A finite element splitting method for a convection-diffusion problem

V. Thomée * and Stig Larsson †

November 12, 2018

1 Introduction

In this paper we shall consider the numerical solution of the convection-diffusion problem in the cube $\Omega = (0,1)^d,\ d \geq 1$,

$$(1.1)\quad \frac{\partial u}{\partial t} = \nabla \cdot (a\nabla u) + b \cdot \nabla u - cu, \quad \text{in } \Omega, \quad \text{for } t > 0, \quad \text{with } u(0) = v,$$

with periodic boundary conditions, where the initial function $v = v(x)$, the positive definite $d \times d$ matrix $a = a(x) = (a_{ij}(x))$, the vector $b = b(x) = (b_1(x), \dots, b_d(x))$, and the positive function $c = c(x)$ are periodic and smooth.

The equation (1.1) is a case of the initial-value problem for the operator equation

$$(1.2)\qquad \frac{du}{dt} = -Au + Bu, \quad \text{for } t \geq 0, \quad \text{with } u(0) = v,$$

where A and B represent different physical processes, in our case $Au = -\nabla \cdot (a\nabla u) + cu,\ Bu = b \cdot \nabla u$. The solution of (1.2) may be formally expressed as

$$u(t) = E(t)v = e^{-t(A-B)}v, \quad \text{for } t \geq 0.$$

*Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE–412 96 Gothenburg, Sweden; E-mail: thomee@chalmers.se

†Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE–412 96 Gothenburg, Sweden; E-mail: stig@chalmers.se

In some sense, in (1.1) A and B may be thought of as representing slow and fast physical processes. To discretize such an equation in time, a common approach is to split $A-B$ into A and $-B$. With k a time step, one introduces $t_n = nk$ and then, on each time interval (t_{n-1}, t_n) one may use the first order splitting

$$E(k) = e^{-k(A-B)} \approx E_k = e^{kB}e^{-kA}, \tag{1.3}$$

or the second order symmetric Strang splitting [5], [4].

$$E(k) = e^{-k(A-B)} \approx E_k = e^{\frac{1}{2}kB}e^{-kA}e^{\frac{1}{2}kB}, \tag{1.4}$$

which thus locally involves solutions of

$$u_t = -Au \quad \text{and} \quad u_t = Bu, \tag{1.5}$$

see e.g. Hundsdorfer and Verwer [2] and references therein. In this paper we will restrict ourselves to the first order splitting (1.3). The equations in (1.5) then suggest replacing the exact solution $u(t)$ of (1.1) at $t_n = nk$ by the time discrete solution $u_k(nk) = E_k^n v$.

We note that if A and B commute, which for (1.1) holds when a and b are independent of x, then $e^{-k(A-B)} = e^{kB}e^{-kA}$ so that the error in (1.3) vanishes. When A and B do not commute, then formally, by Taylor expansion, $e^{-k(A-B)} - e^{kB}e^{-kA} = O(k^2)$. Error estimates for the time splitting, depending on the regularity of the initial values may be found in Jahnke and Lubich [3], Hansen and Ostermann [1] and references therein.

Setting

$$\|v\|_s = \|v\|_{H^s}, \quad s \geq 0,$$

we note that

$$\|Av\|_s \leq C\|v\|_{s+2}, \quad \|Bv\|_s \leq C\|v\|_{s+1},$$

and

$$\|e^{-t(A-B)}v\|_s + \|e^{-tA}v\|_s + \|e^{tB}v\|_s \leq C_T\|v\|_s, \quad \text{for } s \geq 0, \quad t \leq T. \tag{1.6}$$

For the first order splitting (1.3) we shall show the following local error estimate

$$\|E_k v - E(k)v\| \leq Ck^2\|v\|_4. \tag{1.7}$$

and, as a result, the following global error estimate

$$(1.8) \qquad \|E_k^n v - E(nk)v\| \le CTk\|v\|_4, \quad \text{for } nk \le T.$$

The regularity requirement on the initial data may be reduced by using a smoothing property of $E(t)$.

To define the spatially semidiscrete version of (1.1), which will be discussed in detail in Section 2 below, we let S_h be the periodic, continuous piecewise linear functions on the triangulation $\mathcal{T}_h$. We set

$$A(\psi,\chi) = (a\cdot\nabla\psi,\nabla\chi) + (c\psi,\chi), \quad B(\psi,\chi) = (b\cdot\nabla\psi,\chi), \quad \forall\ \psi,\chi\in S_h,$$

The Standard Galerkin spatially semidiscrete version of (1.1) is then to find $u_h(t)\in S_h$ for $t\ge 0$ such that

$$(1.9) \qquad (u_{h,t},\chi) + A(u_h,\chi) - B(u_h,\chi) = 0, \quad \forall\chi\in S_h, \quad \text{for } t\ge 0,$$

with $u_h(0) = v_h \approx v$. Introducing the operators $A_h, B_h : S_h \to S_h$ by

$$(A_h\psi,\chi) = A(\psi,\chi), \quad (B_h\psi,\chi) = B(\psi,\chi), \quad \forall\psi,\chi\in S_h,$$

the equation (1.9) may be written

$$(1.10) \qquad u_{h,t} + A_h u_h - B_h u_h = 0, \quad \text{for } t>0, \quad \text{with } u_h(0) = v_h.$$

We note that, in particular, A_h is positive definite and $B_h\psi = P_h(b\cdot\nabla\psi)$ for $\psi\in S_h$. We recall that the solution operator $E_h(t) = e^{-t(A_h-B_h)}$ of (1.10) is stable, and that the solution of (1.10) satisfies a $O(h^2)$ error estimate for suitable v_h. Combined with the backward Euler time discretization a $O(k+h^2)$ error estimate holds.

As a first step in defining our splitting finite element solution we set $E_{h,k} = e^{kB_h}e^{-kA_h}$, where the two factors are the solution operators of the hyperbolic and parabolic finite element analogues of (1.5),

$$(1.11) \qquad u_{h,t} - B_h u_h = 0 \quad \text{and} \quad u_{h,t} + A_h u_h = 0,$$

respectively. We shall think of the time discrete solution operator $E_{h,k}$ as the spatial discretization of the operators e^{kB} and e^{-kA} rather than as the splitting of $E_h(k) = e^{k(B_h-A_h)}$. As is well known the spatial discretization of the parabolic equation in (1.5) has a $O(h^2)$ error, whereas, in general, only a $O(h)$ error estimate holds for the hyperbolic equation. We shall show that for suitable v_h,

$$\|E_{h,k}v_h - E(t_n)v\| \le C_T(v)(k+h), \quad \text{where } t_n \le T, \tag{1.12}$$

which is thus of nonoptimal first order in h.

The equations (1.11) may also be expressed in matrix form as follows. With $\{\Phi_j\}_{j=1}^N$ the pyramid function basis for S_h, let $\mathcal{M} = (m_{ij}), m_{ij} = (\Phi_i, \Phi_j)$, $\mathcal{A} = (a_{ij}), a_{ij} = A(\Phi_i, \Phi_j)$, and $\mathcal{B} = (b_{ij}), b_{ij} = B(\Phi_j, \Phi_i)$. Setting $u_h(t) = \sum_{j=1}^N \alpha_j(t)\Phi_j$ and $\alpha = (\alpha_1, \ldots, \alpha_N)^T$, (1.11) may then be written

$$\mathcal{M}\alpha' - \mathcal{B}\alpha = 0 \quad \text{and} \quad \mathcal{M}\alpha' + \mathcal{A}\alpha = 0, \quad \text{for } t > 0.$$

In the one-dimentional case of the equation $u_t = u_{xx} + u_x - u$ this gives

$$\mathcal{M} = h \begin{bmatrix} \frac{2}{3} & \frac{1}{6} & \cdots & \frac{1}{6} \\ \frac{1}{6} & \frac{2}{3} & \cdots & 0 \\ \vdots & \vdots & \ddots & \frac{1}{6} \\ \frac{1}{6} & 0 & \cdots \frac{1}{6} & \frac{2}{3} \end{bmatrix},$$

and

$$\mathcal{A} = \frac{1}{h} \begin{bmatrix} 2 & -1 & \cdots & -1 \\ -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \ddots & -1 \\ -1 & 0 & \cdots -1 & 2 \end{bmatrix} + \mathcal{M}, \quad \mathcal{B} = \frac{1}{2} \begin{bmatrix} 0 & 1 & \cdots & -1 \\ -1 & 0 & 1 \cdots & 0 \\ \vdots & \vdots & \ddots & 1 \\ 1 & 0 & \cdots -1 & 0 \end{bmatrix}.$$

For a computable discretization in space and time, we thus need to replace e^{-kA_h} and e^{kB_h} by rational functions of A_h and B_h. This will be done in Section 3 below. For the hyperbolic solution operator we shall use the explicit approximation

$$H_{h,k} = I + kB_h - \gamma k^2 A_h \approx e^{kB_h}, \tag{1.13}$$

under a mesh-ratio condition for k/h, and for the parabolic part the implicit backward Euler operator

$$Q_{h,k} = (I + kA_h)^{-1} \approx e^{-kA_h}. \tag{1.14}$$

Thus the time stepping operator is then

$$\mathcal{E}_{h,k} = H_{h,k}Q_{h,k} \approx e^{kB_h}e^{-kA_h} \approx e^{k(B-A)}. \tag{1.15}$$

We shall show that, for v_h appropriate,

$$\|\mathcal{E}_{h,k}^n v_h - E(t_n)v\| \le C_T(v)(k+h), \quad \text{for } t_n \le T. \tag{1.16}$$

In matrix form the equations for one time step of the splitting are

$$\mathcal{M}\alpha^{n+1} = (\mathcal{M} + k\mathcal{B} - \gamma k^2 \mathcal{A})\alpha^n \quad \text{and} \quad (\mathcal{M} + k\mathcal{A})\alpha^{n+1} = \mathcal{M}\alpha^n. \tag{1.17}$$

The first equation in (1.17) is not fully explicit in that it requires the application of the inverse of the mass matrix. In Section 4 we shall therefore combine the above approach with the application of the Lumped Mass Method. This uses the approximate inner product $(\cdot,\cdot)_h$ on S_h for the first term in the semidiscrete equation (1.9) and its completely discrete hyperbolic and parabolic counterpars, which results in replacing $\mathcal{M}$ by a diagonal matris $\mathcal{D}$ in (1.17). The error estimate (1.16) will then remain valid.

2 Splitting of the semidiscrete problem

We now want to analyze the result of splitting of the spatially semidiscrete problem to approximate the exact solution of (1.1), or $e^{k(B-A)} \approx e^{kB_h}e^{-kA_h}$. We start by splitting the continuous problem on time intervals of length k into a parabolic problem and a hyperbolic problem and then apply finite element semidiscretization to each of these. We first show (1.7) and (1.8).

Lemma 2.1. *We have, with C independent of k,*

$$\|E_k v - E(k)v\|_j \le Ck^{2-j}\|v\|_{4-j}, \quad \textit{for } j = 0, 1,$$

Proof. Setting $F(t) = e^{-t(A-B)} - e^{tB}e^{-tA}$ and noting that $F(0) = F'(0) = 0$, we may use Taylor's formula to obtain

$$\|F(k)v\| = \|(F(k) - F(0) - kF'(0))v\| \le \tfrac{1}{2}k^2 \sup_{s\le k}\|F''(s)v\|.$$

Here, for $s \le k$,

$$\begin{aligned}\|F''(s)v\| &\le \|(A-B)^2 e^{-s(A-B)}v\| + C\sum_{i_1+i_2=2}\|B^{i_1}e^{sB}A^{i_2}e^{-sA}v\| \\ &\le C\big(\|v\|_4 + \sum_{i_1+i_2=2}\|v\|_{i_1+2i_2}\big) \le C\|v\|_4,\end{aligned}$$

which shows the inequality for $j = 0$. For $j = 1$ we note that

$$\|F(k)v\|_1 = \|(F(k) - F(0))v\|_1 \le k \sup_{s \le k} \|F'(s)v\|_1.$$

In this case,

$$\|F'(s)v\|_1 \le \|(A-B)e^{-s(A-B)}v\|_1 + \|Be^{sB}e^{-sA}v\|_1 + \|e^{sB}Ae^{-sA}v\|_1 \le C\|v\|_3,$$

which completes the proof. □

Theorem 2.1. *We have, with C independent of k,*

$$\|E_k^n v - E(t_n)v\| \le C_T k \|v\|_4, \quad \text{for } t_n \le T.$$

Proof. We find, using the stability of E_k and Lemma 2.1 with $j = 0$,

$$\begin{aligned}\|E_k^n v - E(t_n)v\| &= \Big\| \sum_{j=0}^{n-1} E_k^{n-j-1}(E_k - E(k))E(t_j)v \Big\| \\ &\le Ck^2 \sum_{j=0}^{n-1} \|E(t_j)v\|_4 \le Cnk^2\|v\|_4 \le CTk\|v\|_4.\end{aligned}$$

□

We now discuss the splitting of the spatially semidiscrete problem (1.10). Recalling that A_h is positive definite, we introduce the discrete norm

$$\|\psi\|_{h,s} = \|A_h^{s/2}\psi\| = (A_h^s\psi, \psi)^{1/2}, \quad \forall \chi \in S_h, \quad \text{for } s \ge 0.$$

We begin with some bounds for the operator B_h with respect to this norm.

Lemma 2.2. *We have*

$$\|B_h\psi\|_{h,j} \le C\|\psi\|_{h,j+1}, \quad \text{for } \psi \in S_h, \ j = 0, 1. \tag{2.1}$$

and

$$\|B_h^2\psi\| \le C\|\psi\|_{h,2}, \quad \text{for } \psi \in S_h. \tag{2.2}$$

Proof. By our definitions we find

$$\|B_h\psi\|^2 = \|P_h(b \cdot \nabla\psi)\|^2 \le C\|\nabla\psi\|^2 \le C(A_h\psi, \psi) = C\|\psi\|_{h,1}^2,$$

which shows (2.1) for $j = 0$. For $j = 1$, and for (2.2), it then suffices to show

$$\|B_h\psi\|_1 \le \|A_h^{1/2}B_h\psi\| \le C\|\psi\|_{h,2}, \quad \text{for } \psi \in S_h.$$

With $T_h = A_h^{-1}$, this is equivalent to

$$\|B_hT_h\psi\|_1 \le C\|\psi\|, \quad \text{for } \psi \in S_h.$$

We have

$$\|B_hT_h\psi\|_1 \le \|B_hP_hT\psi\|_1 + \|B_h(T_h - P_hT)\psi\|_1 = I + II.$$

Here

$$I \le \|P_h(b \cdot \nabla T\psi\|_1 + \|P_h(b \cdot \nabla((I - P_h)T)\psi\|_1 = I_1 + I_2,$$

and, since $\|P_h\varphi\|_1 \le C\|\varphi\|_1$ for $\varphi \in H^1$,

$$I_1 \le C\|\nabla T\psi\|_1 \le C\|T\psi\|_2 \le C\|\psi\|$$

and

$$I_2 \le Ch^{-1}\|(I - P_h)T\psi\|_1 \le Ch^{-1}h\|T\psi\|_2 \le C\|\psi\|.$$

Further,

$$II \le Ch^{-1}\|\nabla(T_h - P_hT)\psi\| \le Ch^{-1}h\|\psi\| \le C\|\psi\|.$$

Together these estimates show the lemma. □

We begin the discussion of the splitting of the spatially semidiscrete problem (1.10) with the following stability result.

Lemma 2.3. *We have, with C independent of h,*

$$\|e^{-t(A_h-B_h)}v_h\|+\|e^{-tA_h}v_h\|+\|e^{tB_h}v_h\| \le e^{Ct}\|v_h\|, \quad \text{for } v_h \in S_h,\ t \ge 0. \tag{2.3}$$

Moreover,

$$\|e^{tB_h}v_h\|_{h,2} \le e^{Ch^{-1}t}\|v_h\|_{h,2}, \quad \text{for } v_h \in S_h,\ t \ge 0. \tag{2.4}$$

Proof. Since $(B_h\chi, \chi) = (b\cdot\nabla\chi, \chi) = -\frac{1}{2}(\nabla\cdot b\chi, \chi) \le C\|\chi\|^2$, we have for the solution $u_h(t) = e^{-t(A_h-B_h)}v_h$ of (1.10),

$$(u_{h,t}, u_h) + (A_h u_h, u_h) = (B_h u_h, u_h) \le C\|u_h\|^2.$$

Hence $(d/dt)\|u_h\|^2 \le C\|u_h\|^2$, from which the first part of (2.3) follows. The other parts follow analogously. For $u_h(t) = e^{tB_h}v_h$, we have similarly, by (2.1) and an inverse inequality,

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|u_h\|_{h,2} &= (u_{h,t}, A_h^2 u_h) = (B_h u_h, A_h^2 u_h)\\ &\le \|B_h u_h\|_{h,1}\|u_h\|_{h,3} \le Ch^{-1}\|u_h\|_{h,2}^2.\end{aligned}$$

This proves (2.4). □

We note that, by Lemma 2.3, $E_{h,k} = e^{kB_h}e^{-kA_h}$ is stable:

$$\|E_{h,k}^n\psi\| \le e^{nkC}\|\psi\| \le e^{CT}\|\psi\|, \quad \text{for } \psi\in S_h, t_n \le T. \tag{2.5}$$

We would like to show that the discrete solution $U_{h,k}^n = E_{k,h}^n v_h$ differs from the semidiscrete solution $u_h(t_n) = E_h(t_n)v_h$ at $t_n = nk$, by $O(k)$, under the appropriate regularity assumptions.

We will now need two lemmas corresponding to the discretization of the parabolic and hyperbolic parts, respectively. Here and below, P_h denotes the L_2-projection and $R_h = T_h A = A_h^{-1}P_h A$ the Ritz projection onto S_h.

Lemma 2.4. *We have*

$$\|e^{-kA_h}R_h v - R_h e^{-kA}v\| \le Ckh^j\|v\|_{2+j}, \quad j = 1, 2. \tag{2.6}$$

and

$$\|e^{kB_h}R_h v - R_h e^{kB}v\| \le Ckh\|v\|_3.$$

Proof. With $u(t) = e^{-tA}v$ and $u_h(t) = e^{-tA_h}R_h v$ we have

$$(u_{h,t}, \chi) + A(u_h, \chi) = 0, \quad \forall\chi\in S_h,\ t\ge 0, \quad \text{with } u_h(0) = R_h v.$$

Our goal is to bound $\theta(k)$ where $\theta = u_h - R_h u$, with $\theta(0) = u_h(0) - R_h v = 0$. We have, with $\rho = R_h u - u$,

$$(\theta_t, \chi) + A(\theta, \chi) = -(\rho_t, \chi), \quad \forall\chi\in S_h,$$

and choosing $\chi = \theta$ we find

$$\|\theta(k)\| \le \|\theta(0)\| + C\int_0^k \|\rho_t(s)\|ds \le Ch^j \int_0^k \|u_t(s)\|_j ds \le Ckh^j\|v\|_{2+j}.$$

For the hyperbolic problem we have, now with $u(t) = e^{tB}v$ and $u_h(t) = e^{tB_h}R_hv$,

$$(u_{h,t}, \chi) - (b\cdot\nabla u_h, \chi) = 0, \quad \forall \chi \in S_h,\ t \ge 0, \quad \text{with } u_h(0) = R_hv.$$

and this time we get for $\theta = u_h - R_hu$, with $\theta(0) = u_h(0) - R_hv = 0$.

$$(\theta_t, \chi) - (b\cdot\nabla\theta, \chi) = -(\rho_t, \chi) + (b\cdot\nabla\rho, \chi), \quad \forall \chi \in S_h,$$

and thus

$$\|\theta(k)\| \le \|\theta(0)\| + C\int_0^k (\|\rho_t(s)\| + \|\rho(s)\|_1)ds \le Ckh\|v\|_3. \tag{2.7}$$

This completes the proof. □

This implies the following error estimate for the combined time stepping operator.

Lemma 2.5. *With $E_{h,k} = e^{kB_h}e^{-kA_h}$ and $E(k) = e^{k(B-A)}$ we have*

$$\|E_{h,k}R_hv - R_hE(k)v\| \le C(k^2 + kh)\|v\|_4.$$

Proof. Recalling $E_k = e^{kB}e^{-kA}$ we have

$$E_{h,k}R_h - R_hE_k = e^{kB_h}(e^{-kA_h}R_h - R_he^{-kA}) + (e^{kB_h}R_h - R_he^{kB})e^{-kA}.$$

Hence, by Lemmas 2.3 and 2.4,

$$\|E_{h,k}R_hv - R_hE_kv\| \le Ckh\|v\|_3.$$

Further,

$$\begin{aligned}\|E_{h,k}R_hv - R_hE(k)v\| &\le \|E_{h,k}R_hv - R_hE_kv\| \\ &\quad + \|(R_h - I)(E_kv - E(k))v\| + \|(E_kv - E(k))v\| = I + II + III.\end{aligned}$$

Here, using also Lemma 2.1 we find

$$I \le Ckh\|v\|_3,\ II \le Ch\|(E_kv - E(k))v\|_1 \le Ckh\|v\|_3,\ III \le Ck^2\|v\|_4,$$

which completes the proof. □

We now show the following global error estimate.

Theorem 2.2. *Assume that* $\|v_h - v\| \le Ch\|v\|_4$. *Then we have*

$$\|E_{h,k}^n v_h - E(t_n)v\| \le C_T(k+h)\|v\|_4, \quad \text{for } t_n = nk \le T.$$

Proof. We find, using (2.5) and Lemma 2.5,

$$\begin{aligned}
\|E_{h,k}^n R_h v - R_h E(t_n)v\| &= \Big\| \sum_{j=0}^{n-1} E_{h,k}^{n-j-1}(E_{h,k}R_h - R_h E(k))E(t_j)v \Big\| \\
&\le C \sum_{j=0}^{n-1} \|(E_{h,k}R_h - R_h E(k))E(t_j)v\| \le C(k^2+kh)\sum_{j=0}^{n-1}\|E(t_j)v\|_4 \\
&\le Cn(k^2+kh)\|v\|_4 \le CT(k+h)\|v\|_4.
\end{aligned}$$

Hence

$$\begin{aligned}
\|E_{h,k}^n v_h - E(t_n)v\| &\le \|E_{h,k}^n(v_h - R_h v)\| + \|E_{h,k}^n R_h v - R_h E(t_n)v\| \\
&\quad + \|(R_h - I)E(t_n)v\| \le C_T(k+h)\|v\|_4. \qquad \square
\end{aligned}$$

3 Complete discretization in time and space

We shall now turn to the analysis of the complete discretization using the operators (1.13) and (1.14) for the hyperbolic and parabolic parts of $E_h(k) = e^{k(B_h - A_h)}$, respectively. We shall first turn to the stability. It is clear that, for the parabolic part, $\|Q_{h,k}\chi\| \le \|\chi\|$, but we shall also need stability property for $H_{h,k}$. By Lemma 2.2, with β suitable,

$$\|B_h\chi\|^2 \le \beta(A_h\chi,\chi). \tag{3.1}$$

Lemma 3.1. *Assume that* $\gamma > \beta$. *Then*

$$\|H_{h,k}\chi\| \le (1+Ck)\|\chi\|, \quad \text{for } k/h \le \rho_0 = \sqrt{(\gamma-\beta)/(\gamma^2\kappa)}.$$

Proof. Setting $\psi = H_{h,k}\chi$, we have $\psi - \chi + \gamma k^2 A_h\chi = kB_h\chi$. Recalling $(B_h\chi,\chi) \le C\|\chi\|^2$ for $\chi \in S_h$ we have

$$(\psi-\chi,\chi) + \gamma k^2(A_h\chi,\chi) = k(B_h\chi,\chi) \le Ck\|\chi\|^2, \tag{3.2}$$

or

$$(\psi - \chi, \tfrac{1}{2}(\psi + \chi) - \tfrac{1}{2}(\psi - \chi)) + \gamma k^2 (A_h \chi, \chi) \le Ck\|\chi\|^2.$$

Hence

$$\tfrac{1}{2}(\|\psi\|^2 - \|\chi\|^2) - \tfrac{1}{2}\|\psi - \chi\|^2 + \gamma k^2 (A_h \chi, \chi) \le Ck\|\chi\|^2.$$

Here, if $\|A_h\chi\|^2 \le \kappa h^{-2}(A_h\chi, \chi)$, we have, for $k/h \le \rho_0$, using (3.1)

$$\|\psi - \chi\|^2 \le 2k^2\|B_h\chi\|^2 + 2\gamma^2 k^4 \|A_h\chi\|^2 \le 2k^2(\beta + \gamma^2\kappa\rho_0^2)(A_h\chi, \chi).$$

Thus, if $\beta < \gamma$ and $\rho_0 = \sqrt{(\gamma - \beta)/(\gamma^2\kappa)}$,

$$\tfrac{1}{2}\|\psi - \chi\|^2 \le k^2(\beta + \gamma^2\kappa\rho_0^2)(A_h\chi, \chi) \le \gamma k^2 (A_h\chi, \chi).$$

Hence $\|\psi\|^2 \le (1 + Ck)\|\chi\|^2$, which shows our claim. $\square$

We note that the choice of γ which makes ρ_0 as large as possible is $\gamma = 2\beta$, in which case $\rho_0 = 1/\sqrt{4\kappa\beta}$. In order to increase the limit ρ_0 for the mesh ratio one may replace the operator $H_{h,k}$ by $H^m_{h,k/m}$ for some positive integer m. This means that the approximate solution of the hyperbolic part of the time step is obtained in m steps of length k/m, and we note that the stability estimate Lemma 3.1 then implies

$$\|H^m_{h,k/m} v\| \le (1 + Ck/m)^m \|v\| \le e^{Ck}\|v\|, \quad \text{for } k/h \le m\rho_0. \tag{3.3}$$

In the sequel we will now use

$$\mathcal{E}_k = H^m_{h,k_m} Q_k, \quad \text{and} \quad k_m := k/m, \ k_m/h \le \rho_0, \quad \text{with } m \ge 1.$$

We have the following error estimate for one step of the hyperbolic and parabolic approximations.

Lemma 3.2. *Assume $k/h \le m\rho_0$. We have*

$$\|(H^m_{h,k_m} - e^{kB_h})\chi\| \le Ck^2 \|A_h\chi\|. \tag{3.4}$$

and

$$\|(Q_{h,k} - e^{-kA_h})\chi\| \le \tfrac{1}{2}k^2\|A_h^2\chi\|. \tag{3.5}$$

Proof. Using Taylor expansion and Lemma 2.2, we have

$$\begin{aligned}
\|(e^{kB_h} - H_{h,k})v\| &\le \|(e^{kB_h} - I - kB_h)v\| + \gamma k^2\|A_h v\| \\
&\le \Big\| \int_0^k (k-s)e^{sB_h}B_h^2 v\,ds \Big\| + Ck^2\|v\|_{h,2} \le Ck^2\|v\|_{h,2},
\end{aligned}$$

which shows (3.4) for $m = 1$.

To show (3.4) for $m > 1$ we write

$$H_{h,k_m}^m v - e^{kB_h}v = \sum_{j=0}^{m-1} H_{h,k_m}^{m-j-1}(H_{h,k_m} - e^{k_m B_h})e^{jk_m B_h}v. \tag{3.6}$$

By replacing k by k_m in the already proved case of (3.4), we get

$$\|(e^{k_m B_h} - H_{h,k_m})v\| \le Ck_m^2\|v\|_{h,2}. \tag{3.7}$$

From (2.4) with $t = k_m$ and $k_m \le \rho_0 h$ we have

$$\|e^{jk_m}v\|_{h,2} \le Ce^{Cjk_m h^{-1}}\|v\|_{h,2} \le C\|v\|_{h,2},$$

because $jk_m h^{-1} \le \rho_0 j/m \le \rho_0$. Using (3.3), (3.6), and (3.7) we then conclude, for $k_m/h \le \rho_0$,

$$\begin{aligned}
\|H_{h,k_m}^m v - e^{kB_h}v\| &\le e^{Ck}\sum_{j=0}^{m-1}\|(H_{h,k_m} - e^{k_m B_h})e^{jk_m B_h}v\| \\
&\le Ck_m^2\sum_{j=0}^{m-1}\|e^{jk_m B_h}v\|_{h,2} \\
&\le Cmk_m^2\|v\|_{h,2} \le Ck^2\|v\|_{h,2}.
\end{aligned}$$

For (3.5) we have, with $r(\lambda) = (1+\lambda)^{-1}$,

$$Q_{h,k} = r(kA_h) = I - kA_h + \int_0^k (k-s)r''(sA_h)\,ds$$

and hence, since $r''(\lambda) = 2(1+\lambda)^{-3}$,

$$\begin{aligned}
\|e^{-kA_h}v - Q_{h,k}v\| &= \Big\| \int_0^k (k-s)(e^{-sA_h} - r''(sA_h))A_h^2 v\,ds \Big\| \\
&\le Ck^2\|A_h^2 v\| \le Ck^2\|v\|_{h,4},
\end{aligned}$$

which completes the proof. $\square$

As a result, we have the following error estimate for the completely discrete time stepping operator.

Lemma 3.3. *Under the above assumptions about γ and ρ_0 we have with $\mathcal{E}_{h,k}$ defined by* (1.15) *and $E_{h,k} = e^{kB_h}e^{-kA_h}$,*

$$\|(\mathcal{E}_{h,k} - E_{h,k})R_h v\| \le Ck^2\|v\|_4.$$

Proof. We have, for $\chi \in S_h$, by Lemmas 3.1 and 3.2,

$$\begin{aligned} &\|(\mathcal{E}_{h,k} - E_{h,k})\chi\| \le \|H^m_{h,k_m}(Q_{h,k} - e^{-kA_h})\chi\| \\ &\quad + \|(H^m_{h,k_m} - e^{kB_h})e^{-kA_h}\chi\| \le Ck^2(\|A_h^2\chi\| + \|A_h\chi\|) \le Ck^2\|A_h^2\chi\|. \end{aligned}$$

To complete the proof we note that

$$\begin{aligned} \|A_h^2 R_h v\| &= \|A_h P_h A v\| \le \|A_h R_h A v\| + \|A_h(P_h - R_h)Av\| \\ &\le \|P_h A^2 v\| + Ch^{-2}h^2\|Av\|_2 \le C\|v\|_4. \end{aligned}$$

□

We can now show the following complete error estimate.

Theorem 3.1. *Assume that $\|v_h - v\| \le Ch\|v\|_4$. Then we have*

Proof. We find, using Lemmas 2.5 and 3.3,

$$\begin{aligned} \|\mathcal{E}_{h,k}R_h v - R_k E(k)v\| &\le \|(\mathcal{E}_{h,k} - E_{h,k})R_h v\| + \|(E_{h,k}R_h - R_h E(k))v\| \\ &\le C(k^2 + hk)\|v\|_4, \quad \text{for } t_n \le T, \end{aligned}$$

and hence, as in the proof of Theorem 2.2,

$$\begin{aligned} \|\mathcal{E}^n_{h,k}R_h v - R_h E(t_n)v\| &= \|\sum_{j=0}^{n-1} \mathcal{E}^{n-j-1}_{h,k}(\mathcal{E}_{h,k}R_h - R_h E(k))E(t_j)v\| \\ &\le C(k^2 + kh)\sum_{j=0}^{n-1}\|E(t_j)v\|_4 \le CT(k+h)\|v\|_4. \end{aligned}$$

The result for $\|v_h - v\| \le Ch\|v\|_4$ is then completed as in Theorem 2.2. □

4 Explicit solution of the hyperbolic problem

The above approximate hyperboic solution operator $H_{h,k}$ is not fully explicit in that for the solution of the forward Euler problem (1.17) one needs to apply the inverse of the mass matrix. We shall therefore combine the above thoughts with the application of the Lumped Mass Method, using the approximate inner product on S_h defined by

$$(4.1)\quad (\psi,\chi)_h = \sum_{\tau\in\mathcal{T}_h} Q_{\tau,h}(\psi\chi), \quad \text{where } Q_{\tau,h}(f) = \tfrac{1}{3}\sum_{j=1}^{3} f(P_{\tau,h}) \approx \int_\tau f\,dx.$$

where for a triangle τ of the triangulation $\mathcal{T}_h$, $P_{\tau,j}$, $j = 1, 2, 3$, are its vertices. We note that the corresponding norm $\|\cdot\|_h$ is equivalent to $\|\cdot\|$, or

$$(4.2)\qquad \mu^{-1}\|\chi\| \le \|\chi\|_h \le \mu\|\chi\|, \quad \text{for } \chi\in S_h, \quad \text{with } \mu > 1.$$

The matrix formulations of (1.17) are now

$$(4.3)\qquad \mathcal{D}\alpha^{n+1} = (\mathcal{D} + k\mathcal{B} - \gamma k^2\mathcal{A})\alpha^n \quad \text{and} \quad (\mathcal{D} + k\mathcal{A})\alpha^{n+1} = \mathcal{D}\alpha^n,$$

respectively. Thus, in contrast to (1.17), the first equation in (4.3) is fully explicit in that the mass matrix $\mathcal{M}$ has been replaced by the diagonal matrix $\mathcal{D}$ with diagonal elements $d_i = \|\Phi_i\|_h^2$.

We introduce the positive definite operator $I_h : S_h \to S_h$ by

$$(I_h\psi, \chi) = (\psi, \chi)_h, \quad \forall \psi, \chi \in S_h.$$

The semidiscrete problems analogous to (1.11) may then be written

$$I_h u_{h,t} - B_h u_h = 0 \quad \text{and} \quad I_h u_{h,t} + A_h u_h = 0, \quad \text{with } u_h(0) = v_h.$$

Recalling that

$$|(\psi,\chi)_h - (\psi,\chi)| \le Ch^2\|\nabla\psi\|\,\|\nabla\chi\|, \quad \forall\psi,\chi\in S_h,$$

we find

$$\|(I_h - I)\psi\| \le Ch\|\psi\|_1 \le Ch(A_h\psi,\psi)^{1/2}.$$

In fact,

$$|((I_h - I)\psi, \chi)| \le Ch^2\|\nabla\psi\|\,\|\nabla\chi\| \le Ch\|\psi\|_1\|\chi\|,$$

which implies our claim. It also follows that

$$\|(I_h^{-1} - I)\psi\| = \|I_h^{-1}(I - I_h)\psi\| \le C\|(I_h - I)\psi\| \le Ch(A_h\psi, \psi)^{1/2}. \tag{4.4}$$

The solution operator of the spatially semidiscrete hyperbolic problem is then $H_h(t) = e^{tI_h^{-1}B_h}$, and we have the following error bound.

Lemma 4.1. *We have*

$$\|H_h(k)R_h v - R_h e^{kB} v\| \le Ckh\|v\|_3.$$

Proof. By the dfinition of $H_h(k)$ and by Taylors formula we have for $\chi \in S_h$,

$$\|(e^{kI_h^{-1}B_h} - e^{kB_h})\chi\| \le k\|(I_h^{-1} - I)B_h\chi\| + Ck^2(\|B_h I_h^{-1} B_h\chi\| + \|B_h^2\chi\|).$$

Here, using Lemma 2.2 and (4.4),

$$\begin{aligned}\|B_h I_h^{-1} B_h\chi\| &\le \|B_h^2\chi\| + \|B_h(I_h^{-1} - I)B_h\chi\| \le C\|A_h\chi\| \\ &+ Ch^{-1}\|(I_h^{-1} - I)B_h\chi\| \le C\|A_h\chi\| + Ch^{-1}h\|B_h\chi\|_1 \le C\|A_h\chi\|.\end{aligned} \tag{4.5}$$

and thus, with $\chi = R_h v$,

$$\begin{aligned}\|H_h(k)R_h v - e^{kB}R_h v\| &\le Ckh\|B_h R_h v\|_1 + Ck^2\|A_h R_h v\| \\ &\le Ck(h+k)\|A_h R_h v\| \le Ck(h+k)\|v\|_2.\end{aligned}$$

In view of Lemma 2.4 this completes the proof. □

The corresponding result for the parabolic problem is the following.

Lemma 4.2. *With $Q_h(k) = e^{-kI_h^{-1}A_h}$ we have*

$$\|Q_h(k)R_h v - R_h e^{-kA} v\| \le Ckh^2\|v\|_{2+j}, \quad j = 1, 2.$$

Proof. Setting $\theta(t) = Q_h(t)R_h v - R_h e^{-tA} v$, with $\theta(0) = 0$, and $\rho = R_h u - u$, we now have

$$(\theta_t, \chi)_h + A(\theta, \chi) = -(\rho_t, \chi)_h, \quad \forall \chi \in S_h.$$

Hence, with $\chi = \theta$,

$$\frac{d}{dt}\|\theta\|_h \le \|\rho_t\|_h \le Ch^j\|u_t\|_j \le Ch^j\|v\|_{2+j},\ j = 1, 2,$$

so that $\|\theta(k)\|_h \le Ckh^j\|v\|_{2+j}$, which shows the lemma. □

We now apply discretization in space to the first order split time step operator and begin with the error in one time step, analogous to Lemma 2.5.

Lemma 4.3. *For $\mathcal{E}_{h,k} = H_h(k)Q_h(k)$ we have*

$$\|\mathcal{E}_{h,k}R_h v - R_h E(k)v\| \le C(k^2 + kh)\|v\|_4.$$

Proof. With $E_k = e^{tB}e^{-tA}$ we have

$$\mathcal{E}_{h,k}R_h - R_h E_k = H_h(k)(Q_h(h)R_h - R_h e^{-kA}) + (H_h(k)R_h - R_h e^{kB})e^{-kA}.$$

Hence, by Lemmas 4.1 and 4.2,

$$\|\mathcal{E}_{h,k}R_h v - R_h E_k v\| \le Ckh\|v\|_3.$$

The proof is now completed as in Lemma 2.5. □

We will now analyze the following fully discrete numerical method based on the above first order splitting of the continuous problem and the corresponding spatially semidiscrete method considered above. For the hyperbolic solution operator we shall use the forward Euler approximation $\psi = H_{h,k}\chi \approx H_h(k)\chi$ defined by

$$\Big(\frac{\psi - \chi}{k}, \varphi\Big)_h + \gamma k(A_h\chi, \varphi) = (B_h\chi, \varphi), \quad \forall \varphi \in S_h, \tag{4.6}$$

or $\psi = \chi + kI_h^{-1}(B_h - \gamma k A_h)\chi$, and for the parabolic part the implicit backward Euler operator, with $\psi = Q_{h,k}\chi$ defined by

$$\Big(\frac{\psi - \chi}{k}, \varphi\Big)_h + (A_h\psi, \varphi) = 0, \quad \forall \varphi \in S_h,$$

or $\psi = (I_h + kA_h)^{-1}I_h\chi = (I + kI_h^{-1}A_h)^{-1}\chi$. The time stepping operator is then

$$\mathcal{E}_{h,k} = H_{h,k}Q_{h,k} \approx H_h(k)Q_h(k) \approx e^{kB}e^{-kA}. \tag{4.7}$$

We begin with a stability property for $H_{h,k}$.

Lemma 4.4. *For γ large enough there is a ρ_0 such that for $k/h \le \rho_0$, we have*

$$\|H_{h,k}\chi\|_h \le (1 + Ck)\|\chi\|_h, \quad \textit{for } \chi \in S_h.$$

Proof. We follow the proof of Lemma 3.1. Setting $\psi = H_{h,k}\chi$, we have now, Instead of (3.2),

$$\tfrac{1}{2}(\|\psi\|_h^2 - \|\chi\|_h^2) - \tfrac{1}{2}\|\psi - \chi\|_h^2 + \gamma k^2 (A_h\chi, \chi) \le Ck\|\chi\|^2. \tag{4.8}$$

Here we have

$$\|\psi - \chi\|_h^2 \le \mu^2 \|\psi - \chi\|^2 \le \mu^2 \Big(4k^2 \|B_h\chi\|^2 + 4\gamma^2 k^4 \|A_h\chi\|^2 \Big).$$

For $k/h \le \rho_0$, if $\|A_h\chi\|^2 \le \kappa h^{-2}(A_h\chi, \chi)$, we have

$$\mu^2\gamma^2 k^4 \|A_h\chi\|^2 \le \mu^2\gamma^2 k^4 \kappa h^{-2} (A_h\chi, \chi) \le \mu^2\gamma^2 k^2 \kappa \rho_0^2 (A_h\chi, \chi),$$

and thus, if $2\mu^2\beta < \gamma$ and ρ_0 sufficiently small,

$$\tfrac{1}{2}\|\psi - \chi\|_h^2 \le 2(\mu^2\gamma^2\kappa\rho_0^2 + \mu^2\beta) k^2 (A_h\chi, \chi) \le \gamma k^2 (A_h\chi, \chi).$$

Hence, (4.8) implies $\|\psi\|_h^2 \le (1 + Ck)\|\chi\|_h^2$, which shows our claim. $\square$

For the parabolic part the corresponding result is as follows.

Lemma 4.5.

$$\|Q_{h,k}\chi\|_h \le \|\chi\|_h, \quad \textit{for } \chi \in S_h.$$

Proof. We have with $\psi = Q_{h,k}\chi$,

$$(\psi - \chi, \varphi)_h + k(A_h\psi, \varphi) = 0, \quad \forall \varphi \in S_h,$$

In particular, with $\varphi = \psi$,

$$\|\psi\|_h^2 \le (\chi, \psi)_h \le \|\chi\|_h \|\psi\|_h,$$

which shows our claim. $\square$

As a result we have the following.

Lemma 4.6. *The operator $\mathcal{E}_{h,k} = H_{h,k}Q_{h,k}$ is stable so that*

$$\|\mathcal{E}_{h,k}^n \chi\| \le C_T \|\chi\|, \quad \forall \chi \in S_h, \quad \textit{for } t_n \le T.$$

Proof. Since $\|Q_{h,k}\chi\|_h \leq \|\chi\|_h$, we have

$$\|\mathcal{E}_{h,k}\chi\|_h \leq (1+Ck)\,\|Q_{h,k}\chi\|_h \leq (1+Ck)\|\chi\|_h.$$

Since the norms $\|\cdot\|_h$ and $\|\cdot\|$ are equivalent, this implies our claim. □

We now show a complete error estimate for the hyperbolic time stepping operator.

Lemma 4.7. *We have*

$$\|(H_{h,k}R_h v - R_h e^{kB} v\| \leq Ck(h+k)\|v\|_3.$$

Proof. By Taylor's formula, using (4.5),

$$\|(H_{h,k} - H_h(k))R_h v\| \leq Ck^2\|B_h I_h^{-1} B_h R_h v\| \leq Ck^2\|A_h R_h v\| \leq Ck^2\|v\|_2$$

Together with Lemma 2.4 this completes the proof. □

The following is the corresponding parabolic estimate.

Lemma 4.8. *We have*

$$\|Q_{h,k}R_h v - R_h e^{-kA} v\| \leq C(hk+k^2)\|v\|_4.$$

Proof. With $\theta^1 = Q_{h,k}R_h v - R_h u(k), \theta^0 = R_h v - R_h u(0) = 0$ and $u^1 = u(k)$ we have

$$(\bar\partial\theta^1,\chi)_h + A(\theta^1,\chi) = (u_t^1,\chi) - (\bar\partial R_h u^1,\chi)_h \quad \text{for } \chi \in S_h,$$

or, with $\chi = \theta^1$,

$$\begin{aligned}(\bar\partial\theta^1,\theta^1)_h + A(\theta^1,\theta^1) &= (u_t^1 - \bar\partial u^1,\theta^1) + (\bar\partial u^1 - \bar\partial R_h u^1,\theta^1)\\ &\quad + \varepsilon_h(\bar\partial R_h u^1,\theta^1) = I + II + III.\end{aligned}$$

Here

$$|I| \leq \|u_t^1 - \bar\partial u^1\|\,\|\theta^1\| \leq C\int_0^k \|u_{tt}\|ds\|\theta^1\| \leq Ck\|v\|_4\|\theta^1\|.$$

Further

$$|II| \leq Ch^2\|\bar\partial u^1\|_2\|\theta^1\| \leq Ch^2k^{-1}\int_0^k \|u_t\|_2 ds\|\theta^1\| \leq Ch^2\|v\|_4\|\theta^1\|.$$

Also

$$|III| \leq h^2\|\nabla R_h \bar{\partial} u^1\| \|\nabla\theta^1\| \leq Ch^2\|\bar{\partial}\nabla u^1\| \|\nabla\theta^1\|$$
$$\leq Ch^2k^{-1}\int_0^k \|\nabla u_t\| ds \|\nabla\theta^1\| \leq Ch\|v\|_3\|\theta^1\|.$$

Since $\theta^0 = 0$ we find

$$\|\theta^1\|_h^2 \leq Ck(k+h)\|v\|_4\|\theta^1\|,$$

which shows our claim. □

We can now show the following complete error estimate.

Theorem 4.1. *Assume that* $\|v_h - v\| \leq Ch\|v\|_4$. *Then we have*

$$\|\mathcal{E}_{h,k}^n v_h - E(t_n)v\| \leq C_T(h+k)\|v\|_4, \quad \textit{for } t_n \leq T.$$

Proof. With $E_k = e^{kB}e^{-kA}$ we have, by Lemmas 4.7 and 4.8,

$$\|(\mathcal{E}_{h,k}R_h v - R_h E(k)v\| \leq \|H_{h,k}(Q_{h,k}R_h - R_h e^{-kA})v\|$$
$$+ \|(H_{h,k}R_h - R_h e^{kB})e^{-kA}v\| \leq Ck(h+k)\|v\|_4$$

The proof is now completed as that of Theorem 3.1, using Lemma 4.6. □

References

[1] E. Hansen and A. Ostermann, *Exponential splitting for unbounded operators.* Math. Comp. **78** (2009), 1485-1496.

[2] W. Hundsdorfer and J. Verwer, *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*, Springer Series in Computational Mathemaics, Springer-Verlag Berline Heidelberg 2003.

[3] T. Jahnke and C. Lubich, *Error bounds for exponential operator splittings.* BIT **40** (2000), 735-744.

[4] S. MacNamara and G. Strang, Gilbert, *Operator splitting.* Splitting methods in communication, imaging, science, and engineering, 95-114, Sci. Comput., Springer, Cham, 2016.

[5] G. Strang, *On the construction and comparison of difference schemes*, SIAM J. Numer. Anal. **5** (1968), 506-517.

[6] V. Thomée and A.S. Vasudeva Murthy, An explicit-impicit splitting method for a convection-diffusion problem, CMAM (to appear).