It follows that (ii) holds with this b since by the assumption (i) of Lemma 5.1,

$$|B_k(v, w)| = \Big|\sum_m b(k\lambda_m) v_m w_m\Big| = \Big|\sum_m r(k\lambda_m)\big(a(k\lambda_m)^{\frac12} v_m\big)\big(a(k\lambda_m)^{\frac12} w_m\big)\Big| \leqslant$$
$$\leqslant \Big(\sum_m a(k\lambda_m) v_m^2\Big)^{\frac12} \Big(\sum_m a(k\lambda_m) w_m^2\Big)^{\frac12} = a_k(v)\, a_k(w) .$$

The consistency requirement (iii) is also satisfied as expressed by (5.2) of Lemma 5.1 and hence we may apply Theorem 1 with the appropriate choice of parameters.

The present Galerkin method can be thought of as consisting of solving approximately at each time step an elliptic problem of the form

$$a(-k\Delta) w = b(-k\Delta) v \ \text{ in } \Omega ,$$
$$\Delta^j w = 0 \ \text{ for } j < \alpha \text{ on } \partial\Omega ;$$

the exact solution operator of this problem defines the operator E_k appearing in the theory in Section 3.

We shall now consider the application of Theorem 2. We have already treated the case $\alpha = 1$ above so that we shall assume below that $\alpha \geqslant 2$. Choosing $\mathcal{H}_k$ and $\{S_h\}$ as above, the assumptions (i′), (ii′), (iii′) and (iv) are again valid with the same parameters as before. We now turn to condition (v). Set

$$g(\tau) = a(\tau) - b(\tau) = \sum_{j=1}^{\alpha} g_j \tau^j ,$$

where the degree of g is at most α and where $g_1 = 1$ by Lemma 5.1, (i), (ii). We shall first make the following additional assumption, namely

$$g(\tau) > 0 \quad \text{for} \quad \tau > 0 , \quad \text{and } g_\alpha \neq 0 . \tag{6.5}$$

By this assumption,

$$\tau^{\alpha} a(\tau) \leqslant C g(\tau)^2 , \qquad \tau > 0 .$$

By (6.5) the bilinear form

$$G_k(v, w) = \big(g(-k\Delta)v, w\big) = \sum_m g(k\lambda_m) v_m w_m$$

is positive definite on $\dot{H}^{\alpha}$. Let Q_h be the projection onto S_h defined by the inner product $G_k(v, w)$. In order to prove condition (v) with $\nu_0 = 0$ we want to prove that

$$a_k\big((I - Q_h)v\big) \leqslant C h^{\nu} \|v\|_{\dot{H}^{\nu}} . \tag{6.6}$$

Notice that with $(I - Q_h)v = \tilde{v}$ we have by definition

$$G_k(\tilde{v}, \chi) = 0 , \qquad \chi \in S_h , \tag{6.7}$$

$$g_k(\tilde{v}) = G_k(\tilde{v}, \tilde{v})^{\frac{1}{2}} \leqslant C \inf_{\chi \in S_h} \|v - \chi\|_{\dot{H}_h^{\alpha}} \leqslant C h^s \|v\|_{\dot{H}^s} , \qquad \alpha \leqslant s \leqslant \nu . \tag{6.8}$$

Define ψ by

$$\psi_m = (\psi, \varphi_m) = \frac{a(k\lambda_m)}{g(k\lambda_m)} \tilde{v}_m .$$

Since a/g is bounded at infinity, $\psi \in \dot{H}^s$ when $\tilde{v} \in \dot{H}^s$. Setting $\tilde{\psi} = (I - Q_h)\psi$ we obtain by (6.7) and Cauchy's inequality

$$a_k(\tilde{v})^2 = G_k(\tilde{v}, \psi) = G_k(\tilde{v}, \tilde{\psi}) \leqslant g_k(\tilde{v}) g_k(\tilde{\psi}) . \tag{6.9}$$

Using (6.8) with $s = \alpha$ we have by our assumption (6.5),

$$g_k(\tilde{\psi}) \leqslant C h^{\alpha} \|\psi\|_{\dot{H}^{\alpha}} \leqslant C \left(\sum_m (k\lambda_m)^{\alpha} \left(\frac{a(k\lambda_m)}{g(k\lambda_m)} \right)^2 \tilde{v}_m^2 \right)^{\frac{1}{2}} \leqslant C a_k(\tilde{v}) ,$$

so that by (6.9) and (6.8) with $s = \nu$,

$$a_k(\tilde{v}) \leqslant C g_k(\tilde{v}) \leqslant C h^{\nu} \|v\|_{\dot{H}^{\nu}} ,$$

which completes the proof of (6.6).

We shall now show that under an additional assumption on $\{S_h\}$ we can relax the condition $\gamma = \operatorname{degree} g = \operatorname{degree} a = \alpha$. We now only assume

$$g(\tau) > 0 \quad \text{for} \quad \tau > 0 ,$$

and make the « inverse » assumption

(6.10) $$\|\chi\|_{\dot H_h^\alpha} \leqslant C\|\chi\|_{\dot H_h^\gamma}, \quad \chi \in S_h .$$

We clearly have

$$(1+\tau)^\gamma \leqslant C(1+g(\tau)),$$

and it follows that

$$\|v\|_{\dot H_h^\gamma} \leqslant C(\|v\| + g_k(v)), \quad v \in \dot H^\gamma,$$

and hence by the inverse assumption (6.10),

(6.11) $$\|\chi\|_{\dot H_h^\alpha} \leqslant C(\|\chi\| + g_k(\chi)), \quad \chi \in S_h .$$

Let Q_h be, as before, the projection with respect to the inner product $G_k(v, w)$ and notice that by Lemma 2.6,

(6.12) $$g_k((I-Q_h)v) \leqslant Ch^s\|v\|_{\dot H^s}, \quad \gamma \leqslant s \leqslant \nu .$$

We shall first prove

$$\|(I-Q_h)v\| \leqslant Ch^\nu\|v\|_{\dot H^\nu} .$$

In fact, let $\tilde v$, ψ and $\tilde\psi$ be defined by

$$\tilde v = (I-Q_h)v, \qquad \psi_m = \frac{\tilde v_m}{g(k\lambda_m)}, \qquad \tilde\psi = (I-Q_h)\psi .$$

We have

(6.13) $$\|\tilde v\|^2 = G_k(\tilde v, \psi) = G_k(\tilde v, \tilde\psi) \leqslant g_k(\tilde v) g_k(\tilde\psi),$$

and by (6.12) with $s=\gamma$,

$$g(\tilde\psi) \leqslant Ch^\gamma\|\psi\|_{\dot H^\gamma} \leqslant C\left(\sum_m \frac{(k\lambda_m)^\gamma}{g(k\lambda_m)^2}\tilde v_m^2\right)^{\frac12} \leqslant C\|\tilde v\| .$$

It follows from (6.13) and (6.12) that

(6.14) $$\|\tilde v\| \leqslant Cg_k(\tilde v) \leqslant Ch^\nu\|v\|_{\dot H^\nu} .$$

Notice that this did not require any inverse assumption.

We have for $\chi \in S_h$,

$$a_k((I-Q_h)v) \leqslant a_k(v-\chi) + a_k(\chi - Q_h v),$$

and by (6.11),

$$a_k(\chi - Q_h v) \leq C\|\chi - Q_h v\|_{\dot H_h^\alpha} \leq C\{\|\chi - Q_h v\| + g_k(\chi - Q_h v)\} \leq$$
$$\leq C\{\|(I - Q_h)v\| + g_k((I - Q_h)v) + \|v - \chi\|_{\dot H_h^\nu}\},$$

so that

$$a_k(\tilde v) \leq C\Big\{\inf_{\chi\in S_h} \|v - \chi\|_{\dot H_h^\alpha} + \|\tilde v\| + g_k(\tilde v)\Big\} \leq Ch^\nu \|v\|_{\dot H^\nu},$$

from (iv) and (6.14). This proves that (v) holds.

The diagonal and subdiagonal Padé approximations are special cases of rational functions satisfying the assumptions of Lemma 5.1 and therefore Theorem 1 applies to all of these. In particular, this means that the class of methods so characterized contains methods of arbitrarily high order of accuracy. The following two Padé approximations correspond to $\mu = 3$ and $\mu = 4$, respectively:

$$r_{21}(\tau) = \frac{1 - (1/3)\,\tau}{1 + (2/3)\,\tau + (1/6)\,\tau^2}, \quad r_{22}(\tau) = \frac{1 - (1/2)\,\tau + (1/12)\,\tau^2}{1 + (1/2)\,\tau + (1/12)\,\tau^2}.$$

The application of Theorem 2 is again possible by (5.3), (5.4) and (6.5) for all subdiagonal and odd order diagonal Padé approximations. With an inverse assumption also the even order diagonal Padé approximations are included in the above theory. In the special case r_{22} the inverse assumption (6.10) takes the form

$$\|\chi\|_{\dot H_h^2} \leq C\|\chi\|_{\dot H_h^1}, \qquad \chi \in S_h.$$

7. – Some methods using subspaces without prescribed boundary conditions.

We shall describe here some methods which are based on work in the elliptic case by NITSCHE [6] and BRAMBLE and NITSCHE [1]. In the first method we shall present a second order scheme in time in which we require the family of subspaces to satisfy inverse assumptions. For this method only Theorem 2 applies since A_k will be definite only on the subspace. A similar fourth order method is then introduced. Finally, a second order method is described which does not require inverse and boundary condition assumptions.

For $\varphi, \psi \in H^2$ and γ positive we define

$$N_\gamma(\varphi, \psi) = D(\varphi, \psi) - \Big\langle \varphi, \frac{\partial\psi}{\partial n}\Big\rangle - \Big\langle \frac{\partial\varphi}{\partial n}, \psi\Big\rangle + \gamma h^{-1}\langle\varphi, \psi\rangle.$$

We shall consider a family of subspaces $\{S_h\} \in \mathcal{S}_{2,\nu}$ for which the following inverse assumption holds: There is a constant C_0 independent of h such that

$$\left|\frac{\partial\chi}{\partial n}\right| \leq C_0 h^{-\frac12} D(\chi, \chi)^{\frac12}, \qquad \chi \in S_h. \tag{7.1}$$

Setting

$$d_h(\varphi) = \left(D(\varphi, \varphi) + h\left|\frac{\partial\varphi}{\partial n}\right|^2 + h^{-1}|\varphi|^2\right)^{\frac{1}{2}},$$

we have the following:

LEMMA 7.1. – Under the assumption (7.1) there is a constant γ_0 such that for $\gamma \geqslant \gamma_0$, N_γ is positive definite on S_h; more precisely, for fixed $\gamma \geqslant \gamma_0$, there are constants c and C with

$$cd_h(\chi) \leqslant N_\gamma(\chi, \chi)^{\frac{1}{2}} \leqslant Cd_h(\chi), \qquad \chi \in S_h.$$

PROOF. – The inequality on the left follows from (7.1) and that on the right from Cauchy's inequality (cf. NITSCHE [6]).

We shall consider the Galerkin equations

$$(U_{n+1} - U_n, \chi) + \frac{k}{2} N_\gamma(U_{n+1} + U_n, \chi) = 0, \qquad \chi \in S_h;$$

that is, we take

$$A_k(\varphi, \psi) = (\varphi, \psi) + \frac{k}{2} N_\gamma(\varphi, \psi),$$

$$B_k(\varphi, \psi) = (\varphi, \psi) - \frac{k}{2} N_\gamma(\varphi, \psi).$$

We shall see that in this case Theorem 2 applies. By Lemma 7.1, A_k is positive definite on S_h and

$$|B_k(\varphi, \psi)| \leqslant a_k(\varphi) a_k(\psi), \qquad \varphi, \psi \in S_h.$$

Let $\mathcal{H}_k$ be the Hilbert space obtained by completing $C^\infty(\bar{\Omega})$ with respect to the norm

$$\|\varphi\|_{\mathcal{H}_k} = (\|\varphi\|^2 + h^2 d_h(\varphi)^2)^{\frac{1}{2}}.$$

As an immediate consequence of Lemma 7.1,

$$c\|\chi\|_{\mathcal{H}_k} \leqslant a_k(\chi) \leqslant C\|\chi\|_{\mathcal{H}_k}, \qquad \chi \in S_h.$$

One easily proves by Lemma 2.7 that

$$\|\varphi\|_{\mathcal{H}_k} \leqslant C\|\varphi\|_{H_h^2}.$$

In particular, $\mathcal{H}_k$ contains H^2 and hence also S_h. For $\varphi, \psi \in \mathcal{H}_k$ one has from the definitions,

$$|B_k(\varphi, \psi)| \leqslant C\|\varphi\|_{\mathcal{H}_k}\|\psi\|_{\mathcal{H}_k},$$

and hence (i′), (ii′), (iv) are all satisfied with $a=b=2$. Consider now the consistency condition (iii′). We have after integration by parts and using the boundary conditions, for $v\in\dot H^\infty$, $\psi\in S_h$,

$$N_\gamma(v,\psi)=-(\Delta v,\psi)\,,$$

and hence for such functions v and ψ,

$$|A_k(E(k)v,\psi)-B_k(v,\psi)|=\left|\left((E(k)-I)v-\frac{k}{2}(E(k)+I)\Delta v,\psi\right)\right|\leqslant$$
$$\leqslant\left\|\left(I-\frac{k}{2}\Delta\right)E(k)v-\left(I+\frac{k}{2}\Delta\right)v\right\|\,\|\psi\|\leqslant Ch^s\|v\|_{\dot H^s}\|\psi\|\,,\qquad 2\leqslant s\leqslant 6\,,$$

with the last estimate following from Lemma 5.1. That (iii′) is satisfied with $b=\mu=2$ now follows since N_γ is positive definite on S_h and hence

$$\|\psi\|\leqslant\left(\|\psi\|^2+\frac{k}{2}N_\gamma(\psi,\psi)\right)^{\frac12}=a_k(\psi)\,,\qquad \psi\in S_h\,.$$

We finally turn to the approximation property (v). We have the following result of NITSCHE [6]:

LEMMA 7.2. – For $v\in\dot H^\nu$ given, the equations

(7.2) $$N_\gamma(v-w,\chi)=0\,,\qquad \chi\in S_h\,,$$

admit a unique solution $w=Q_hv\in S_h$ and

$$\|(I-Q_h)v\|_{\mathcal{H}_k}\leqslant Ch^\nu\|v\|_{\dot H^\nu}\,.$$

In this case we have

$$G_k(\varphi,\psi)=kN_\gamma(\varphi,\psi)\,,$$

so that the equation in condition (v) is exactly (7.2). Hence the conclusion of Lemma 7.2 implies that (v) holds with $\nu_0=0$ and hence Theorem 2 applies. For $s>4$, $\nu=4$ the result is

$$\|E_{kh}^n v-E(nk)v\|_{\mathcal{H}_k}\leqslant Ch^4\|v\|_{\dot H^s}\,.$$

For the purpose of describing also a fourth order method similar to the second order method just introduced we shall employ in addition to the bilinear form $N_\gamma(\varphi,\psi)$ also

$$M_\gamma(\varphi,\psi)=(\Delta\varphi,\Delta\psi)+\left\langle\varphi,\frac{\partial\Delta\psi}{\partial n}\right\rangle+\left\langle\frac{\partial\Delta\varphi}{\partial n},\psi\right\rangle+\gamma h^{-3}\langle\varphi,\psi\rangle\,.$$

We shall assume this time that $\{S_h\} \in \mathrm{S}_{4,\nu}$ with $\nu \geq 4$ and that the following inverse assumptions hold, namely in addition to (7.1),

$$\left|\frac{\partial \Delta \chi}{\partial n}\right| \leq C_0 h^{-\frac{3}{2}} \|\Delta \chi\|, \qquad \chi \in S_h, \tag{7.3}$$

and

$$\|\Delta \chi\| \leq C_0 h^{-1} D(\chi, \chi)^{\frac{1}{2}}, \qquad \chi \in S_h. \tag{7.4}$$

We have the following lemma:

LEMMA 7.3. – Under the assumption (7.3) there is a γ_0 such that for $\gamma \geq \gamma_0$, M_γ is positive definite on S_h; more precisely, for each $\gamma \geq \gamma_0$ there are positive constants c and C with

$$c\left(\|\Delta \chi\| + h^{-\frac{3}{2}}|\chi| + h^{\frac{3}{2}}\left|\frac{\partial \Delta \chi}{\partial n}\right|\right) \leq M_\gamma(\chi, \chi)^{\frac{1}{2}} \leq C\left(\|\Delta \chi\| + h^{-\frac{3}{2}}|\chi| + h^{\frac{3}{2}}\left|\frac{\partial \Delta \chi}{\partial n}\right|\right), \quad \chi \in S_h.$$

PROOF. – We have at once by the inverse assumption (7.3) with $\varepsilon = \varepsilon_0 h^3$ and ε_0 small enough, for $\chi \in S_h$,

$$M_\gamma(\chi, \chi) = \|\Delta \chi\|^2 + 2\left\langle \chi, \frac{\partial \Delta \chi}{\partial n}\right\rangle + \gamma h^{-3}|\chi|^2 \geq$$
$$\geq \|\Delta \chi\|^2 - \varepsilon\left|\frac{\partial \Delta \chi}{\partial n}\right|^2 + (\gamma h^{-3} - \varepsilon^{-1})|\chi|^2 \geq \frac{1}{2}\|\Delta \chi\|^2 + (\gamma - \varepsilon_0^{-1}) h^{-3}|\chi|^2.$$

Using the assumption (7.3) once more, this proves the lemma.

Now let γ be large enough that both N_γ and M_γ are positive definite on S_h as in Lemmas 7.1 and 7.3. Consider the Galerkin equations defined by

$$A_k(\varphi, \psi) = (\varphi, \psi) + \frac{k}{2} N_\gamma(\varphi, \psi) + \frac{k^2}{12} M_\gamma(\varphi, \psi),$$
$$B_k(\varphi, \psi) = (\varphi, \psi) - \frac{k}{2} N_\gamma(\varphi, \psi) + \frac{k^2}{12} M_\gamma(\varphi, \psi).$$

Clearly by Lemmas 7.1 and 7.3, A_k is symmetric positive definite on S_h and

$$|B_k(\varphi, \psi)| \leq a_k(\varphi) a_k(\psi), \qquad \varphi, \psi \in S_h.$$

Let now $\mathcal{H}_k$ be the Hilbert space defined by completion of $C^\infty(\bar{\Omega})$ with respect to

$$\|\varphi\|_{\mathcal{H}_k} = \left(\|\varphi\|^2 + h^2 D(\varphi, \varphi) + h^4 \|\Delta \varphi\|^2 + h^3\left|\frac{\partial \varphi}{\partial n}\right|^2 + h^7\left|\frac{\partial \Delta \varphi}{\partial n}\right|^2\right)^{\frac{1}{2}}.$$

It then follows by Lemmas 7.1 and 7.3 and obvious estimates that (i') and (ii') are valid with $a = b = 4$.

We now turn to consistency. We have for $v \in \dot{H}^\infty$, $\psi \in S_h$,

$$N_\gamma(v, \psi) = -(\Delta v, \psi), \qquad M_\gamma(v, \psi) = (\Delta^2 v, \psi),$$

and hence in the same way as above, by (5.1) of Lemma 5.1,

$$|A_k(E(k)v, \psi) - B_k(v, \psi)| =$$
$$= \left|\left(\left(I - \frac{1}{2}k\Delta + \frac{1}{12}k^2\Delta^2\right)E(k)v - \left(I + \frac{1}{2}k\Delta + \frac{1}{12}k^2\Delta^2\right)v, \psi\right)\right| \leqslant$$
$$\leqslant Ch^s\|v\|_{\dot{H}^s}\|\psi\| \leqslant Ch^s\|v\|_{\dot{H}^s}a_k(\psi), \qquad 4 \leqslant s \leqslant 10,$$

which is (iii') with $b = \mu = 4$.

For the purpose of applying Theorem 2 it remains only to prove the approximation property (v). We have again this time

$$G_k(v, \chi) = kN_\gamma(v, \chi),$$

and the result therefore follows with $\nu_0 = 0$ from the following:

LEMMA 7.4. – There is a constant C such that for $v \in \dot{H}^\nu$ the equations

$$N_\gamma(w - v, \chi) = 0, \qquad \chi \in S_h,$$

admit a unique solution $w = Q_h v \in S_h$ and

$$\|(I - Q_h)v\|_{\mathcal{H}_k} \leqslant Ch^\nu\|v\|_{\dot{H}^\nu}.$$

PROOF. – By Lemma 7.2 it remains only to prove that with $\tilde{v} = (I - Q_h)v$,

$$h^2\|\Delta\tilde{v}\| + h^{\frac{7}{2}}\left|\frac{\partial\Delta\tilde{v}}{\partial n}\right| \leqslant Ch^\nu\|v\|_{\dot{H}^\nu}.$$

We have for arbitrary $\chi \in S_h$,

$$h^2\|\Delta\tilde{v}\| + h^{\frac{7}{2}}\left|\frac{\partial\Delta\tilde{v}}{\partial n}\right| \leqslant$$
$$\leqslant h^2\|\Delta(v - \chi)\| + h^{\frac{7}{2}}\left|\frac{\partial\Delta(v - \chi)}{\partial n}\right| + h^2\|\Delta(\chi - Q_h v)\| + h^{\frac{7}{2}}\left|\frac{\partial\Delta(\chi - Q_h v)}{\partial n}\right|.$$

Using the inverse assumptions (7.3) and (7.4), the last two terms are majorized by

$$ChD(\chi - Q_h v, \chi - Q_h v)^{\frac{1}{2}} \leqslant Ch[D(\chi - v, \chi - v)^{\frac{1}{2}} + D(\tilde{v}, \tilde{v})^{\frac{1}{2}}].$$

Hence

$$h^2\|\Delta\tilde{v}\| + h^{\frac{3}{2}}\left|\frac{\partial\Delta\tilde{v}}{\partial n}\right| \leq C\Big[\inf_{\chi\in S_h}\|v-\chi\|_{H_h^4} + \|(I-Q_h)v\|_{H_h^1}\Big] \leq Ch^{\nu}\|v\|_{\dot{H}^{\nu}},$$

which completes the proof.

A possible drawback with the above methods is the requirement of inverse assumptions. We shall now present a second order method where this demand is eliminated. The price we pay for this is that we use second order derivatives in the bilinear forms.

For $\varphi, \psi \in H^2$ and γ positive we define, following BRAMBLE and NITSCHE [1],

$$K_\gamma(\varphi,\psi) = D(\varphi,\psi) - \left\langle\varphi, \frac{\partial\psi}{\partial n}\right\rangle - \left\langle\frac{\partial\varphi}{\partial n}, \psi\right\rangle + \frac{k}{2}(\Delta\varphi, \Delta\psi) + \gamma[h^{-1}\langle\varphi,\psi\rangle + h\langle\nabla_s\varphi, \nabla_s\psi\rangle],$$

where ∇_s denotes the gradient within $\partial\Omega$. We have the following:

LEMMA 7.5. – *There is a γ_0 such that for $\gamma \geq \gamma_0$, K_γ is positive definite. Furthermore, there is a positive constant c such that*

$$K_\gamma(\varphi,\varphi)^{\frac{1}{2}} \geq ck\|\Delta\varphi\|^{\frac{1}{2}}, \qquad \varphi\in H^2.$$

PROOF. – See BRAMBLE and NITSCHE [1].

With this form, we use the Galerkin equations

$$(U_{n+1} - U_n, \chi) + \frac{k}{2}D(U_{n+1} - U_n, \chi) + \frac{k}{2}K_\gamma(U_{n+1} + U_n, \chi) = 0, \qquad \chi\in S_h,$$

that is we define

$$A_k(\varphi,\psi) = (\varphi,\psi) + \frac{k}{2}D(\varphi,\psi) + \frac{k}{2}K_\gamma(\varphi,\psi),$$

and

$$B_k(\varphi,\psi) = (\varphi,\psi) + \frac{k}{2}D(\varphi,\psi) - \frac{k}{2}K_\gamma(\varphi,\psi).$$

It is an immediate consequence of Lemma 7.5 that $A_k(\varphi,\varphi)$ is positive definite and we also have, using the appropriate trace inequalities,

$$a_k(\varphi) = A_k(\varphi,\varphi)^{\frac{1}{2}} \leq C\|\varphi\|_{H_h^2}. \tag{7.5}$$

Let $\mathcal{H}_k$ be the completion with respect to $a_k(\cdot)$ of $C^\infty(\bar{\Omega})$. Then $H^2 \subset \mathcal{H}_k$ and hence assuming that $\{S_h\}\in \mathcal{S}_{2,\nu}$, conditions (i) and (iv) are satisfied with $a=2$. The stability condition in (ii) is a trivial consequence of Lemma 7.5 and B_k is defined on $\dot{H}^4 \times \mathcal{H}_k$. In the same way as in the above method we have for $v \in \dot{H}^\infty$, $\psi \in C^\infty(\bar{\Omega})$,

$$D(v, \psi) = -(v, \Delta\psi),$$

$$K_\gamma(v, \psi) = -(\Delta v, \psi) + \frac{k}{2}(\Delta v, \Delta\psi),$$

so that

$$A_k(v, \psi) = \left(v - \frac{k}{2}\Delta v, \psi - \frac{k}{2}\Delta\psi\right),$$

$$B_k(v, \psi) = \left(v + \frac{k}{2}\Delta v, \psi - \frac{k}{2}\Delta\psi\right),$$

and hence by Lemma 5.1 for $2 \leqslant s \leqslant 6$,

$$|A_k(E(k)v, \psi) - B_k(v, \psi)| = \left|\left(\left(I - \frac{k}{2}\Delta\right)E(k)v - \left(I + \frac{k}{2}\Delta\right)v, \left(I - \frac{k}{2}\Delta\right)\psi\right)\right| \leqslant$$

$$\leqslant Ch^s\|v\|_{\dot{H}^s}\left\|\left(I - \frac{k}{2}\Delta\right)\psi\right\| \leqslant Ch^s\|v\|_{\dot{H}^s}\, a_k(\psi),$$

where we have used that by Lemma 7.5,

$$\left\|\left(I - \frac{k}{2}\Delta\right)\psi\right\| \leqslant \|\psi\| + \frac{k}{2}\|\Delta\psi\| \leqslant C a_k(\psi).$$

This proves that the consistency condition (iii) is satisfied (with $b = \mu = 2$) so that Theorem 1 applies.

To see that also Theorem 2 applies with the same $\mathcal{H}_k$ as above we only have to discuss condition (v). In this case

$$G_k(\varphi, \psi) = kK_\gamma(\varphi, \psi),$$

and by (7.5) the result follows with $\nu_0 = 0$ from the following:

LEMMA 7.6. – For $v \in \dot{H}^\nu$ given, the equations

$$K_\gamma(w - v, \chi) = 0, \qquad \chi \in S_h,$$

admit a unique solution $w = Q_h v \in S_h$ and

$$\|(I - Q_h)v\|_{\mathcal{H}_k} \leqslant Ch^\nu\|v\|_{\dot{H}^\nu}.$$

PROOF. – See BRAMBLE and NITSCHE [1].

8. – Least squares methods.

In this section we shall describe some methods which contain as special cases the methods in BRAMBLE and THOMÉE [3]. They have the advantage that no boundary behavior will be prescribed for the subspaces and no inverse assumptions will be needed. The class of methods will contain schemes of arbitrarily high order of accuracy.

We first consider a simple example where we do in fact assume that the functions in S_h vanish on the boundary. Thus suppose that $\{S_h\} \in \dot{\mathcal{S}}_{2,r}$ and consider, for U_n given, the problem of minimizing

$$\left\| \left(I - \frac{k}{2}\Delta\right)\varphi - \left(I + \frac{k}{2}\Delta\right)U_n \right\|^2$$

for $\varphi \in S_h$. An obvious calculation shows that the unique minimizing function U_{n+1} is obtained by the Galerkin equations

$$A_k(U_{n+1}, \chi) = B_k(U_n, \chi), \qquad \chi \in S_h,$$

where

$$A_k(\varphi, \psi) = \left(\left(I - \frac{k}{2}\Delta\right)\varphi, \left(I - \frac{k}{2}\Delta\right)\psi\right),$$

$$B_k(\varphi, \psi) = \left(\left(I + \frac{k}{2}\Delta\right)\varphi, \left(I - \frac{k}{2}\Delta\right)\psi\right).$$

We have for φ vanishing on $\partial\Omega$,

$$\left\| \left(I \pm \frac{k}{2}\Delta\right)\varphi \right\|^2 = \|\varphi\|^2 \mp kD(\varphi, \varphi) + \frac{k^2}{4}\|\Delta\varphi\|^2. \tag{8.1}$$

In particular, $a_k(\varphi) = A_k(\varphi, \varphi)^{\frac{1}{2}}$ defines a norm on the set of functions in $C^\infty(\bar{\Omega})$ which vanish on $\partial\Omega$. Let $\mathcal{H}_k$ be the Hilbert space obtained by completion. It is then easy to check that condition (i) holds with $a = 2$ and that the norm in $\mathcal{H}_k$ is equivalent uniformly in k to that in $\dot{H}_h^2$.

From (8.1) we obtain immediately,

$$\left\| \left(I + \frac{k}{2}\Delta\right)v \right\| \leq \left\| \left(I - \frac{k}{2}\Delta\right)v \right\|, \qquad v \in \mathcal{H}_k,$$

and hence the stability condition (ii) follows easily from the definitions.

By Lemma 5.1,

$$|A_k(E(k)v, \psi) - B_k(v, \psi)| = \left|\left(\left(I - \frac{k}{2}\Delta\right)E(k)v - \left(I + \frac{k}{2}\Delta\right)v, \left(I - \frac{k}{2}\Delta\right)\psi\right)\right| \leqslant$$

$$\leqslant \left\|\left(I - \frac{k}{2}\Delta\right)E(k)v - \left(I + \frac{k}{2}\Delta\right)v\right\| a_k(\psi) \leqslant Ch^s\|v\|_{\dot{H}^s} a_k(\psi), \qquad 2 \leqslant s \leqslant 6,$$

which proves the consistency condition (iii) with $b = \mu = 2$.

Since we have assumed condition (iv), it follows that Theorem 1 applies. We shall see that also Theorem 2 applies. Using the same $\mathcal{K}_k$ as in Theorem 1 it remains only to prove that condition (v) is satisfied, with $\nu_0 = 0$.

We have here

$$G_k(\varphi, \psi) = kD(\varphi, \psi) + \frac{k^2}{2}(\Delta\varphi, \Delta\psi).$$

We obtain at once with Q_h the projection with respect to the inner product $G_k(\varphi, \psi)$, and $\tilde{v} = (I - Q_h)v$, where $v \in \dot{H}^\nu$,

$$g_k(\tilde{v}) \leqslant C \inf_{\chi \in S_h} \|v - \chi\|_{H_h^2} \leqslant Ch^s\|v\|_{\dot{H}^s}, \qquad 2 \leqslant s \leqslant \nu. \tag{8.2}$$

In order to estimate $a_k(\tilde{v})$ it remains now only to estimate $\tilde{v}$ in $L^2(\Omega)$. We use again Nitsche's technique and let w be the solution of

$$-\Delta w = \tilde{v} \text{ in } \Omega, \qquad w = 0 \text{ on } \partial\Omega.$$

We obtain then, since $\tilde{v}$ vanishes on $\partial\Omega$,

$$\|\tilde{v}\|^2 = -(\Delta w, \tilde{v}) = D(w, \tilde{v}) = k^{-1}G_k(w, \tilde{v}) - \frac{k}{2}(\Delta w, \Delta\tilde{v}) = k^{-1}G_k(w, \tilde{v}) - \frac{k}{2}(\tilde{v}, \Delta\tilde{v}).$$

Setting $\tilde{w} = (I - Q_h)w$ and using the fact that by the definition of Q_h,

$$G_k(Q_h w, \tilde{v}) = 0,$$

we obtain

$$\|\tilde{v}\|^2 = k^{-1}G_k(\tilde{w}, \tilde{v}) + \frac{k}{2}D(\tilde{v}, \tilde{v}) \leqslant k^{-1}g_k(\tilde{w})g_k(\tilde{v}) + \frac{1}{2}g_k(\tilde{v})^2. \tag{8.3}$$

Application of (8.2) with $s = 2$ to w gives

$$k^{-1}g_k(\tilde{w}) \leqslant C\|w\|_{\dot{H}^2} \leqslant C\|\tilde{v}\|,$$

and we easily conclude from (8.2) and (8.3) that

$$\|\tilde{v}\| \leqslant C g_k(\tilde{v}) \leqslant C h^{\nu} \|v\|_{\dot{H}^{\nu}},$$

which completes the proof.

We shall now turn to the general situation and consider, for a rational function

$$r(\tau) = \frac{b(\tau)}{a(\tau)}$$

with

$$a(\tau) = \sum_{j=0}^{\alpha} a_j \tau^j, \qquad b(\tau) = \sum_{j=0}^{\beta} b_j \tau^j,$$

where $\beta \leqslant \alpha$ and

$$a_0 = b_0 = 1, \qquad |b_j| < a_j, \qquad j = 1, \dots, \alpha, \tag{8.4}$$

the problem of minimizing, for $\varphi \in S_h$,

$$\|a(-k\Delta)\varphi - b(-k\Delta)\, U_n\|^2 .$$

This time we do not want to assume that the elements of S_h vanish on $\partial\Omega$ and in minimizing we therefore add a boundary term to the above expression so that we minimize with a certain positive number γ, the size of which will be made precise below,

$$\|a(-k\Delta)\varphi - b(-k\Delta)\, U_n\|^2 + \gamma \sum_{j=0}^{\alpha-1} k^{2j+\frac{1}{2}} |\Delta^j \varphi|^2 .$$

Notice that for t positive the exact solution of the continuous problem not only vanishes on $\partial\Omega$ but that also $\Delta^j u(x,t) = 0$, $x \in \partial\Omega$, $j = 1, 2, \dots$, so that the requirement that certain $\Delta^j \varphi$ be small on $\partial\Omega$ is natural.

The minimizing function U_{n+1} satisfies the Galerkin equations with

$$A_k(\varphi, \psi) = \big(a(-k\Delta)\varphi, a(-k\Delta)\psi\big) + \gamma \sum_{j=0}^{\alpha-1} k^{2j+\frac{1}{2}} \langle \Delta^j \varphi, \Delta^j \psi \rangle ,$$
$$B_k(\varphi, \psi) = \big(b(-k\Delta)\varphi, a(-k\Delta)\psi\big) .$$

Obviously A_k is positive definite. Letting $\mathcal{H}_k$ be the completion of $C^\infty(\bar{\Omega})$ with respect to $a_k(\cdot)$ and taking $\{S_h\} \in \mathbb{S}_{2\alpha,\nu}$ we find that (i), (iv) are satisfied with $a = 2\alpha$.

We have for $v \in \dot{H}^\infty$, $\psi \in C^\infty(\bar{\Omega})$,

$$|A_k(E(k)v, \psi) - B_k(v, \psi)| = |\big(a(-k\Delta)E(k)v - b(-k\Delta)v, a(-k\Delta)\psi\big)| \leqslant$$
$$\leqslant \|a(-k\Delta)E(k)v - b(-k\Delta)v\| \cdot \|a(-k\Delta)\psi\| \leqslant C h^s \|v\|_{\dot{H}^s} a_k(\psi), \qquad 2\beta \leqslant s \leqslant 2\mu + 2,$$

by Lemma 5.1, which proves (iii) with $b = 2\beta$.

Since B_k is defined on $\dot{H}^{2\beta}\times\mathcal{K}_k$, in order to be able to apply Theorem 1 it remains only to discuss the stability inequality in (ii). We have

$$|B_k(\varphi, \psi)| = |(b(-k\Delta)\varphi, a(-k\Delta)\psi)| \leqslant \|b(-k\Delta)\varphi\|\cdot\|a(-k\Delta)\psi\| .$$

The stability requirement is therefore satisfied for large enough γ by the following lemma:

LEMMA 8.1. – For any positive K there is a γ_0 such that for any rational function $r(\tau)$ satisfying (8.4) and with $\max_j a_j \leqslant K$, we have for $\gamma \geqslant \gamma_0$, $v\in C^\infty(\bar{\Omega})$,

$$\|b(-k\Delta)v\|^2 \leqslant \|a(-k\Delta)v\|^2 + \gamma\sigma^{-\frac{1}{2}}\sum_{j=0}^{\alpha-1} k^{2j+\frac{1}{2}}|\Delta^j v|^2 ,$$

where

$$\sigma = \min_{j=1,\dots,\alpha}(a_j^2 - b_j^2) .$$

For the purpose of the proof we introduce some notation. First let

$$V = \Big(\sum_{j=0}^{\alpha-1} k^{2j+\frac{1}{2}}|\Delta^j v|^2\Big)^{\frac{1}{2}} .$$

Secondly, for v given and $j=1,\dots,\alpha$, let $v = H_j + z_j$ where

$$\Delta^j H_j = 0 \text{ in } \Omega , \qquad \Delta^l H_j = \Delta^l v \text{ on } \partial\Omega , \qquad l = 0, \dots, j-1 ,$$
$$\Delta^j z_j = \Delta^j v \text{ in } \Omega , \qquad \Delta^l z_j = 0 \text{ on } \partial\Omega , \qquad l = 0, \dots, j-1 ,$$

and set

$$Z = \left(\sum_{j=0}^{\alpha-1} k^{2j+\frac{3}{2}}\left|\frac{\partial\Delta^j z_{j+1}}{\partial n}\right|^2\right)^{\frac{1}{2}} .$$

We proceed to prove three lemmas.

LEMMA 8.2. – There is a constant C such that for any $\varepsilon > 0$ we have with $\delta = (\varepsilon/4C)^{\frac{1}{2}}$, $\varkappa = \frac{1}{2}C^{\frac{1}{2}}\varepsilon^{-\frac{1}{2}}$ that

$$k(v, \Delta v) + \delta k^{\frac{3}{2}}\left|\frac{\partial z_1}{\partial n}\right|^2 \leqslant \varepsilon k^2\|\Delta v\|^2 + \varkappa k^{\frac{1}{2}}|v|^2 .$$

PROOF. – We have using Green's formula and the fact that $D(H_1, z_1) = 0$,

$$(v, \Delta v) = (v, \Delta z_1) = \left\langle v, \frac{\partial z_1}{\partial n}\right\rangle - D(v, z_1) = \left\langle v, \frac{\partial z_1}{\partial n}\right\rangle - D(z_1, z_1) . \tag{8.5}$$

Since $z_1 = 0$ on $\partial\Omega$ we have for any $\varepsilon_1 > 0$ (Lemma 4.2 in BRAMBLE and THOMÉE [3]),

$$\left|\frac{\partial z_1}{\partial n}\right|^2 \leq \varepsilon_1 \|\Delta z_1\|^2 + \frac{C}{\varepsilon_1} D(z_1, z_1), \tag{8.6}$$

and hence, since $\Delta z_1 = \Delta v$, by adding the appropriate multiples of (8.5) and (8.6) with $\delta = (\varepsilon/(4C))^{\frac{1}{2}}$, $\varepsilon_1 = 2\delta C k^{\frac{1}{2}}$,

$$k(v, \Delta v) + 2\delta k^{\frac{3}{2}} \left|\frac{\partial z_1}{\partial n}\right|^2 \leq \varepsilon k^2 \|\Delta v\|^2 + k \left\langle v, \frac{\partial z_1}{\partial n}\right\rangle,$$

from which Lemma 8.2 follows with $\varkappa = (4\delta)^{-1}$.

LEMMA 8.3. – With ε, δ and $\varkappa$ as above we have

$$\sum_{j=0}^{\alpha-1} k^{2j+1}(\Delta^j v, \Delta^{j+1} v) + \delta Z^2 \leq \varepsilon \sum_{j=1}^{\alpha} k^{2j} \|\Delta^j v\|^2 + \varkappa V^2.$$

PROOF. – This follows at once by applying Lemma 8.2 to $(k\Delta)^j v$ for $j = 0, \dots, \alpha - 1$ and adding if we notice that

$$\Delta^j v = \Delta^j H_{j+1} + \Delta^j z_{j+1},$$

where $\Delta^j H_{j+1}$ is harmonic and $\Delta^j z_{j+1}$ vanishes on $\partial\Omega$.

LEMMA 8.4. – For $j < l - 1$ we have

$$(\Delta^j v, \Delta^l v) = (\Delta^{j+1} v, \Delta^{l-1} v) + \left\langle \Delta^j v, \frac{\partial \Delta^{l-1} z_l}{\partial n}\right\rangle - \left\langle \Delta^{l-1} v, \frac{\partial \Delta^j z_{j+1}}{\partial n}\right\rangle.$$

PROOF. – Using Green's formula we obtain

$$\begin{aligned}(\Delta^j v, \Delta^l v) - (\Delta^{j+1} v, \Delta^{l-1} v) &= (\Delta^j v, \Delta^l z_l) - (\Delta^{j+1} z_{j+1}, \Delta^{l-1} v) = \\ &= \left\langle \Delta^j v, \frac{\partial \Delta^{l-1} z_l}{\partial n}\right\rangle - \left\langle \Delta^{l-1} v, \frac{\partial \Delta^j z_{j+1}}{\partial n}\right\rangle - D(\Delta^j v, \Delta^{l-1} z_l) + D(\Delta^{l-1} v, \Delta^j z_{j+1}).\end{aligned}$$

Here

$$D(\Delta^j v, \Delta^{l-1} z_l) = D(\Delta^j H_{j+1}, \Delta^{l-1} z_l) + D(\Delta^j z_{j+1}, \Delta^{l-1} z_l) = D(\Delta^j z_{j+1}, \Delta^{l-1} z_l),$$

since $\Delta^j H_{j+1}$ is harmonic and $\Delta^{l-1} z_l$ vanishes on $\partial\Omega$. Similarly

$$D(\Delta^{l-1} v, \Delta^j z_{j+1}) = D(\Delta^{l-1} z_l, \Delta^j z_{j+1}),$$

which completes the proof.

PROOF OF LEMMA 8.1. – Consider

$$R = \|a(-k\Delta)v\|^2 - \|b(-k\Delta)v\|^2 = \sum_{j,l\leqslant\alpha} (-k)^{j+l}(a_j a_l - b_j b_l)(\Delta^j v, \Delta^l v).$$

For $j + l = 2r$ even we obtain by repeated use of Lemma 8.4 and Cauchy's inequality,

$$(-k)^{j+l}(\Delta^j v, \Delta^l v) \geqslant k^{2r}\|\Delta^r v\|^2 - ZV,$$

and for $j + l = 2r + 1$ odd we obtain similarly

$$(-k)^{j+l}(\Delta^j v, \Delta^l v) \geqslant -k^{2r+1}(\Delta^r v, \Delta^{r+1} v) - ZV.$$

Hence there are positive constants c_1 and c_2 such that

$$R \geqslant \sigma \sum_{r=1}^{\alpha} k^{2r}\|\Delta^r v\|^2 - c_1 \sum_{r=0}^{\alpha-1} k^{2r+1}(\Delta^r v, \Delta^{r+1} v) - c_2 ZV.$$

Hence by Lemma 8.3,

$$R \geqslant (\sigma - \varepsilon c_1) \sum_{r=1}^{\alpha} k^{2r}\|\Delta^r v\|^2 + c_1 \delta Z^2 - c_1 \varkappa V^2 - c_2 ZV.$$

Choose now $\varepsilon = \frac{1}{2}\sigma/c_1$. Then using the form of δ and $\varkappa$ we obtain

$$R \geqslant \tfrac{1}{2}\sigma \sum_{r=1}^{\alpha} k^{2r}\|\Delta^r v\|^2 - \gamma\sigma^{-\frac{1}{2}} V^2,$$

for sufficiently large γ, which completes the proof.

We now turn to the application of Theorem 2. We shall use the same space $\mathcal{H}_k$ as above so that it only remains to discuss condition (v). We have here

$$G_k(\varphi, \psi) = \big(g(-k\Delta)\varphi, a(-k\Delta)\psi\big) + \gamma \sum_{j=0}^{\alpha-1} k^{2j+\frac{1}{2}} \langle \Delta^j \varphi, \Delta^j \psi \rangle,$$

where

$$g(\tau) = a(\tau) - b(\tau) = \sum_{j=1}^{\alpha} g_j \tau^j, \qquad g_j > 0,\ j = 1, \dots, \alpha.$$

LEMMA 8.5. – For $\tilde{\gamma}$ sufficiently large there is a positive constant c such that for $\gamma \geqslant \tilde{\gamma}$,

$$G_k(v, v) \geqslant c\left\{\sum_{j=1}^{\alpha} k^{2j}\|\Delta^j v\|^2 + \sum_{j=0}^{\alpha-1} k^{2j+\frac{1}{2}}|\Delta^j v|^2\right\}, \qquad v \in C^\infty(\bar{\Omega}).$$

PROOF. – In the same way as in the proof of Lemma 8.1 we see that

$$\tilde{R} = \big(g(-k\Delta)v,\, a(-k\Delta)v\big) = \sum_{j,l} (-k)^{j+l} g_j a_l(\Delta^j v, \Delta^l v) \geq c \sum_{j=1}^{\alpha} k^{2j} \|\Delta^j v\|^2 - \tilde{\gamma} V^2 ,$$

and the result threfore follows at once.

It follows that with $g_k(v) = G_k(v, v)^{\frac{1}{2}}$ we have

$$G_k(\varphi, \psi) \leq C g_k(\varphi) a_k(\psi) . \tag{8.7}$$

As a consequence of the positivity of $G_k(v, v)$ the equations

$$G_k(w - v, \chi) = 0 , \qquad \chi \in S_h ,$$

admit, for $v \in \dot{H}^\nu$ given, a unique solution $w = Q_h v \in S_h$. It remains to estimate $a_k(\tilde{v})$ where $\tilde{v} = (I - Q_h)v$. As a consequence of Lemma 8.5,

$$a_k(v) \leq C\big(\|v\| + g_k(v)\big) .$$

Hence, since

$$g_k(\tilde{v}) \leq C \inf_{\chi \in S_h} \|v - \chi\|_{H_h^{2\alpha}} \leq C h^\nu \|v\|_{\dot{H}^\nu} ,$$

it remains to obtain a similar estimate for $\tilde{v}$ in $L^2(\Omega)$.

For this purpose, let $w \in \dot{H}^{2\alpha}$ be the solution of

$$g(-k\Delta) w = \tilde{v} \text{ in } \Omega ,$$
$$\Delta^j w = 0 , \qquad j < \alpha , \qquad \text{on } \partial\Omega .$$

This solution exists and is unique by the positivity of g. We may then write

$$\|\tilde{v}\|^2 = \big(\tilde{v}, g(-k\Delta) w\big) = \big(g(-k\Delta)\tilde{v}, w\big) + \Gamma_k(\tilde{v}, w) , \tag{8.8}$$

where $\Gamma_k(\tilde{v}, w)$ are the boundary terms obtained in the integration by parts.

Consider first the first term on the right in (8.8). It may be written

$$\big(g(-k\Delta)\tilde{v}, w\big) = \big(g(-k\Delta)\tilde{v}, a(-k\Delta) w\big) + \Big(g(-k\Delta)\tilde{v}, \big(I - a(-k\Delta)\big) w\Big) .$$

Now by the definition of v we have for $\chi \in S_h$,

$$\big(g(-k\Delta)\tilde{v}, a(-k\Delta) w\big) = G_k(\tilde{v}, w) = G_k(\tilde{v}, w - \chi) ,$$

so that by (8.7), using Lemmas 2.7 and 2.6,

$$|(g(-k\Delta)\tilde{v}, a(-k\Delta)w)| \leqslant Cg_k(\tilde{v}) \inf_{\chi\in S_h} a_k(w-\chi) \leqslant Cg_k(\tilde{v}) \inf_{\chi\in S_h} \|w-\chi\|_{B_h^{2\alpha}} \leqslant Cg_k(\tilde{v})h^{2\alpha}\|w\|_{\dot{H}^{2\alpha}}.$$

For the purpose of estimating the last factor on the right we notice that since

$$\tau^\alpha \leqslant Cg(\tau), \qquad \tau > 0,$$

we have

$$h^{2\alpha}\|w\|_{\dot{H}^{2\alpha}} \leqslant C\big(\sum_m (k\lambda_m)^{2\alpha}w_m^2\big)^{\frac12} \leqslant C\big(\sum_m g(k\lambda_m)^2 w_m^2\big)^{\frac12} \leqslant C\|g(-k\Delta)w\| = C\|\tilde{v}\|,$$

so that

$$|(g(-k\Delta)\tilde{v}, a(-k\Delta)w)| \leqslant Cg_k(\tilde{v})\|\tilde{v}\|.$$

Similarly, since

$$|a(\tau)-1| \leqslant Cg(\tau), \qquad \tau > 0,$$

we obtain

$$\|(I-a(-k\Delta))w\| \leqslant C\|g(-k\Delta)w\| = C\|\tilde{v}\|,$$

and hence

$$\big|\big(g(-k\Delta)\tilde{v}, (I-a(-k\Delta))w\big)\big| \leqslant Cg_k(\tilde{v})\|\tilde{v}\|,$$

so that altogether,

$$|(g(-k\Delta)\tilde{v}, w)| \leqslant Cg_k(\tilde{v})\|\tilde{v}\|. \tag{8.9}$$

We now want to consider the boundary term $\Gamma_k(\tilde{v}, w)$. Assume first $\{S_h\} \in \dot{S}_{2\alpha,\nu}$. Then $\tilde{v} \in \dot{H}^{2\alpha}$ and $\Gamma_k(\tilde{v}, w) = 0$. In this case we may thus conclude from (8.8) and (8.9) that

$$\|\tilde{v}\| \leqslant Cg_k(\tilde{v}),$$

and hence

$$a_k(\tilde{v}) \leqslant Cg_k(\tilde{v}) \leqslant Ch^\nu\|v\|_{\dot{H}^\nu}.$$

In this case Theorem 2 applies with $\nu_0 = 0$.

Consider now the general case $\{S_h\} \in S_{2\alpha,\nu}$. We have for any $\varepsilon > 0$,

$$|\Gamma_k(\tilde{v}, w)| \leqslant C\sum_{j=1}^{\alpha}\sum_{l=0}^{j-1} k^j \left|\left\langle \frac{\partial}{\partial n}\Delta^{j-l-1}w, \Delta^l\tilde{v}\right\rangle\right| \leqslant C\left[\sum_{j=0}^{\alpha-1} \varepsilon k^{2j+\frac32}\left|\frac{\partial}{\partial n}\Delta^j w\right|^2 + \varepsilon^{-1}g_k(\tilde{v})^2\right].$$

But since, for $j<\alpha$, $\Delta^j w=0$ on $\partial\Omega$ we have for these j,

$$\left|\frac{\partial}{\partial n}\Delta^j w\right|\leqslant C\|\Delta^{j+1}w\| .$$

Since $\tau^{j+1}\leqslant Cg(\tau)$ we obtain as above

$$k^{j+1}\|\Delta^{j+1}w\|=\big(\sum_m(k\lambda_m)^{2(j+1)}w_m^2\big)^{\frac12}\leqslant C\|\tilde v\| ,$$

so that with ε a small multiple of $k^{\frac12}$,

$$|\Gamma_k(\tilde v, w)|\leqslant\tfrac12\|\tilde v\|^2+Ck^{-\frac12}g_k(\tilde v)^2 .$$

This gives with (8.8) and (8.9),

$$\|\tilde v\|\leqslant Ck^{-\frac12}g_k(\tilde v) ,$$

so that finally

$$a_k(\tilde v)\leqslant Ch^{\nu-\frac12}\|v\|_{\dot H^\nu} .$$

In this case Theorem 2 applies with $\nu_0=\frac12$.

REMARK. – Consider now the case in which we only have

$$a_j>0 , \qquad |b_j|\leqslant a_j , \qquad j=0, \dots, \alpha .$$

This for instance is the case with the diagonal Padé approximations. We may then apply Lemma 8.1 to $(1+\beta k)a(\tau)$ and $b(\tau)$ for some $\beta>0$ and obtain with a new γ,

$$\|b(-k\Delta)v\|^2\leqslant(1+\beta k)^2\Big[\|a(-k\Delta)v\|^2+\gamma\sum_{j=0}^{\alpha-1}k^{2j}|\Delta^j v|^2\Big] .$$

Defining this time

$$A_k(\varphi, \psi)=\big(a(-k\Delta)\varphi, a(-k\Delta)\psi\big)+\gamma\sum_{j=0}^{\alpha-1}k^{2j}\langle\Delta^j\varphi, \Delta^j\psi\rangle ,$$

we obtain with B_k as before

$$|B_k(\varphi, \psi)|\leqslant(1+\beta k)a_k(\varphi)a_k(\psi) .$$

In the special case

$$r(\tau)=\frac{1-\frac12\tau}{1+\frac12\tau} ,$$

this was used in BRAMBLE and THOMÉE [3] to obtain results for the corresponding scheme with the assumption $kh^{-2}=$ constant replaced by $k^2h^{-3}\geqslant$ constant.

REFERENCES

[1] J. BRAMBLE - J. A. NITSCHE, *A generalized Ritz-least-squares method for Dirichlet problems*, SIAM J. Numer. Anal., **10** (1973), pp. 81-93.

[2] J. H. BRAMBLE - A. H. SCHATZ, *Rayleigh-Ritz-Galerkin methods for Dirichlet's problem using subspaces without boundary conditions*, Comm. Pure. Appl. Math., **23** (1970), pp. 653-675.

[3] J. BRAMBLE - V. THOMÉE, *Semidiscrete-least squares method for a parabolic boundary value problem*, Math. Comp., **26** (1972), pp. 633-647.

[4] J. DOUGLAS - T. DUPONT, *Galerkin methods for parabolic equations*, SIAM J. Numer. Anal., **7** (1970), pp. 575-626.

[5] J. A. NITSCHE, *Linear Spline-Funktionen und die Methoden von Ritz für elliptische Randwertprobleme*, Arch. Rational Mech. Anal., **36** (1970), pp. 348-355.

[6] J. A. NITSCHE, *Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind*, Abh. Math. Sem. Univ. Hamburg, **36** (1971), pp. 9-15.

[7] H. S. PRICE - R. S. VARGA, *Error bounds for semi-discrete Galerkin approximations of parabolic problems with application to petroleum reservoir mechanics*, Numerical Solution of Field Problems in Continuum Physics, AMS Providence, R.I. (1970), pp. 74-94.

[8] R. S. VARGA, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1962.