MATHEMATICS OF COMPUTATION
VOLUME 47, NUMBER 175
JULY 1986, PAGES 55-75

Galerkin Approximations of Abstract Parabolic Boundary Value Problems With Rough Boundary Data–L_p Theory*

By I. Lasiecka

Abstract. Galerkin approximations of an abstract parabolic boundary value problem with "rough" boundary data are considered. The optimal rates of convergence in $L_p[0T; L_2(\Omega)]$ norms for $L_p[0T; L_2(\Gamma)]$ boundary terms are derived.

1. Introduction. Let Ω be an open, bounded domain in R^n with smooth boundary Γ. As a motivation for the present paper, let us consider the following two canonical examples of parabolic problems with "rough" boundary data:

$$\text{(1.1.D)} \qquad \begin{cases} y_t(t) = \Delta y & \text{in } Q \equiv \Omega \times [0, T], \\ y(0) = 0, \\ y|_\Gamma = u \in L_p[0T; L_2(\Gamma)] \end{cases}$$

and

$$\text{(1.1.N)} \qquad \begin{cases} y_t(t) = (\Delta - 1) & \text{in } Q, \\ y(0) = 0, \\ \left.\dfrac{\partial y}{\partial \eta}\right|_\Gamma = u \in L_p[0T; H^{-1}(\Gamma)], & 1 \leqslant p \leqslant \infty. \end{cases}$$

Our interest is a study of the rates of convergence of Galerkin approximations to (1.1.D), (1.1.N) in the $L_p(0T; L_2(\Omega))$-norms, with boundary data either in $L_p(0T; L_2(\Gamma))$ (Dirichlet case) or else in $L_p(0T; H^{-1}(\Gamma))$ (Neumann case).

A standard technique of treating nonhomogeneous boundary conditions consists in subtracting the effect of the boundary term and then considering the corresponding nonhomogeneous equation with homogeneous boundary conditions (see [5]). Application of these techniques requires, however, a certain smoothness of the boundary function (at least $H^{1/2}(\Gamma)$ for the Dirichlet case). This requirement is needed to carry over standard variational arguments based on H^1-coercivity of the bilinear form associated with the differential operator Δ. Thus, our assumption that u is only in $L_p(0T; L_2(\Gamma))$ (resp. $L_p(0T; H^{-1}(\Gamma))$) in the Dirichlet case (resp. Neumann case) is the *distinctive* feature of this work.

Received November 19, 1984; revised October 2, 1985.
1980 *Mathematics Subject Classification*. Primary 65N30.
*Research partially supported by the National Science Foundation under grant DMS-8301668.

We shall construct a Galerkin approximation of (1.1.D) and (1.1.N) which will yield the *optimal* rates of convergence equal to $\mathcal{O}(\sqrt{h})$ measured in the $L_p(0T; L_2(\Omega))$-topology for $1 < p < \infty$, with $L_p[0T; L_2(\Gamma)]$-boundary terms in (1.1.D) (resp. $L_p(0T; H^{-1}(\Gamma))$-boundary terms in (1.1.N)). In the limit case, when $p = 1$ or $p = \infty$, the corresponding optimal rates of convergence are proved to be $\mathcal{O}(\sqrt{h}\ \ln h)$.

The approach taken in this paper is based on semigroup theory combined with the theory of singular integrals.

We shall first consider a general abstract model of the form:

$$\tag{1.2} \begin{cases} y_t(t) = -Ay(t) + A^Q Bu(t);^{**} & 0 \leqslant Q \leqslant 1 \text{ on } \left(D(A^*)\right)', \\ y(0) = 0, \end{cases}$$

where $-A$ is the generator of an analytic semigroup $S(t)$ on a Hilbert space H and B is a bounded operator from another Hilbert space U into H. $D(A^*)'$ stands for the dual space to $D(A^*)$ with respect to $L_2(Q)$, equipped with the graph topology.

Model (1.2) is suitable to treat nonhomogeneous boundary problems with "rough" boundary data, and in particular it covers as a special case the two canonical examples given by (1.1.D) and (1.1.N), as illustrated below.

Dirichlet Case. Here, in order to represent (1.1.D) in the form (1.2), we introduce the operator $A\colon L_2(\Omega) \to L_2(\Omega)$ defined by

$$\tag{1.3} -Ay = \Delta y, \qquad y \in \mathscr{D}(A) \equiv H_0^1(\Omega) \cap H^2(\Omega).$$

It is well known that $-A$ generates an analytic semigroup $S(t)$ on $L_2(Q)$. Next, let us define the "Dirichlet" map $D\colon L_2(\Gamma) \to L_2(\Omega)$ to be just a harmonic extension of the boundary data g, i.e., $Dg = v$ if and only if

$$\tag{1.4} \Delta v = 0 \quad \text{on } \Omega, \qquad v|_\Gamma = g \quad \text{on } \Gamma.$$

An abstract version of problem (1.1.D) is given by the following semigroup formula (see, e.g., [2], [11])

$$\tag{1.5} y(t) = A\int_0^t S(t-z)Du(z)\,dz \quad \text{on } L_2(\Omega).$$

Knowing that $D \in \mathscr{L}(L_2(\Gamma) \to H^{1/2}(\Omega))$ [15] and that $H^{1/2}(Q) \subset \mathscr{D}(A^{1/4-\varepsilon})$, $\varepsilon > 0$, [6] we have

$$\tag{1.6} A^{1/4-\varepsilon}D \in \mathscr{L}\left(L_2(\Gamma) \to L_2(\Omega)\right).$$

After setting $H \equiv L_2(\Omega)$; $U \equiv L_2(\Gamma)$; $B = A^{1/4-\varepsilon}D$; $Q = 3/4 + \varepsilon$, we can rewrite (1.5) in differential form as

$$\tag{1.7} \begin{cases} y_t(t) = -Ay(t) + ADu(t) = -Ay + A^Q Bu, \\ y(0) = 0 \quad \text{on } \mathscr{D}(A^*)'. \end{cases}$$

Thus, (1.7) (a special case of (1.2)) can be interpreted as an abstract version of (1.1.D).

** Without loss of generality we assume that the spectrum of $-A$ lies in the left complex plane, hence the fractional powers A^Q are well defined.

Neumann Case. To treat Neumann boundary conditions, we proceed in a similar fashion. With $-Ay \equiv \Delta y - y$ defined on

$$\mathscr{D}(\mathscr{A}) \equiv \{ y \in L_2(\Omega); \Delta y \in L_2(\Omega); \partial y/\partial\eta|_\Gamma = 0 \},$$

we associate the corresponding semigroup $S(t)$ and we define the "Neumann map" N: $L_2(\Gamma) \to L_2(\Omega)$ by $Ng = v$, where

$$(1.8) \qquad (\Delta - 1)v = 0 \quad \text{on } \Omega, \qquad \frac{\partial}{\partial\eta} v \bigg|_\Gamma = g \quad \text{on } \Gamma.$$

Since [15] $N \in \mathscr{L}(H^{-1}(\Gamma) \to H^{1/2}(\Omega)) \equiv \mathscr{D}(A^{1/4})$ [6], with $H \equiv L_2(\Omega)$, $U = H^{-1}(\Gamma)$, we have

$$B \equiv A^{1/4}N \in \mathscr{L}(U \to H).$$

The solution y of (1.1.N) can now be written as (e.g. [11])

$$(1.9) \qquad \begin{cases} y_t(t) = -Ay(t) + ANu(t) = -Ay(t) + A^Q Bu(t), \\ y(0) = 0, \end{cases}$$

where $Q = 3/4$.

Thus, here again, (1.9) is an abstract version of the boundary problem (1.1.N). The same procedure described above applies to an arbitrary elliptic operator (see Section 5). We are here justified in viewing (1.2) as an *abstract* model for an arbitrary parabolic boundary value problem with nonhomogeneous boundary conditions, after giving an appropriate meaning to the operators A and B.

For the abstract model we shall now define *Galerkin approximations* and we shall establish the optimal rates of convergence of the approximate solutions. These abstract results will then be applied to the original parabolic problem.

A brief description of our general setting for Galerkin approximations to model (1.2) follows. We introduce a suitably chosen family of finite-dimensional approximating subspaces $V_h \subset \mathscr{D}(B^*A^{*Q})$, as well as a sequence of finite-dimensional operators A_h: $V_h \to V_h$ approximating (in the sense described later) the generator A.

As Galerkin approximation of the abstract model (1.2) we then take:

Find $y_h(t) \subset V_h$ such that

$$(1.10) \qquad \begin{cases} (\dot{y}_h(t), v_h)_H = -(A_h y_h(t), v_h)_H + \langle u(t), B^*A^{*Q}v_h \rangle_U, \\ y_h(0) = 0 \quad \text{for all } v_h \in V_h. \end{cases}$$

By crucially using the analyticity of the original semigroup $S(t)$, as well as the *uniform analyticity* of the underlined Galerkin approximation $S_h(t) = e^{A_h t}$, we shall prove that Galerkin approximations to (1.2) yield the optimal rates of convergence (optimal with respect to the maximal regularity of the solutions). The main tools used at this stage are: (i) the theory of singular integrals combined with interpolation theory and (ii) estimates for initial value problems with "rough" data. The latter follow from the above-mentioned uniform analyticity of $S_h(t)$.

We now specialize the above procedure to our two canonical examples. In the *Dirichlet case* (1.1.D), we can take as space V_h the space V_h^0 of linear splines vanishing on the boundary Γ and as approximation A_h the standard Galerkin approximation of A. Thus, the version of (1.10) for the special case of Dirichlet data, i.e., the Galerkin approximation of (1.7), will take the form

$$(1.11) \quad (\dot{y}_h(t), v_h)_Q = -(\nabla y_h(t), \nabla v_h)_\Omega - \langle u(t), D^*A^*v_h \rangle_\Gamma, \qquad v_h \in V_h^0.$$

Using $D^*A^* = \partial v/\partial\eta|_\Gamma$ ([2], [12]), we see that (1.11) is equivalent to

$$(1.11') \quad (\dot{y}_h(t), v_h)_\Omega = -(\nabla y_h(t), \nabla v_h)_\Omega - \left\langle u(t), \frac{\partial}{\partial\eta} v_h \right\rangle_\Gamma, \qquad v_h \in V_h^0.$$

The convergence result obtained for the abstract model (1.10) will yield in this Dirichlet case the following result:

With y (resp. y_h) the solution to (1.1.D) (resp. (1.11′)), we shall prove that

$$(1.12) \quad \begin{cases} |y - y_h|_{L_p[0T;\, L_2(\Omega)]} \leqslant Ch^{1/2}|u|_{L_p[0T;\, L_2(\Gamma)]}, & 1 < p < \infty, \\ |y - y_h|_{L_p[0T;\, L_2(\Omega)]} \leqslant Ch^{1/2} \ln h|u|_{L_p[0T;\, L_2(\Gamma)]}, & p = 1, \infty. \end{cases}$$

To treat *Neumann* boundary conditions, we can take for the approximating subspace V_h the space of linear splines with no requirements of vanishing on the boundary and for A_h the standard Galerkin approximation of A. The version of the Galerkin approximation (1.10) in the Neumann case will then be

$$(1.13) \quad \begin{cases} (\dot{y}_h(t), v_h)_\Omega = -(\nabla y_h(t), \nabla u_h)_\Omega - \langle u(t), N^*A^*v_h\rangle_\Gamma, & v_h \in V_h, \\ y_h(0) = 0. \end{cases}$$

Since $N^*A^*v = v|_\Gamma$, we have

$$(1.13') \quad \begin{cases} (\dot{y}_h(t), v_h)_\Gamma = -(\nabla y_h(t), \nabla v_h)_\Omega - \langle u(t), v_h\rangle_\Gamma, & v_h \in V_h, \\ y_h(0) = 0. \end{cases}$$

With y (resp. y_h) the solution to (1.1.N) (resp. (1.13′)), we will prove the following rates of convergence,

$$(1.14) \quad \begin{cases} |y - y_h|_{L_p[0T;\, L_2(\Omega)]} \leqslant Ch^{1/2}|u|_{L_p[0T;\, H^{-1}(\Gamma)]}, & 1 < p < \infty, \\ |y - y_h|_{L_p[0T;\, L_2(\Omega)]} \leqslant Ch^{1/2} \ln h|u|_{L_p[0T;\, H^{-1}(\Gamma)]}, & p = 1, \infty. \end{cases}$$

Remarks. 1. The convergence results with "rough" data obtained in (1.12) and (1.14) are optimal. They reflect, in fact, the maximal regularity of the continuous solutions, and they are of the same order as the "best approximation" to $y(t)$.

2. If one considers smoother boundary data, then one would expect to obtain higher than $\sqrt{h}$ rates of convergence. In fact, for the Neumann problem this is indeed the case. One can show that the algorithm (1.13′), when applied to smooth u, will yield the optimal (with respect to the optimal regularity of the solutions y) rates of convergence. In contrast, in the Dirichlet case, the algorithm (1.11) limits its accuracy to $\sqrt{h}$, no matter how smooth the boundary data "u" are. This is because in the Dirichlet case the approximating elements are forced to approximate (in some sense) the generator A, hence they satisfy zero (or more generally nearly zero) boundary conditions. This fact makes it impossible to achieve higher order of accuracy (higher than $\sqrt{h}$) for the approximations of the solutions to (1.1.D) with nonhomogeneous boundary data.

The outline of the paper is as follows: Section 2 deals with the abstract model (1.2) and provides the maximal regularity results for the original solutions. In Section 3, Galerkin approximations of the abstract problem (1.2) are introduced and the main abstract convergence results are formulated. Section 4 is devoted to the proofs of the results of Section 3. Section 5 deals with the application of the abstract results to the parabolic problem (1.1), where it provides the optimal rates of convergence for the approximation of (1.1), expressed in terms of $L_p(L_2)$-norms.

Notation. X' is the dual of X.

$\|\cdot\|_{H\to H}$ = norm in $\mathscr{L}(H \to H)$.

$\|x\|$, $(\cdot,)$ are the norm and scalar product in $L_2(\Omega)$.

$|\ |$, $\langle\ ,\ \rangle$ are the norm and scalar product in $L_2(\Gamma)$.

$\|\cdot\|_s$, $|\cdot|_s$ are the norms in $H^s(\Omega)$ and $H^s(\Gamma)$, respectively.

$\hat{K}(\lambda)$ is the Fourier transform on $K(t)$.

2. Abstract Parabolic Boundary Value Problem. Let A be the generator of an analytic semigroup $S(t)$ on a Banach space H. It is well known (see [17]) that there exist constants $a, b > 0$ and $C > 0$ such that for the resolvent set $\rho(A)$ of A one has

$$\rho(A) \supset \Sigma \equiv \{\lambda; \operatorname{Re}\lambda > a - b\operatorname{Im}|\lambda|\}$$

and for all $\lambda \in \Sigma$ the resolvent operator $(\lambda I - A)^{-1}$ of A satisfies

$$\|R(\lambda, A)\|_{H\to H} \leqslant C(1+|\lambda|)^{-1}, \tag{2.1}$$

or equivalently,

$$\|A^n S(t)\|_{H\to H} \leqslant \frac{C}{t^n}e^{at}; \qquad t > 0.^{***} \tag{2.1'}$$

Without loss of generality we may assume that $a < 0$, so $0 \in \rho(A)$, and fractional powers A^Q, $0 \in Q \leqslant 1$, are well defined. We shall consider the following abstract model:

$$\begin{cases} \dot{y}(t) = -Ay(t) + A^Q Bu(t), \\ y(0) = 0, \end{cases} \tag{2.2}$$

where $B \in \mathscr{L}(U \to H)$, with U another Banach space.

Remark. Since $\mathscr{D}(A^Q B)$ (considered on H) may be empty, Eq. (2.2) should be understood in a sense of $(\mathscr{D}(A^*))'$-topology. Analyticity of the semigroup $S(t)$ will guarantee that for any $u \in L_p[0T; U]$ there exists a solution $y(t)$ defined a.e. for $t \in [0, T]$. More precisely, let

$$L\colon L_2[0T; U] \to L_2[0T; H]$$

be defined as

$$(Lu)(t) \equiv \int_0^t S(t-\tau)A^Q Bu(\tau)\,d\tau.$$

Clearly, L is densely defined as $H^1[0T; U] \subset \mathscr{D}(L)$. Moreover, the following result holds.

THEOREM 2.1. *Let* $0 \leqslant Q \leqslant 1$. *Then*

(i) $L \in \mathscr{L}(L_\infty[0T; U] \to \mathrm{BMO}[0T; \mathscr{D}(A^{1-Q})])$,†

(ii) $L \in \mathscr{L}(L_1[0T; U] \to L^1_*[0T; \mathscr{D}(A^{1-Q})])$,†

(iii) $L \in \mathscr{L}(L_p[0T; U] \to L_p[0T; \mathscr{D}(A^{1-Q-\varepsilon})])$, $p = 1, \infty$, $\varepsilon > 0$.

***C stands for a generic constant.

†For the definition of BMO (bounded mean oscillations) and L^1_* (L_1-weak) spaces we refer to [15], [8] and [18].

Interpolating between BMO$[0T; \mathscr{D}(A^{1-\varrho})]$ and $L^1_*[0T; \mathscr{D}(A^{1-\varrho})]$ (see [18]) we obtain

COROLLARY 2.2.

$$L \in \mathscr{L}\big(L_p[0T; U] \to L_p[0T; \mathscr{D}(A^{1-\varrho})]\big), \qquad 1 < p < \infty. \tag{2.3}$$

Remark. The result stated in Corollary 2.2 was proved in [20] and [4] in the case when H is a Hilbert space [4] or H is a reflexive Banach space [20]. Our proof, although similar in spirit to [20] and [4], is, however, more general and technically different. Also, elements of this proof will be used later in treating approximation problems.

Proof of Theorem 2.1. In order to prove Theorem 2.1 we shall need some results from the theory of singular integrals. For the convenience of the reader we shall state them below.

THEOREM 2.3. (S. SPANNE–E. STEIN [22]). *Let* $Tf = K * f$ *where* $(K * f)(t) \equiv \int_{-\infty}^{\infty} K(t-\tau) f(\tau)\, d\tau$ *and* $K \in \mathscr{L}(U \to L_1[R, H])$ *satisfies the following properties*:
There exists a constant $M > 0$ (*independent of* $\|K\|_{U \to L_1[R, H]}$) *such that*

$$|\hat{K}(\lambda) u|_H \leqslant M |u|_U, \qquad \lambda = \beta,\ \beta \in R, \tag{2.4}$$

$$\int_{|x| \geqslant 2|y|} \|K(x-y) - K(x)\|_{U \to H}\, dx \leqslant M \quad \forall y \neq 0 \in R. \tag{2.5}$$

Then T maps $L^\infty(R; U)$ *into* BMO$(R; H)$, *and the following inequality holds*,

$$|Tf|_{\mathrm{BMO}(R;\,H)} \leqslant CM |f|_{L^\infty(R;\,U)},$$

where CM does not depend on $\|K\|_{U \to L_1[R;\,H]}$.

Remark. Theorem 2.3 was originally proved in [21] under the assumption that $H = R^n$. Analysis of the proof in [21] reveals, however, that the generalization of the result to vector-valued functions represents no extra difficulties.

In order to apply Theorem 2.3 to our situation, we define

$$K_\delta(t) \stackrel{\Delta}{=} \begin{cases} AS(t)B, & t \geqslant \delta, \\ 0, & t < \delta. \end{cases} \tag{2.6}$$

It can be easily verified that for all $u \in L_1[0T; U]$ and extended by zero outside $[0, T]$, we have

$$(T_\delta u)(t) \stackrel{\Delta}{=} \int_{-\infty}^{+\infty} K_\delta(t-\tau) u(\tau)\, d\tau = \int_0^{t-\delta} AS(t-\tau) B u(\tau)\, d\tau$$

and

$$T_\delta u \underset{\delta \to 0}{\rightarrow} A^{1-\varrho} L u \quad \text{in } C[0T; H] \text{ for all } u \in H^1[0T; U]. \tag{2.7}$$

Also, for each $\delta > 0$, the kernel $K_\delta(t) \in \mathscr{L}(U \to L_1[R; H])$. This follows from the analyticity of the semigroup $S(t)$, which, in particular, implies that $AS(t)$: $H \to H$ is bounded for all $t \geqslant \delta$, and from the fact that $|AS(t)x|_{H \to H} \leqslant Ce^{-wt}$ for t large, $w > 0$ (since we assume without loss of generality that a in the definition of Σ is negative).[††] Now we shall verify assumptions (2.4) and (2.5). As for (2.4), we have

$$\hat{K}_\delta(\beta)u = \int_{-\infty}^{+\infty} e^{-i\beta t} K_\delta(t)u\,dt = \int_\delta^\infty e^{-i\beta t} AS(t)Bu\,dt.$$

From the analyticity of the semigroup, it follows that we can shift the path of integration from the positive real axis to a ray $te^{i\phi}$ starting at the origin. If $\beta > 0$ we choose $\phi < 0$ and obtain

$$\begin{aligned}\hat{K}_\delta(\beta)u &= -e^{i\beta\delta}S(\delta)Bu + i\beta\int_\delta^\infty e^{-i\beta t}S(t)Bu\,dt\\ &= -e^{-i\beta\delta}S(\delta)Bu + i\beta\int_\delta^\infty e^{-i\beta t\cos\phi+\beta t\sin\phi}S(te^{i\phi})Bue^{i\phi}\,dt.\end{aligned}$$

Hence,

$$|\hat{K}_\delta(\beta)u|_H \leqslant C\left[|u|_U + \beta\frac{e^{+\beta\delta\sin\phi}}{\beta}|u|_U\right] \leqslant C|u|_U.$$

For $\beta < 0$ we choose the ray of integration with $\phi > 0$ and obtain a similar estimate. Therefore,

$$\|\hat{K}_\delta(\beta)\|_{U\to H} \leqslant M \quad \textit{uniformly} \text{ in } \delta > 0. \tag{2.8}$$

Next, we shall show that (2.5) is verified for $K_\delta(t)$. To this end, let us write with $y > 0$,

$$\int_{|x|\geqslant 2y} \|K_\delta(x-y) - K_\delta(x)\|_{U\to H}\,dx =$$

$$\begin{cases} ① \displaystyle\int_{2y}^\infty \|(AS(x-y)-AS(x))B\|_{U\to H}\,dx, & y \geqslant \delta,\\ ② \displaystyle\int_{2y}^{y+\delta} \|AS(x)B\|_{U\to H}\,dx + \int_{y+\delta}^\infty \|(AS(x-y)-AS(x))B\|\,dx, & \delta/2 \leqslant y \leqslant \delta,\\ ③ \displaystyle\int_{\delta}^{y+\delta} \|AS(x)B\|_{U\to H}\,dx + \int_{y+\sigma}^\infty \|(AS(x-y)-AS(x))B\|_{U\to H}\,dx, & 0 \leqslant y \leqslant \delta/2.\end{cases}$$

As for the term ①, we shall use

$$\begin{aligned}\int_{2y}^\infty \|(AS(x-y)-AS(x))f\|_{H\to H}\,dx &\leqslant \int_{2y}^\infty \sup_{\substack{f\in D(A^2)\\ |f|_H=1}} \left|A\int_x^{x-y}\frac{d}{d\tau}S(\tau)f\,d\tau\right|_H dx\\ &\leqslant \int_{2y}^\infty \sup_{\substack{f\in D(A^2)\\ |f|_H=1}} \left|\int_x^{x-y} A^2S(\tau)f\,d\tau\right|_H dx\end{aligned}$$

[††] Regularity of the map L clearly will not be affected by translating the spectrum of A to the left of the complex plane.

(by (2.1′) applied with $n = 2$)

$$\leqslant C\int_{2y}^{\infty} \sup_{\substack{f\in D(A^2)\\ |f|_H=1}} \int_x^{x-y} \frac{1}{\tau^2}|f|_H\,d\tau\,dx \leqslant C\int_{2y}^{\infty}\left|\frac{1}{x-y}-\frac{1}{x}\right|dx \leqslant C\ln 2.$$

The same line of arguments applies to the second integrals in the terms ② and ③. As for the first integral in ② and ③, we simply use (2.1′) with $n = 1$. Repeating the same estimates with $y < 0$ yields

$$\int_{|x|\geqslant -2y}^{\infty} \|K_\delta(x-y) - K_\delta(x)\|_{U\to H}\,dx \leqslant M \quad \text{uniformly in } \delta. \tag{2.9}$$

In view of Theorem 2.3, (2.8) and (2.9), we have

$$|T_\delta u|_{\mathrm{BMO}(R,H)} \leqslant C|u|_{L^\infty(R,U)} \quad \text{uniformly in } \delta. \tag{2.10}$$

(2.7), (2.10) and standard density argument yield

$$\left|A^{1-\varrho}Lu\right|_{\mathrm{BMO}(0,T;H)} \leqslant C|u|_{L^\infty[0T;U]},$$

which completes part (i) of Theorem 2.1.

As for part (ii), we shall use the Theorem of Schwartz, which can be stated as follows:

THEOREM 2.4. [21]. *With $Tf = K * f$ as in Theorem 2.3, assume that (2.4) is satisfied. Moreover, assume that there exist a real number $a > 1$ and a constant $M < \infty$ such that for all $\mu > 0$*

$$\int_{|x|\geqslant a} \|K(\mu(x-y) - K(\mu x)\|\,dx \leqslant \frac{M}{\mu} \tag{2.11}$$

for all $|y| < 1/a$.

Then T: $L_1[R;U] \to L_1^*[R,H]$ and

$$|Tf|_{L_1^*[R,H]} \leqslant \frac{M}{a}|f|_{L_1[R,U]}. \tag{2.12}$$

In view of (2.12), to prove part (ii) of Theorem 2.1, it is enough to verify (2.11) with $K_\delta(t)$, where M should be *independent* on δ. To accomplish this, we shall use Corollary 5 in [21]. According to this corollary the sufficient condition for (2.11) to hold is

$$\int_{|x|\geqslant a\mu} \left\|\frac{\partial}{\partial x}K_\delta(x)\right\|_{U\to H} dx \leqslant \frac{M}{\mu}. \tag{2.11′}$$

To check (2.11′), it suffices to write

$$\int_{|x|\geqslant a\mu} \left\|\frac{\partial}{\partial x}K_\delta(x)\right\|_{U\to H} dx = \begin{cases} ①\displaystyle\int_{a\mu}^{\infty} \|A^2S(x)B\|_{U\to H}\,dx, & a\mu \geqslant \delta, \\ ②\displaystyle\int_{\delta}^{\infty} \|A^2S(x)B\|_{U\to H}\,dx, & a\mu \leqslant \delta. \end{cases}$$

To estimate ① and ②, we shall use analyticity of the semigroup $S(t)$. In fact,

$$① \leqslant C\int_{a\mu}^{\infty} \frac{1}{x^2}\,dx \leqslant \frac{C}{a\mu} \leqslant \frac{C}{\mu} \quad \text{for any } a \geqslant 1.$$

Similarly,

$$②\leqslant C\int_{\delta}^{\infty}\frac{1}{x^2}\,dx\leqslant\frac{C}{\delta}\leqslant\frac{C}{\mu}.$$

Thus, (2.11′), and consequently (2.11), are satisfied. From Theorem 2.4 we now obtain

$$(2.13)\qquad |T_\delta f|_{L_1^*[R,H]}\leqslant C|f|_{L_1[R;U]}\quad \textit{uniformly in } \delta>0.$$

Equation (2.13) together with (2.7) imply that

$$|A^{1-Q}Lu|_{L_1^*[0T;H]}\leqslant C|u|_{L_1[0T;U]},$$

which completes the proof of part (ii) of Theorem 2.1. Proof of part (iii) is straightforward. In fact,

$$|(A^{1-Q-\varepsilon}Lu)(t)|_H=\left|\int_0^t A^{1-\varepsilon}S(t-\tau)Bu(\tau)\,d\tau\right|_H$$

(by (2.1′) applied with $n=1-\varepsilon$)

$$\leqslant C_T\int_0^t\frac{1}{(t-\tau)^{1-\varepsilon}}\,d\tau\,|u|_{L_\infty[0T;U]}\leqslant C_T|u|_{L_\infty[0T;U]}.$$

Similarly, for $p=1$ we have

$$\int_0^T|A^{1-Q-\varepsilon}Lu(t)|_H\,dt\leqslant C\int_0^T\int_0^t\frac{1}{(t-\tau)^{1-\varepsilon}}|u(\tau)|_U\,d\tau\,dt$$

(changing the order of integration)

$$\leqslant C_T\int_0^T\int_\tau^T\frac{1}{(t-\tau)^{1-\varepsilon}}\,dt\,|u(\tau)|_U\,d\tau\leqslant C_T|u|_{L_1[0T;U]}.$$

The proof of Theorem 2.1 is thus completed. □

3. Galerkin Approximation of an Abstract Boundary Value Problem. Let $0<h<1$, $h\to 0$ be a parameter of discretization. Let $V_h\subset\mathscr{D}(B^*A^{*Q})$ be a family of finite-dimensional subspaces of H. We shall assume that the subspaces V_h enjoy the following approximation properties. There exists a constant $m\geqslant 0$ such that

$$(3.0)\qquad |x-P_hx|_H\leqslant Ch^{m\alpha}|x|_{D(A^\alpha)},\qquad 0\leqslant\alpha\leqslant 1,$$

where P_h stands for the orthogonal projection of H onto V_h,

$$(3.1)\qquad |B^*A^{*Q}(I-P_h)x|_U\leqslant Ch^{m(1-\bar{Q})}|x|_{D(A^*)}$$

for some $\bar{Q}\leqslant Q$,

$$(3.1')\qquad |B^*A^{*Q}(I-P_h)x|_H\leqslant Ch^{m(\alpha-Q)}|x|_{D(A^{*\alpha})}\quad\text{for } Q\leqslant\alpha\leqslant 1,$$

$$(3.2)\qquad |B^*A^{*Q}V_h|_U\leqslant Ch^{-m\bar{Q}}|v_h|_H\quad\text{(inverse approximation property).}$$

Remark. Since $B\in\mathscr{L}(U\to H)$, properties (3.1) and (3.2) follow from

$$(3.1'')\qquad |(I-P_h)x|_{D(A^{*Q})}\leqslant Ch^{m(1-\bar{Q})}|x|_{D(A^*)},$$

$$(3.2'')\qquad |v_h|_{D(A^{*Q})}\leqslant Ch^{-m\bar{Q}}|v_h|_H.$$

In the special case when A represents a differential operator of order m and such that for $x \in D(A)$ we have $|x|_{D(A)} \sim \|x\|_m$, then the properties (3.0), (3.1), (3.1″) and (3.2″) (with $\overline{Q} = Q$) are the standard approximation requirements satisfied by the spaces of splines defined on uniform meshes. Consequently, (3.0)–(3.2) with $\overline{Q} = Q$ are the *weaker* versions of the standard approximation properties, weaker in the sense that (3.1″) and (3.2″) require *a priori* that $V_h \subset D(A^{*Q})$, while (3.1) and (3.2) require only that $V_h \subset D(B^*A^{*Q})$. This fact will be crucially used in Section 5, where our general theory is applied to a parabolic boundary value problem. As we shall see, when working with linear splines, we shall have $V_h \subset D(B^*A^{*Q})$ *but not* in $D(A^{*Q})$.

The generator A will be approximated by the sequence of finite-dimensional operators A_h: $V_h \to V_h$ satisfying the following properties:

$$(3.3) \qquad \left|\left(A^{-1} - A_h^{-1}\right) R_h x\right|_H \leqslant Ch^m |x|_H \quad \text{(convergence)};$$

with $S_h(t) = e^{A_h t}$,

$$(3.4) \qquad \left|A_h^\beta S_h(t)\right|_{H \to H} \leqslant \frac{C_\beta}{t^\beta}, \qquad 0 \leqslant \beta \leqslant 2, \text{ uniformly in } h$$

(uniform analyticity).

We shall also assume that the properties (3.1)–(3.4) hold for A^*.

Remark. If A is a selfadjoint strongly elliptic operator of order m with appropriate boundary conditions, then A_h defined by

$$(3.5) \qquad \left(A_h y_h, x_h\right)_H \equiv \left(A y_h, x_h\right)_H \quad \forall y_h,\ x_h \in V_h \subset D(A^{1/2}),$$

complies with the requirements (3.3), (3.4) (see [9]). Similarly, if A is coercive in the norm of $\mathscr{D}(A^{1/2})$, which is the case in the parabolic situation, then A_h defined by (3.5) also satisfies (3.3) and (3.4).

An equivalent version of (3.4) is that there exist $a, b > 0$, $C > 0$ such that

$$(3.4') \qquad \rho(A_h) \supset \Sigma \equiv \{\lambda; \operatorname{Re}\lambda > a - b \operatorname{Im}|\lambda|\}$$

and

$$\|R(\lambda, A_h)\| \leqslant \frac{C}{(|\lambda| + 1)} \quad \textit{uniformly in } h.$$

Let $y(t)$ be the solution of (1.2). The *approximate schemes* we shall consider are as follows: Find $y_h(t) \in V_h$ such that

$$(3.6) \qquad \begin{cases} \left(\dot{y}_h(T), v_h\right)_H = -\left(A_h y_h(t), v_h\right)_H + \left(u(t), B^*A^{*Q} v_h\right)_U \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for every } v_h \in V_h, \\ y_h(0) = 0. \end{cases}$$

Let $e(t) \equiv y(t) - y_h(t)$. The following theorem gives the error estimates for the error function e in the schemes defined above.

Theorem 3.1.
(i) $|e|_{L^p[0T;\, H]} \leqslant Ch^{m(1-\overline{Q})} |u|_{L^p[0T;\, U]}$ *for* $1 < p < \infty$,
(ii) $|e|_{L^p[0T;\, H]} \leqslant C[\ln h h^{m(1-\overline{Q})} + h^{m(1-Q+\varepsilon)}] |u|_{L^p[0T;\, U]}$ *for* $p = 1, \infty$.

Remark. The convergence results obtained in Theorem 3.1 are *optimal*. In fact, in view of the optimal regularity of the solution y, with $u \in L_p[0T; H]$, $1 < p < \infty$, we have that $y \in L_p[0T; \mathcal{D}(A^{1-Q})]$. Thus, the estimate in part (i) essentially reproduces the approximation property (3.0) (with $\alpha = (1 - \overline{Q})$). Similarly, for $p = 1, \infty$ the maximal regularity of the solution is $y \in L_p[0T; \mathcal{D}(A^{1-Q-\epsilon})]$. This fact is reflected by the presence of the $\ln h$ term in part (ii) of the theorem.

4. Proof of Theorem 3.1. In the proof of Theorem 3.1 the key role will be played in the following "rough data" estimates.

THEOREM 4.1. ([9], [10]). *Let A_h, A_h^* satisfy* (3.3) *and* (3.4). *Then*

$$\|S_h(t)P_h - P_hS(t)\|_{H\to H} \leqslant \frac{C_T}{t^l}h^{ml}, \qquad 0 \leqslant l \leqslant 1, \tag{4.1}$$

$$\|R(\lambda, A) - R(\lambda, A_h)P_h\|_{H\to H} \leqslant Ch^m \quad \textit{for all } \lambda \in \Sigma. \tag{4.2}$$

Theorem 4.1 implies the following

COROLLARY 4.1. $\|AS(t) - A_hS_h(t)P_h\|_{H\to H} \leqslant Ch^m/t^2$.

Proof. With $x \in D(A)$ we have

$$[AS(t) - A_hS_h(t)P_h]x = \int_\Gamma e^{\lambda t}\lambda[R(\lambda, A) - R(\lambda, A_h)P_h]x\,d\lambda,$$

where Γ denotes the contour of Σ. Hence, in view of (4.2),

$$|AS(t) - A_hS_h(t)P_h|_{H\to H} \leqslant Ch^m\int_\Gamma e^{\operatorname{Re}\lambda t}|\lambda|\,d\lambda.$$

Straightforward evaluation of the last integral now yields the desired bound. □

Remark. In the special case when A represents a selfadjoint (resp. slightly nonselfadjoint) second-order strongly elliptic operator, (4.1) in Theorem 4.1 was proved in [3] (resp. [7] and [19] and [13]). In [9] this result was extended to a more general case of analytic semigroups with "uniformly analytic" generator A_h. This is the case, for example, for an arbitrary strongly elliptic operator where the bilinear form $(A_hy_h, x_h)_H$ is coercive in the $D(A^{1/2})$-topology.

We shall start by proving the "easy" part of Theorem 3.1, i.e., part (ii).

With L_h: $L_2[0T; U] \to L_2[0T; U_h]$ given by

$$(L_hu)(t) \equiv \int_0^t S_h(t-\tau)P_h(A^QBu(\tau))\,d\tau,$$

we observe that (3.6) is equivalent to $y_h(t) = (L_hu)(t)$. Thus,

$$e(t) = (L - L_h)(u)(t). \tag{4.3}$$

By taking adjoints L^* and L_h^* to L and L_h (with respect to $L_2[0T; U] \to L_2[0T; H]$ topology) we obtain

$$(L^*f)(t) = \int_t^T B^*A^{*Q}S^*(\tau - t)f(\tau)\,d\tau$$

and

$$(L_h^*f_h)(t) = \int_t^T B^*A^{*Q}S_h^*(\tau - t)f_h(\tau)\,d\tau.$$

Notice that in view of duality, part (ii) of Theorem 3.1 is equivalent to the following estimate

$$
(4.4)\qquad \begin{aligned} &|(L^* - L_h^* P_h f)|_{L_p[0T;U]} \\ &\quad \leqslant C\left[\ln h h^{m(1-\bar{Q})} + h^{m(1-Q+\varepsilon)}\right]|f|_{L_p[0T;H]} \quad \text{for } p = 1, \infty. \end{aligned}
$$

We start by proving (4.4) with $p = \infty$. To this end, let us write

$$
(L^* - L_h^* P_h f)(t) = e_1^*(t) + e_2^*(t),
$$

where

$$
e_1^*(t) = B^* A^{*Q}(I - P_h)\int_t^T S^*(\tau - t) f(\tau)\, d\tau,
$$

$$
e_2^*(t) = B^* A^{*Q}\int_t^T \left[P_h S^*(\tau - t) - S_h^*(\tau - t) P_h\right] f(\tau)\, d\tau.
$$

Consider first $t \leqslant T - h^{2m}$,

$$
\begin{aligned} |e_1^*(t)|_U \leqslant &\left| B^* A^{*Q}(I - P_h)\int_t^{t+h^{2m}} S^*(\tau - t) f(\tau)\, d\tau\right|_U \\ &+ \left| B^* A^{*Q}(I - P_h)\int_{t+h^{2m}}^T S^*(\tau - t) f(\tau)\, d\tau\right|_U. \end{aligned}
$$

Applying to the first term (3.1′) with $\alpha = 1 - \varepsilon$, and (3.1) to the second term, yields the further bounds

$$
\begin{aligned} &C h^{-mQ} h^{m(1-\varepsilon)}\left|\int_t^{t+h^{2m}} A^{*(1-\varepsilon)} S^*(\tau - t) f(\tau)\, d\tau\right|_H \\ &\quad + C h^{-m\bar{Q}} h^m \left|\int_{t+h^{2m}}^T A^* S^*(\tau - t) f(\tau)\, d\tau\right|_H \end{aligned}
$$

(by analyticity of the semigroup $S(t)$)

$$
\begin{aligned} &\leqslant C h^{m(1-Q)} h^{-m\varepsilon}\int_t^{t+h^{2m}} \frac{1}{(\tau - t)^{1-\varepsilon}}\, d\tau |f|_{L^\infty[0T;H]} \\ &\quad + C h^{m(1-\bar{Q})}\int_{t+h^{2m}}^T \frac{1}{(\tau - t)}\, d\tau |f|_{L_\infty[0T;H]}. \end{aligned}
$$

Thus, for each t such that $t \leqslant T - h^{2m}$ we have

$$
(4.5)\qquad |e_1^*(t)|_U \leqslant C\left[h^{m(1-\bar{Q})}\ln h + h^{m(1-Q+\varepsilon)}\right]|f|_{L_\infty[0T;H]}.
$$

Similarly, for $t \geqslant T - h^{2m}$ we have

$$
|e_1^*(t)|_U \leqslant \left| B^* A^{*Q}(I - P_h)\int_t^T S^*(\tau - t) f(\tau)\, d\tau\right|_U
$$

(by (3.1′) applied with $\alpha = 1 - \varepsilon$)

$$
\begin{aligned} &\leqslant C h^{-mQ} h^{m(1-\varepsilon)}\left|\int_t^T A^{*1-\varepsilon} S^*(\tau - t) f(\tau)\, d\tau\right|_H \\ (4.6)\qquad &\leqslant C h^{m(1-Q)} h^{-m\varepsilon}\int_{T-h^{2m}}^T \frac{1}{(\tau - t)^{\varepsilon}}\, d\tau |f|_{L_\infty[0T;H]} \\ &\leqslant C h^{m(1-Q+\varepsilon)}|f|_{L_\infty[0T;H]}. \end{aligned}
$$

Thus, (4.5) and (4.6) give

$$|e_1^*|_{L_\infty[0T;U]} \leqslant C\left[h^{m(1-\bar{Q})}\ln h + h^{m(1-Q+\varepsilon)}\right]|f|_{L_\infty[0T;H]}. \tag{4.7}$$

Similar analysis applies to the term $e_2^*(t)$. In fact, by (3.2) we have

$$|e_2^*(t)|_U \leqslant Ch^{-m\bar{Q}}\int_t^T |P_hS^*(\tau - t) - S_h^*(\tau - t)|_{H\to H}\,d\tau|f|_{L_\infty[0T;H]}.$$

Thus, for $t \leqslant T - h^m$,

$$|e_2^*(t)|_U \leqslant Ch^{-m\bar{Q}}\left[\int_t^{t+h^m} h^{m(1-\varepsilon)}\frac{d\tau}{(\tau - t)^{1-\varepsilon}} + \int_{t+h^m}^T h^m\frac{d\tau}{\tau - t}\right]|f|_{L_\infty},$$

where in the first integral we have applied (4.1) with $l = 1 - \varepsilon$ and in the second with $l = 1$. Hence, for $t \leqslant T - h^m$,

$$|e_2^*(t)|_U \leqslant Ch^{m(1-\bar{Q})}\ln h|f|_{L_\infty[0T;H]}. \tag{4.8}$$

For $t \geqslant T - h^m$, we obtain similarly as before

$$|e_2^*(t)|_U \leqslant Ch^{m(1-\bar{Q})}|f|_{L_\infty[0T;H]}. \tag{4.9}$$

(4.8), (4.9), and (4.7) together yield

$$|e_1^*(t) + e_2^*(t)|_U \leqslant C\left[h^{m(1-\bar{Q})}\ln h + h^{m(1-Q+\varepsilon)}\right]|f|_{L_\infty[0T;H]}, \tag{4.10}$$

which completes the proof of (4.4) for $p = \infty$. Now, we consider case $p = 1$:

$$\begin{aligned}
\int_0^T |e_1^*(t)|_U\,dt &\leqslant \int_0^T\int_0^\tau |B^*A^{*Q}(I - P_h)S^*(\tau - t)f(\tau)|_U\,dt\,d\tau\\
&\leqslant \int_0^{h^{2m}}\int_0^\tau |B^*A^{*Q}(I - P_h)S^*(\tau - t)f(\tau)|_U\,dt\,d\tau\\
&\quad + \int_{h^{2m}}^T\left[\int_0^{\tau-h^{2m}} + \int_{\tau-h^{2m}}^\tau\right] |B^*A^{*Q}(I - P_h)S^*(\tau - t)f(\tau)|_U\,dt\,d\tau\\
&\leqslant Ch^{-mQ+m(1-\varepsilon)}\int_0^{h^{2m}}\int_0^\tau |A^{*1-\varepsilon}S^*(\tau - t)f(\tau)|_H\,dt\,d\tau\\
&\quad + Ch^{-m\bar{Q}+m}\int_{h^{2m}}^T\int_0^{\tau-h^{2m}} |A^{*1-\varepsilon}S^*(\tau - t)f(\tau)|_H\,dt\,d\tau\\
&\quad + Ch^{-mQ+m(1-\varepsilon)}\int_{h^{2m}}^T\int_{\tau-h^{2m}}^\tau |A^{*1-\varepsilon}S^*(\tau - t)f(\tau)|_H\,dt\,d\tau,
\end{aligned} \tag{4.11}$$

where in the last inequality (4.11) we have used (3.1′) applied with $\alpha = 1 - \varepsilon$ to the first and third integral, and (3.1) to the second. From the analyticity of $S^*(t)$ it now follows that

$$\begin{aligned}
\int_0^T |e_1^*(t)|_U\,dt\\
&\leqslant Ch^{m(1-Q)}h^{-m\varepsilon}\left[\int_0^{h^{2m}}\int_0^\tau \frac{dt}{(\tau - t)^{1-\varepsilon}}|f(\tau)|_H\,d\tau\right.\\
&\qquad\left. + \int_{h^{2m}}^T\int_{\tau-h^{2m}}^\tau \frac{dt}{(\tau - t)^{1-\varepsilon}}|f(\tau)|_H\,d\tau\right]\\
&\quad + Ch^{m(1-\bar{Q})}\int_{h^{2m}}^T\int_0^{\tau-h^{2m}} \frac{dt}{\tau - t}|f(\tau)|\,d\tau.
\end{aligned} \tag{4.12}$$

Hence,

$$\int_0^T |e_1^*(t)|_U dt \leqslant C\left[h^{m(1-\bar{Q})}\ln h + h^{m(1-Q+\varepsilon)}\right]|f|_{L_\infty[0T;H]}. \tag{4.13}$$

To show that the estimate (4.13) also holds for $e_2^*(t)$, we follow the same line of arguments as before. The only difference now is that instead of using (3.1′) we use (3.2) combined with (4.1), applied with $l = 1 - \varepsilon$ and $l = 1$ (see also the proof of (4.9)). The proof of (4.4) is thus *completed.*

Proceeding with the proof of Theorem 3.1 we shall now prove part (i). To accomplish this we shall need the following dual formulations of the assumptions (3.1) and (3.2).

COROLLARY 4.2. (i) $|A^{-1}(I - P_h)A^Q Bu|_H \leqslant Ch^{m(1-\bar{Q})}|u|_U$,
(ii) $|P_h A^Q Bu|_H \leqslant Ch^{-m\bar{Q}}|u|_U$.

To continue *with the proof* of our Theorem 3.1 we write

$$\begin{aligned} e(t) = (L - L_h)u(t) &= \int_0^t \underbrace{(S(t-\tau) - S_h(t-\tau))P_h\left(A^Q Bu(\tau)\right) d\tau}_{① \equiv e_1(t)} \\ &\quad + \int_0^t \underbrace{AS(t-\tau)A^{-1}(I - P_h)(A^Q Bu)(\tau)\, d\tau}_{② \equiv e_2(t)}. \end{aligned} \tag{4.14}$$

To estimate the term $e_2(t)$, we use Corollary 2.2 with $Q = 1$ and $B = I$. This yields

$$|e_2(t)|_{L_p[0T;H]} \leqslant C_T\left|A^{-1}(I - P_h)A^Q Bu\right|_{L_p[0T;H]}$$

(by Corollary 4.2(i))

$$\leqslant C_T h^{m(1-\bar{Q})}\left|A^{-Q}A^Q Bu\right|_{L_p[0T;H]} \leqslant C_T h^{m(1-\bar{Q})}|u|_{L_p[0T;U]}. \tag{4.15}$$

Thus, in order to complete the proof, it is enough to show that the same estimate holds for $e_1(t) = (E_h u)(t)$, where we introduced the notation

$$(E_h u)(t) \equiv \int_0^t [S(t-\tau) - S_h(t-\tau)]\, P_h\left[A^Q Bu(\tau)\right] d\tau. \tag{4.16}$$

We shall prove

LEMMA 4.3.

$$|E_h u|_{\mathrm{BMO}[0T;H]} \leqslant C_T h^{m(1-\bar{Q})}|u|_{L^\infty[0T;U]}.$$

LEMMA 4.4.

$$|E_h u|_{L_1^*[0T;H]} \leqslant C_T h^{m(1-\bar{Q})}|u|_{L^1[0T;U]}.$$

Once Lemmas 4.3 and 4.4 are proved, the result of part (i) of Theorem 3.1 follows immediately by interpolating the results of Lemmas 4.3 and 4.4 (see [18]). In fact, we obtain

$$|E_h u|_{L_p[0T;H]} \leqslant C_{T,p} h^{m(1-\bar{Q})}|u|_{L^p[0T;H]},$$

which together with (4.15) and (4.14) gives

$$|e(t)|_{L_p[0T;H]} \leqslant C_{T,p} h^{m(1-\bar{Q})}|u|_{L_p[0T;H]}, \qquad p \neq 1, \infty.$$

Thus, to complete the proof of the theorem, it is enough to establish the validity of Lemmas 4.3 and 4.4.

Proof of Lemma 4.3. We apply Theorem 2.3 with

$$K_h(t) \triangleq \begin{cases} [S(t) - S_h(t)] P_h[A^Q Bu], & t \geq 0, \\ 0, & t < 0. \end{cases}$$

Clearly, for $h > 0$, $K_h \in \mathscr{L}(U \to L_1[0T; H])$, and with $u \in L_1[0T; U]$ and $u(t) = 0$, $t < 0$, we have

$$(E_h u)(t) = K_h * u.$$

In order to apply Theorem 2.3 we need to verify that (2.4) and (2.5) are valid. To see this, let us compute

$$\begin{aligned} \hat{K}_h(\lambda) u &= [R(\lambda, A) - R(\lambda, A_h)] P_h(A^Q Bu) \\ &= [R(\lambda, A) - R(\lambda, A_h) P_h] P_h(A^Q Bu). \end{aligned}$$

Hence,

$$|\hat{K}_h(\lambda) u|_H \leq |R(\lambda, A) - R(\lambda, A_h) P_h|_{H \to H} |P_h(A^Q Bu)|_H$$

by (4.2)

$$\leq C h^m |P_h(A^Q Bu)|_H$$

(by part (ii) of Corollary 4.2)

$$\leq C h^{m(1-\bar{Q})} |u|_U.$$

Thus,

$$|\hat{K}_h(\lambda)|_{U \to H} \leq C h^{m(1-\bar{Q})}. \tag{4.17}$$

Next, we establish the validity of (2.5). In fact,

$$\begin{aligned} K_h(x-y)u - K(x)u = {} & [S(x-y) - S(x)] P_h(A^Q Bu) \\ & -(S_h(x-y) - S_h(x)) P_h(A^Q Bu) \end{aligned}$$

(by the semigroup property)

$$\begin{aligned} &= \int_x^{x-y} AS(\tau) P_h(A^Q Bu)\, d\tau - \int_x^{x-y} A_h S_h(\tau) P_h(A^Q Bu)\, d\tau \\ &= \int_x^{x-y} |AS(\tau) - A_h S_h(\tau)| P_h(A^Q Bu)\, d\tau. \end{aligned}$$

Thus, in view of Corollary 4.1 we have

$$|K_h(x-y)u - K_x(x)(u)|_H \leq C \int_x^{x-y} \frac{1}{\tau^2}\, d\tau h^m P_h(A^Q Bu)$$

(by Corollary 4.2(ii))

$$\leq C \left| \frac{1}{x-y} - \frac{1}{x} \right| h^{m(1-\bar{Q})} |u|_U. \tag{4.18}$$

Therefore (with $y \geq 0$),

$$\begin{aligned} \int_{|x| \geq 2|y|} |K(x-y) - K(x)|_{U \to H}\, dx &\leq C h^{m(1-\bar{Q})} \int_{x=2y}^{\infty} \left(\frac{1}{x-y} - \frac{1}{x} \right) dx \\ &\leq C h^{m(1-\bar{Q})} \ln 2. \end{aligned} \tag{4.19}$$

The same argument gives an estimate for $y < 0$. Thus, (4.17), (4.19) and Theorem 2.3 yield

$$|K_h * u|_{\mathrm{BMO}[0T;\, H]} \leqslant C_T h^{m(1-\bar{Q})} |u|_{L^\infty[0T;\, U]},$$

which completes the proof of Lemma 4.3. □

Proof of Lemma 4.4. The proof of this lemma relies on the application of Schwartz's Theorem 2.4. All we need to check is the validity of (2.11), with $K_h(t)$ defined as above. The computations, in fact, are also very similar to those before. Set $a = 2$ in Theorem 2.4. Then, as in (4.18),

$$\|K_h(\mu(x-y)) - K_h(\mu x)\|_{U \to H} \leqslant \frac{C h^{m(1-\bar{Q})}}{\mu} \left| \frac{1}{x-y} - \frac{1}{x} \right|$$

(first with $x > 0$)

$$\begin{aligned} &\int_{|x| \geqslant 2} \|K_h(\mu(x-y)) - K_h(\mu x)\|_{U \to H}\, dx \\ &\quad = \int_2^\infty \|K_h(\mu(x-y)) - K_h(\mu x)\|_{U \to H}\, dx \\ &\quad \leqslant \frac{c}{\mu} h^{m(1-\bar{Q})} \int_2^\infty \left[\frac{1}{x-y} + \frac{1}{x} \right] dx = \frac{c}{\mu} h^{m(1-\bar{Q})} \ln \frac{3}{4}. \end{aligned} \tag{4.20}$$

Similar computations apply to the case $x \leqslant 0$. Thus, (4.20) and Theorem 3.2 imply that

$$\|K_h * u\|_{L_1^*[0T;\, H]} \leqslant C h^{m(1-\bar{Q})} |u|_{L_1[0T;\, U]},$$

which is precisely the statement of Lemma 4.4. □

The proof of Theorem 3.1 is now completed.

5. Applications to Parabolic Boundary Value Problems. The purpose of this section is to show how the abstract approximation result formulated in Theorem 3.1 can be applied to yield the optimal rate of convergence of Galerkin approximation to parabolic problems with "rough" (i.e., $L_p(L_2)$) Dirichlet boundary data. To begin with, let Ω be a bounded domain in R^n with a smooth boundary Γ. Let $A(x, \partial)$ denote a uniformly strongly elliptic operator

$$A(x, \partial) f \equiv \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(a_{ij}(x) \frac{\partial f}{\partial x_j} \right) + \sum_{i=1}^n a_i(x) \frac{\partial f}{\partial x_i} + a_0 f(x),$$

where

$$\sum_{i=1}^n \sum_{j=1}^n a_{ij}(x) \xi_i \xi_j \geqslant \alpha \sum_{i=1}^n \xi_i^2, \qquad \alpha > 0,$$

for all $x \in \Omega$, and all coefficients are assumed to be in $C^\infty(\Omega)$. Consider the following parabolic equation:

$$\begin{cases} \dfrac{\partial y(x,t)}{\partial t} = A(x, \partial) y(x,t), \\ y(0) = 0, \\ y|_\Gamma = u \in L_p[0T;\, L_2(\Gamma)], \qquad 1 \leqslant p < \infty. \end{cases} \tag{5.1}$$

In order to express the solution of the parabolic problem (5.1) as a singular integral (as in Section 3), we introduce an operator A: $L_2(\Omega) \to L_2(\Omega)$ defined by

$$-Ay \equiv A(x,\partial)y, \quad y \in \mathscr{D}(A),$$

where

$$\mathscr{D}(A) = \{ y \in L_2(\Omega); Ay \in L_2(\Omega); y|_\Gamma = 0 \}.$$

It is well known that $-A$ generates an analytic semigroup $S(t)$ on $L_2(\Omega)$. Let us now define the "Dirichlet" map D: $L_2(\Gamma) \to L_2(\Omega)$, where

$$\begin{cases} A(x,\partial)\,Dg = 0, \\ Dg|_\Gamma = 0. \end{cases} \tag{5.2}$$

It is well known [14] that

$$D\colon H^s(\Gamma) \to H^{s+1/2}(\Omega) \quad \text{is bounded for all real } s. \tag{5.3}$$

An abstract version of (5.1) can now be expressed by the following formula (see [2], [11]):

$$y(t) = \int_0^t AS(t-\tau)\,Du(\tau)\,d\tau. \tag{5.4}$$

Knowing that

$$\mathscr{D}(A^\alpha) = H^{2\alpha}(\Omega), \quad 0 \leqslant \alpha < 1/4, \quad [6],$$

we have

$$D\colon D \in \mathscr{L}\big(L_2(\Gamma) \to \mathscr{D}(A^{1/4-\varepsilon})\big) \quad \text{for every } \varepsilon > 0.$$

Therefore, (5.4) can be rewritten equivalently as

$$y(t) = \int_0^t A^{3/4-\varepsilon} S(t-\tau) A^{1/4-\varepsilon} Du(\tau)\,d\tau. \tag{5.5}$$

Thus, we are now exactly in the situation described in Section 1, with $H \equiv L_2(\Omega)$, $U \equiv L_2(\Gamma)$, $B \equiv A^{1/4-\varepsilon}D \in \mathscr{L}(L_2(\Gamma) \to L_2(\Omega)]$, $Q = 3/4 + \varepsilon$ and $m = 2$.

In order to formulate an approximating scheme for (5.1) (or equivalently (5.5)), we introduce finite-dimensional subspaces $V_h \subset \mathscr{D}(D^*A^*)$.

Since

$$D^*A^*v = \frac{\partial}{\partial\eta_A} v|_\Gamma \quad (\text{see } [2], [12]),$$

clearly $H^{3/2+\varepsilon}(\Omega) \subset \mathscr{D}(D^*A^*)$.

Remark. Linear splines, although they *are not* in $H^{3/2+\varepsilon}(\Omega)$, still belong to $\mathscr{D}(D^*A^*)$. This fact will be used in the sequel.

Using the identifications [6]

$$\mathscr{D}(A^Q) = \mathscr{D}(A^{*Q}) = H_0^{2Q}(\Omega), \qquad 0 \leqslant Q < 3/4,$$
$$\mathscr{D}(A) = \mathscr{D}(A^*) = H_0^1(\Omega) \cap H^2(\Omega),$$

and

$$|x|_{D(A^\beta)} \leqslant C\|x\|_{2\beta} \quad \text{for } x \in D(A^\beta),$$

one can easily check that (3.0)–(3.2) with $\overline{Q} = Q$ are equivalent to the well-known approximation properties of spaces of linear (and higher-order) splines defined on uniform meshes.

As an approximation A_h of A we take an arbitrary operator A_h: $V_h \to V_h$ such that

$$(5.6) \qquad \begin{cases} \|(A^{-1} - A_h^{-1}R_h)x\| \leqslant Ch^2\|x\|, \\ \|A_h^{\beta}v_h\| \leqslant C\|A^{\beta}v_h\|, \qquad 0 \leqslant \beta \leqslant 1/2, \end{cases}$$

and

$$(5.7) \qquad |(A_h u_h, v_h)| \leqslant C\|u_h\|_1\|v_h\|_1, \qquad (A_h u_h, u_h) \geqslant \gamma\|u_h\|_1^2.$$

It was shown in [9] and [7] that most of the well-known approximations A_h to elliptic problems comply with (5.6) and (5.7). For instance, the standard Galerkin method, where

$$(5.8) \qquad (A_h u_h, v_h) \equiv (A(x,\partial)u_h, v_h) \quad \forall u_h, v_h \in V_h,$$

satisfies all the desired properties (5.6) and (5.7). Also, Babuška's method [1] and Nitsche's method [16] of approximating A_h (which, in fact, do not require subspaces V_h to satisfy zero boundary conditions) can be used.

It was also shown in [9] that with A_h satisfying (5.6) and (5.7), the corresponding semigroup $S_h(t)$ is *uniformly* analytic, i.e.,

$$\|A_h^{\beta}S_h(t)x_h\| \leqslant \frac{c}{t^{\beta}}\|x_h\| \quad \text{uniformly in } h.$$

Consequently, the following "rough" data estimates hold [9]:

$$(5.9) \qquad \|(S_h(t)P_h - P_hS(t))x\| \leqslant \frac{C_T}{t^l}h^{2l}\|x\|, \qquad 0 \leqslant l \leqslant 1,$$

$$(5.10) \qquad \|(R(\lambda, A_h)P_h - R(\lambda, A))x\| \leqslant Ch^2\|x\|, \qquad \lambda \in \Sigma.$$

Remark. Estimates (5.9) were also proved in [7], [19], and [13] for the case where $a_{ij} = a_{ji}(x)$ in the definition of $A(x, \partial)$.

Thus, we are exactly in the situation described in Section 3, where approximation assumptions (3.3) and (3.4) were satisfied with $m = 2$. The algorithm for computing the approximation of y takes the form:

Find $y_h(t) \in V_h$ such that

$$(5.11) \qquad \begin{cases} (\dot{y}_h(t), v_h)_\Omega = -(A_h y_h(t), v_h)_\Omega + \langle u(t), D^*A^*v_h\rangle_\Gamma, \\ y_h(0) = 0, \quad \textit{all } v_h \in V_h. \end{cases}$$

Noticing that (see [12], [2])

$$(5.12) \qquad D^*A^*v = \frac{\partial}{\partial\eta_A}v,$$

where $\partial/\partial\eta_A = a_{ij}(x)n_i(x)\partial/\partial x_j$ and n_i are the components of the outward unit vector normal to the boundary Γ, we can rewrite (5.11) as

$$(5.11') \quad (\dot{y}_h(t), v_h)_\Omega = -(A_h y_h(t), v_h)_\Omega + \left\langle u(t), \frac{\partial}{\partial\eta A}v_h\right\rangle_\Gamma, \qquad y_h(0) = 0.$$

Remark. If one takes for A_h a standard Galerkin approximation (see (5.8)) then (5.11′) becomes

$$(\dot{y}_h(t), v_h)_\Omega = -(A(x,\partial)y_h(t), v_h)_\Omega + \left\langle u(t), \frac{\partial}{\partial\eta A}v_h\right\rangle_\Gamma \quad \text{all } v_h \in V_h.$$

Remark. One should note that (5.11) does not require the elements of V_h to be conformal (as long as A_h is appropriately defined on nonconformal subspaces).

Applying the result of Theorem 3.1 to our scheme (5.11), and taking $\overline{Q} = Q$ in (3.1) and (3.2), yields the following error estimates,

$$(5.13) \qquad |e(t)|_{L_p[0T;\, L_2(\Omega)]} \leq C_T h^{1/2-\varepsilon} |u|_{L_p[0T;\, L_2(\Gamma)]}, \qquad 1 < p < \infty,$$

with $e(t) \equiv y(t) - y_h(t)$, where $y(t)$ (resp. $y_h(t)$) is the solution of (5.1) (resp. (5.11)). Note that the error estimates (5.13) are nonoptimal (modulo h^{ε}). In fact, parabolic theory provides us with the following regularity results

$$(5.14) \quad \begin{cases} \text{for } u \in L_p[0T; L_2(\Gamma)] \text{ we have } \begin{cases} y \in L_p[0T; H^{1/2}(\Omega)], & 1 < p < \infty, \\ y \in L_p[0T; H^{1/2-\varepsilon}(\Omega)], & p = 1, \infty. \end{cases} \end{cases}$$

Therefore, the optimal rate of convergence for $e(t)$ should be $\mathcal{O}(h^{1/2})$ for $1 < p < \infty$. The nonoptimality of the estimate (5.13) is the result of a nonoptimal identification of B with $A^{1/4-\varepsilon}D$ (we cannot take $\varepsilon = 0$!). On the other hand, it is known that $R(D) \subset H^{1/2}(\Omega)$ (but *not* in $\mathcal{D}(A^{1/4})$!). We will be able to improve the estimate (5.13), by imposing slightly stronger requirements on approximation properties of V_h. More precisely we have the following result.

THEOREM 5.1. *Let $V_h \subset H^1(\Omega)$ be such that*

$$(5.15) \quad \begin{array}{lll} \text{(a)} & \|v - P_h v\|_s \leq C h^{\alpha - s} \|v\|_\alpha, & 0 \leq \alpha \leq 2, 0 \leq s \leq 1, \alpha - s \geq 0, \\ \text{(b)} & |\nabla(P_h v - v)| \leq C h^{1/2} \|v\|_2, & \\ \text{(c)} & \left|\dfrac{\partial v_h}{\partial \eta A}\right| \leq C h^{-3/2} \|v_h\|, & \\ \text{(d)}^{\dagger\dagger\dagger} & |\nabla(P_h v - v)| \leq C h^{\alpha} \|v\|_{(2\alpha+3+\varepsilon)/2}, & 0 \leq \alpha < 1/2. \end{array}$$

Let $y(y_h)$ be the solution of (5.1) ((5.11)). *Then, with $e(t) \equiv y(t) - y_h(t)$, we have*

(i) $|e(t)|_{L_p[0T;\, L_2(\Omega)]} \leq C h^{1/2} |u|_{L_p[0T;\, L_2(\Gamma)]}$, $1 < p < \infty$,

(ii) $|e(t)|_{L_p[0T;\, L_2(\Omega)]} \leq C h^{1/2} \ln h |u|_{L_p[0T;\, L_2(\Gamma)]}$, $p = 1, \infty$.

Remarks. 1. In view of the regularity of the solution $y(1) \in L_p[0T; H^{1/2}(\Omega)]$ for boundary data $u \in L_p[0T; L_2(\Gamma)]$, as described by $u(t)$ (5.14), we deduce that the results of Theorem 5.1 are *optimal*; indeed they reproduce the approximation property (5.15)(a) with $\alpha = 1/2$ and $s = 0$.

2. Spaces which comply with (5.15) are, for example, linear (or higher-order) splines defined on uniform meshes (see also [16]).

The proof of Theorem 5.1 follows immediately from Theorem 3.1 after taking $\overline{Q} = 3/4$, $Q = 3/4 + \varepsilon$ and making use of assumptions (b), (c) and (d). In fact, (5.15)(b), (c), (d) readily imply that

$$(\text{b}') \qquad |D^* A^* (R_h v - v)|_\Gamma \leq C h^{1/2} \|v\|_2 \leq C h^{1/2} |v|_{D(A^*)} \quad \text{for } v \in D(A^*)$$

†††If $V_h \subset H^{3/2+\varepsilon}(\Omega)$, then (5.15)(d) follows automatically from the Trace Theorem and (5.15) (a), (b).

and

$$\text{(c}') \qquad |D^*A^*v|_\Gamma \leqslant Ch^{-3/2}\|v_h\|,$$

$$\text{(d}') \qquad \left|D^*A^*(R_h v - v)\right|_\Gamma \leqslant Ch^{\alpha}\|v\|_{(2\alpha+3+\varepsilon)/2}, \qquad 0 \leqslant \alpha \leqslant 1/2.$$

Thus, assumptions (3.1), (3.2), and (3.1′) are satisfied with $\overline{Q} = 3/4$, $Q = 3/4 + \varepsilon$.

6. Concluding Remarks. The same technique can be used to approximate parabolic equations with different types of boundary conditions. For instance, in the case of Neumann boundary conditions we simply replace the operator D by $N \in \mathscr{L}(L_2(\Gamma) \to H^{3/2}(\Omega))$, where N is an appropriate "Neumann" extension. In this case the optimal rate of convergence is $\mathcal{O}(h^{3/2})$, which reflects the optimal regularity of the solution. In order to obtain the convergence results for the Neumann problem with $H^{-1}(\Gamma)$ boundary data (see (1.14)), we simply invoke Theorem 3.1 with $Q = \overline{Q} = 3/4$, which gives us immediately the error estimates (1.14).

Department of Mathematics
University of Florida
Gainesville, Florida 32611

1. I. BABUŠKA, "The finite element method with Lagrangian multipliers," *Numer. Math.*, v. 20, 1973, pp. 179–192.

2. A. V. BALAKRISHNAN, *Applied Functional Analysis*, Springer-Verlag, Berlin, 1976.

3. J. BRAMBLE, A. SCHATZ, V. THOMÉE & L. B. WAHLBIN, "Some convergence estimates for semidiscrete Galerkin type approximations for parabolic equations," *SIAM J. Numer. Anal.*, v. 14, no. 2, 1977, pp. 218–241.

4. P. CANNARSA & V. VESPRI, "On maximal L^p regularity for the abstract Cauchy problem," *Pubbl. Sci. Accad. Navale*, v. 52, Livorno, 1984.

5. J. DOUGLAS & T. DUPONT, "Galerkin methods for parabolic equations," *SIAM J. Numer. Anal.*, v. 4, 1970, pp. 575–626.

6. D. FUJIWARA, "Concrete characterizations of the domains of fractional powers of some elliptic differential operators of the second order," *Proc. Japan Acad.*, v. 43, 1967, pp. 82–86.

7. M. HUANG & V. THOMÉE, "Some convergence estimates for semidiscrete type schemes for time-dependent nonselfadjoint parabolic equations," *Math. Comp.*, v. 37, 1981, pp. 327–346.

8. F. JOHN & L. NIRENBERG, "On functions with bounded mean oscillations," *Comm. Pure Appl. Math.*, v. 14, 1961, pp. 415–426.

9. I. LASIECKA, "Convergence estimates for semidiscrete approximations of nonselfadjoint parabolic equations," *SIAM J. Numer. Anal.*, v. 21, 1984, pp. 894–909.

10. I. LASIECKA, "Approximations of analytic and differentiable semigroups—Rate of convergence with nonsmooth initial conditions," *Infinite Dimensional Systems*, Proc. Conf. on Operator Semigroups and Applications, Lecture Notes in Math., vol. 1076, Springer-Verlag, Berlin and New York, 1984, pp. 123–138.

11. I. LASIECKA, "A unified theory for abstract parabolic boundary problems: a semigroup approach," *Appl. Math. Optim.* v. 6, 1980, pp. 287–333.

12. I. LASIECKA, J. P. LIONS & R. TRIGGIANI, "Regularity results for hyperbolic systems—Trace theory," *J. Math. Pures Appl.* (To appear.)

13. M. LUSKIN & R. RANNACHER, "On the smoothing property of the Galerkin method for parabolic equations," *SIAM J. Numer. Anal.*, v. 19, 1981, pp. 93–113.

14. J. NEČAS, *Les Méthodes Directes en Théorie des Équations Elliptiques*, Eds. Acad. Czech. des Sciences, Prague, 1967.

15. V. NERI, "Some properties of functions with bounded mean oscillations," *Studia Math.*, v. 61, 1977, pp. 63–75.

16. J. NITSCHE, "Über ein Variationsprinzip zur Lösung von Dirichlet Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind," *Abh. Math. Sem. Univ. Hamburg*, v. 36, 1971, pp. 9–15.

17. A. PAZY, "Semigroups of linear operators and applications to partial differential equations," *Appl. Math. Sci.*, 44, 1983.

18. N. M. RIVIÈRE, "Interpolation à la Marcinkiewicz," *Rev. Un. Mat. Argentina*, v. 25, 1970/71, pp. 363–377.

19. P. H. SAMMON, "Convergence estimates for semidiscrete parabolic equation approximations," *SIAM J. Numer. Anal.*, v. 19, 1982, pp. 68–81.

20. L. DE SIMON, "Un' applicazione della teoria degli integrali singolari allo studio delle equazioni differenziali lineari astratte del primo ordine," *Rend. Sem. Mat. Univ. Padova*, v. 34, 1964, pp. 205–223.

21. J. SCHWARTZ, "A remark on inequalities of Calderon-Zygmund type for vector-valued functions," *Comm. Pure Appl. Math.*, v. 14, 1961, pp. 785–799.

22. E. STEIN, *Singular Integrals, Harmonic Functions, and Differentiable Properties of Functions of Several Variables*, Proc. Sympos. Pure Math., vol. 10, Amer. Math. Soc., Providence, R.I., 1967.