Confluentes Mathematici, Vol. 3, No. 3 (2011) 495–521

DOI: 10.1142/S1793744211000436

STABILITY, CONVERGENCE AND ORDER OF THE EXTRAPOLATIONS OF THE RESIDUAL SMOOTHING SCHEME IN ENERGY NORM

MAGALI RIBOT

Laboratoire Dieudonné,
Université de Nice-Sophia Antipolis,
Parc Valrose, 06108 Nice Cedex 02, France
ribot@unice.fr

MICHELLE SCHATZMAN

Institut Camille Jordan,
Université Claude Bernard–Lyon 1,
21 Avenue Claude Bernard,
69622 Villeurbanne Cedex, France

Received 24 February 2011

The Residual Smoothing Scheme is a numerical method which consists in preconditioning at each time step the method of lines. In this paper, RSS is defined and analyzed in an abstract linear parabolic case, i.e. for an abstract ordinary differential equation of the form

$$du/dt + Au = 0,$$

with A a self-adjoint non negative operator, and it can be written

$$(U^{n+1} - U^n)/\Delta t + \tau B(U^{n+1} - U^n) + AU^n = 0,$$

where B is a preconditioner of A.

We show that RSS is stable, convergent and of order one in energy norm. We also prove that its kth Richardson's extrapolation is stable and of order k.

Keywords: Residual smoothing scheme; preconditioner; extrapolation; energy norm.

AMS Subject Classification: 65L20, 65B05, 65M12

1. Introduction

The time integration of parabolic systems of equations is dominated by the dilemma between explicit methods, subject to the Courant–Friedrichs–Lewy (CFL) condition and implicit methods which require the use of efficient solvers, and make use of preconditioners.

Preconditioners are often left for the computer science and implementation side of scientific computation; for elliptic problems, preconditioners have been actively studied, the aim being to obtain a better convergence rate for iterative methods.

In the case of time-dependent problems, most of the literature considers that after applying the method of lines (i.e. semidiscretization in time), the preconditioning for time-dependent problems is reduced to preconditioning for space-dependent problems with a matrix $I + \Delta t A$ instead of A, as in Bornemann [3, 4], Mulholland and Sloan [22], Brown and Woodward [6], Gerace *et al.* [13] or Dutt *et al.* [12].

We can find some examples of preconditioners specially proposed to discretize efficiently time-dependent partial differential equations: in [16], Jin and Chan propose a circulant preconditioner to discretize second order hyperbolic equations; in [21], Moore and Dillon compare different preconditioners for parabolic equations discretized by finite element Galerkin method. Then, in [5], Brill and Pinder use a red-black lower block triangular preconditioner for the heat equation discretized by Hermite collocation method. Finally, Mardal, Nilssen and Staff [20, 31] precondition the Runge–Kutta scheme for the heat equation, thanks to Jacobi and Gauss–Seidel type methods. However, when a preconditioning is included in the time integration, the error due to its use is rarely studied in an analytic point of view.

In this paper, we consider that preconditioning is an essential step from the analytical and numerical points of view, and we give a convergence and error analysis for a class of time integration schemes. More precisely, let A be a self-adjoint operator in a Hilbert space; we assume that A is bounded from below and we consider the problem

$$\begin{cases} \dfrac{du}{dt} + Au = f(t), \\ u(0) = u_0. \end{cases} \tag{1.1}$$

Without loss of generality, we may assume that for all u in the domain $D(A)$ of A, we have

$$(Au, u) \geq |u|^2. \tag{1.2}$$

Indeed, if A is bounded from below, there exists C in $\mathbb{R}$ such that

$$(Ax, x) \geq C|x|^2,$$

we set $v = ue^{-\lambda t}$ in (1.1) and we obtain

$$\begin{cases} \dfrac{dv}{dt} + (A + \lambda)v = 0, \\ v(0) = u_0, \end{cases}$$

that is to say a system analogous to (1.1) with $\tilde{A} = A + \lambda$ instead of A. We choose λ such that $C + \lambda \geq 1$ and therefore inequality (1.2) holds for $\tilde{A}$.

Denote by V the closure of $D(A)$ for the pre-Hilbertian norm (Au, u). Assume that B is a self-adjoint unbounded operator which has the same domain as A and which satisfies

$$c^{-1}(Bu, u) \leq (Au, u) \leq c(Bu, u) \tag{1.3}$$

for some strictly positive constant c.

The residual smoothing scheme has been considered in Averbuch *et al.* [1] as an alternative to the backward Euler scheme; it is given by

$$\frac{U_{n+1}-U_n}{\Delta t}+\tau B(U_{n+1}-U_n)+AU_n=F_n, \tag{1.4}$$

where τ is a parameter which can be chosen to enforce stability.

In [1], the authors show that the scheme (1.4) is unconditionally stable for τ large enough if (1.3) holds.

Let us also mention another interesting article by Costa *et al.* [8], where they discretize parabolic equations, thanks to a Fourier collocation method improving the stability of the high frequency component. To do so, they consider an implicit correction to the eigenvalue shifting technique in order to converge to the appropriate steady state. This correction is very similar to scheme (1.4) and possesses the same kind of stability properties.

Now, considering scheme (1.4), we define $P(t)$ as

$$P(t)=\mathbf{1}-t(\mathbf{1}+t\tau B)^{-1}A. \tag{1.5}$$

In this paper, we write $P(t)$ more simply under the following form:

$$P(t)=\mathbf{1}-t\beta(t)A,$$

with

$$\beta(t)=(\mathbf{1}+t\tau B)^{-1}.$$

Thanks to the definition (1.4) of the residual smoothing scheme, if F vanishes, we have

$$U_n=(P(\Delta t))^nU_0.$$

Given any choice of integers $1\le n_1<n_2<\cdots<n_k$, the Richardson extrapolation of $P(t)$ is

$$P_k(t)=\sum_{j=1}^{k}\ell_j^k(0)P(t/n_j)^{n_j}$$

where ℓ_j^k are the elements of the Lagrange interpolation basis with knots $1/n_j$.

We observe that once the scheme defined by P is computed, it is very easy to compute its extrapolations; moreover, if P is of order 1, its Richardson extrapolation P_k is of order k. Therefore, computing the Richardson extrapolations is an easy way to obtain high order schemes provided that they are stable.

In this paper, A and B are abstract operators, but we apply our strategy in [28, 29] to a spectral method preconditioned by a finite elements method. In [28, 29], we prove that the matrices A and B are equivalent in this particular case and we calculate the consistency error, using some expansions of ultra-spherical polynomials proved in [27]. A similar result was proved by Parter [24, 25] using different techniques. The preconditioning of spectral methods by finite differences or finite elements method has been widely studied in the literature, theoretically by Orszag [23] or Haldenwang *et al.* [14] and numerically, for example, by Canuto and Quarteroni [7] or Deville and Mund [9, 10]. However, these articles only deal with elliptic problems.

In [28], we already proved the unconditional stability of the Residual Smoothing Scheme (1.4) and its extrapolations provided that τ is large enough for the usual norm. In this paper, we prove the stability in energy norm, defined by $|x|_A = (x^* Ax)^{1/2}$; let us remark that this norm is finer than the usual norm. This norm is also convenient since in order to study the energy norm of $P(t)$, we have to study the operator $A^{-1/2}P^*APA^{-1/2}$ which is self-adjoint unlike P.

We therefore define an order relation between two operators to make precise the equivalence of two operators as in Eq. (1.3). We then prove that if A is dominated by B and if τ is large enough the Residual Smoothing Scheme is stable, that is to say that the energy norm of P is bounded by 1. We then extend this result to Richardson's extrapolations of P. If furthermore A is equivalent to B, we can prove the convergence of RSS and of its Richardson's extrapolations for τ large enough. To show this result, we use the theory of approximation of continuous semigroups by discrete semigroups described in Kato [17]. Finally, we show that if $A^{-k}B^k$ and $B^{-k}A^k$ are bounded, the scheme defined with the help of P_k is of order k; we use for that purpose the paper of Dia and Schatzman [11] dealing with the algebraic point of view on extrapolation.

In [18], Laevsky obtains weaker results under weaker assumptions, using norms of the form $(u + \tau\Delta tAu, u)^{1/2}$ or $(u + \tau\Delta tBu, u)^{1/2}$. In Laevsky's paper, B does not have to be positive, only dominated by A (in the sense of quadratic forms) and moreover he assumes that $\gamma(\mathbf{1} + \tau\Delta tA) \leq \mathbf{1} + \tau\Delta tB$ for some $\gamma > 0$ and $\Delta t \leq 1$. Laevsky's paper also contains applications to domain decomposition and to fictitious domains methods, but he does not consider the extrapolations of the Residual Smoothing Scheme.

Let us explain the organization of the paper: in Sec. 2, we define an order relation between self-adjoint operators and we study its properties; it enables us to define an equivalence relation between operators, which we use to say that the operator A and its preconditioner B are equivalent. In Sec. 3, we define some regularity spaces related to operators A and B and we give conditions for their equivalence. After these preliminary results, in Sec. 4, we prove the stability of RSS in energy norm and we extend the proof of the stability to the extrapolations of RSS in Sec. 5. Then, in Sec. 6, we study the conditional stability; in Sec. 7, we prove the convergence of RSS and its extrapolations and eventually in Sec. 8, the orders of these schemes. Finally, in Sec. 9, a few perspectives are given.

2. An Order on Self-Adjoint Operators

Let us first define in a very precise way the equivalence of two operators and study the properties of the order relation.

In this paper, we denote by **1** the identity operator in any vector space. We recall that every self-adjoint operator T in a Hilbert space possesses a spectral decomposition

$$T = \int_{\mathbb{R}} \lambda dP(\lambda),$$

where $dP(\lambda)$ is the spectral measure associated to T. We will say that a self-adjoint operator T is positive semi-definite if for all $x \in D(T)$, $x^*Tx \geq 0$. If T is positive semi-definite, the square root of T is defined by

$$\sqrt{T} = \int_{\mathbb{R}} \sqrt{\lambda} dP(\lambda).$$

We define as follows a partial order relation between self-adjoint and bounded from below operators in a Hilbert space H:

$$T_1 \prec T_2 \Rightarrow D(T_2) \subset D(T_1) \quad \text{and} \quad \forall\, x \in D(T_2), \quad x^*T_1x \leq x^*T_2x. \tag{2.1}$$

We will also need a less precise order relation, when $T_2 \geq 0$:

$$T_1 \precsim T_2 \Rightarrow \exists\, r \in (0, \infty) : \ T_1 \prec rT_2. \tag{2.2}$$

If T_2 is positive and no assumption on the sign of the self-adjoint operator T_1 is made, definitions (2.1) and (2.2) still make sense; moreover, we may define for $T_1 \precsim T_2$:

$$[T_1 : T_2] = \sup_{T_2 x \neq 0} \frac{x^*T_1x}{x^*T_2x}. \tag{2.3}$$

With this definition, we always have for $T_2 \geq 0$ and $T_1 \precsim T_2$

$$T_1 \prec [T_1 : T_2] T_2.$$

We define the relations $\succ$ and $\succsim$ to be the opposite relations to $\prec$ and $\precsim$; if T_1 and T_2 are positive self-adjoint operators in H, the relation $T_1 \sim T_2$ means that $T_1 \precsim T_2 \precsim T_1$.

We may relate these equivalence relations to algebraic operations; in particular, if S is a self-adjoint operator which is bounded from below, it is plain that

$$T_1 \prec T_2 \Rightarrow T_1 + S \prec T_2 + S.$$

If S is any bounded operator from a Hilbert space H_1 to H, and if the domain of S^*T_jS for $j = 1, 2$ is defined as $S^{-1}D(T_j)$, we also have:

$$T_1 \prec T_2 \Rightarrow S^*T_1S \prec S^*T_2S. \tag{2.4}$$

The proof is performed through the change of variable $x = Sy$.

Another important fact is the following:

Lemma 2.1. *If T_1 and T_2 are positive self-adjoint and injective, then*

$$T_1 \prec T_2 \Rightarrow T_2^{-1} \prec T_1^{-1}. \tag{2.5}$$

Proof. This can be deduced from the proof of Theorem VI.2.21 in Kato's book [17]. □

Observe that if $T_1 \prec T_2$ then for any powers $\alpha \in]0,1[$, $T_1^\alpha \prec T_2^\alpha$. Indeed a formula of Balakrishnan in [2] which is given in Yosida's book [32] gives the representation of T^α:

$$x \in D(T) \Rightarrow T^\alpha x = \frac{\sin(\alpha\pi)}{\pi} \int_0^\infty \lambda^{\alpha-1}(\lambda \mathbf{1} + T)^{-1} T x d\lambda.$$

The relation

$$(\lambda \mathbf{1} + T)^{-1} T = \mathbf{1} - \lambda(\lambda \mathbf{1} + T)^{-1} \tag{2.6}$$

is classical; thus, we infer from relation (2.6) and Lemma 2.1 that

$$(\lambda \mathbf{1} + T_1)^{-1} T_1 \prec (\lambda \mathbf{1} + T_2)^{-1} T_2;$$

therefore, it is plain that $T_1^\alpha \prec T_2^\alpha$.

However, $T_1 \prec T_2$ does not imply $T_1^n \prec T_2^n$ for all n in $\mathbb{N}$; a counterexample is, for instance

$$T_1 = \begin{pmatrix} 2\varepsilon & 0 \\ 0 & \dfrac{2}{\varepsilon} \end{pmatrix}, \quad T_2 = \begin{pmatrix} \varepsilon & \dfrac{1}{2} \\ \dfrac{1}{2} & \dfrac{1}{\varepsilon} \end{pmatrix}.$$

The reader will check that for all positive ε, $T_1 \succ T_2$, while for all small enough ε, it is not true that $T_1^2 \succ T_2^2$. However, if the self-adjoint, positive operators T_1 and T_2 commute, and in particular if one of them is scalar, the conclusion is true and this can be checked simply with the help of the spectral theorem.

3. Some Preliminary Results

Let us introduce the domains of the powers of A and B; we define, as in [26], $\mathcal{H}_A^n$ for $n \in \mathbb{Z}$ as

$$\begin{aligned} \mathcal{H}_A^0 &= H, \\ \mathcal{H}_A^n &= A^{-n/2} H = D(A^{n/2}), \qquad \text{for } n \in \mathbb{N}^* \end{aligned}$$

and

$$\mathcal{H}_A^n = (\mathcal{H}_A^{-n})^*, \qquad \text{for } n \in \mathbb{Z} \backslash \mathbb{N},$$

where the star stands for the dual space.

The norm over $\mathcal{H}_A^n$ is defined by

$$|u|_{\mathcal{H}_A^n} = |A^{n/2} u|_0.$$

The same definitions hold replacing A by B. We will write for simplicity

$$\mathcal{H}^0 = \mathcal{H}_A^0 = \mathcal{H}_B^0.$$

We have the following inclusions

$$\cdots \mathcal{H}_A^{-n} \supset \cdots \mathcal{H}_A^{-2} \supset \mathcal{H}_A^{-1} \supset \mathcal{H}_A^0 = H \supset \mathcal{H}_A^1 \supset \mathcal{H}_A^2 \cdots \supset \mathcal{H}_A^n \cdots$$

and

$$\cdots \mathcal{H}_B^{-n} \supset \cdots \mathcal{H}_B^{-2} \supset \mathcal{H}_B^{-1} \supset \mathcal{H}_B^0 = H \supset \mathcal{H}_B^1 \supset \mathcal{H}_B^2 \cdots \supset \mathcal{H}_B^n \cdots$$

and for $s > t$, $\mathcal{H}_A^s$ (respectively $\mathcal{H}_B^s$) is dense in $\mathcal{H}_A^t$ (respectively $\mathcal{H}_B^t$).

Indeed, if n is even, $n = 2p$, for the density of $\mathcal{H}_A^n$ in $\mathcal{H}^0$, it suffices to consider

$$x(t) = \frac{p!}{t^p} \int_0^p \int_0^{t_1} \cdots \int_0^{t_{p-1}} e^{-t_p A} x \, dt_p \cdots dt_1$$

which belongs to $D(A^p)$ and converges to $x(0) = x$ as t tends to zero. Moreover, if n is odd, $\mathcal{H}_A^n$ contains $\mathcal{H}_A^{n+1}$ which is dense in $\mathcal{H}^0$ from the previous case.

Moreover, $A \sim B$ implies the equality $\mathcal{H}_A^1 = \mathcal{H}_B^1$ and the equivalence of the norms $|A^{1/2}x|_0$ and $|B^{1/2}x|_0$ over $\mathcal{H}_A^1$. Let us give now some conditions for the isomorphism between $\mathcal{H}_A^n$ and $\mathcal{H}_B^n$.

Lemma 3.1. *For $n \in \mathbb{N}$, the following propositions are equivalent*:

(1) $B^{-n/2}A^{n/2}$ *and* $A^{-n/2}B^{n/2}$ *are bounded in* $\mathcal{L}(\mathcal{H}^0)$,
(2) $\mathcal{H}_A^n$ *and* $\mathcal{H}_B^n$ *are isomorphic and their norms are equivalent*,
(3) $\mathcal{H}_A^{-n}$ *and* $\mathcal{H}_B^{-n}$ *are isomorphic and their norms are equivalent.*

Proof. Observe that *a priori* $A^{-n/2}B^{n/2}$ is defined on $\mathcal{H}_B^n$; hypothesis (1) states that this operator is bounded for the operator norm subordinate to the norm of $\mathcal{H}^0$; consequently, it admits a unique extension to all $\mathcal{H}^0$.

It is immediate by duality that (2) is equivalent to (3). Let us prove that hypothesis (1) implies (3).

Assume $y \in \mathcal{H}^0$ and $x = A^{-n/2}y \in \mathcal{H}_A^n \subset \mathcal{H}^0$. Since

$$\|B^{-n/2}A^{n/2}\|_{\mathcal{L}(\mathcal{H}^0)} \leq C_n,$$

we have

$$|B^{-n/2}y|_0 = |B^{-n/2}A^{n/2}x|_0 \leq C_n|x|_0 = C_n|A^{-n/2}y|_0,$$

that is to say

$$|y|_{\mathcal{H}_B^{-n}} \leq C_n|y|_{\mathcal{H}_A^{-n}};$$

and thus by density of $\mathcal{H}^0$ in $\mathcal{H}_A^{-n}$

$$\mathcal{H}_A^{-n} \subset \mathcal{H}_B^{-n}.$$

The opposite inclusion is obtained by exchanging A and B.

To complete the proof, let us prove that hypothesis (3) implies (1): $A^{n/2}$ is an isomorphism from $\mathcal{H}^0$ to $\mathcal{H}_A^{-n}$ and $B^{n/2}$ is an isomorphism from $\mathcal{H}^0$ to $\mathcal{H}_B^{-n}$, thus using hypothesis (3), $B^{-n/2}A^{n/2}$ is an isomorphism from $\mathcal{H}^0$ to $\mathcal{H}^0$ and consequently the same holds true for $A^{-n/2}B^{n/2}$. □

Lemma 3.2. *If for an integer n, $\mathcal{H}^n_A$ is isomorphic to $\mathcal{H}^n_B$, then for all $s \in \{-|n|, \dots, |n|\}$, $\mathcal{H}^s_A$ is isomorphic to $\mathcal{H}^s_B$.*

Proof. This is a result of interpolation; see, for example Definition 2.1 and Remark 2.3 of [19]. □

Define

$$\beta(t) = (\mathbf{1} + t\tau B)^{-1} \tag{3.1}$$

such that $P(t) = \mathbf{1} - t\beta(t)A$.

Let us prove that it converges strongly to $\mathbf{1}$ as t tends to zero.

Lemma 3.3. *If $B^{-n/2}A^{n/2}$ and $A^{-n/2}B^{n/2}$ are bounded in $\mathcal{L}(\mathcal{H}^0)$, then for all $s \in \{-n, \dots, 0\}$,*

$$\beta(t) \to \mathbf{1} \quad \textit{strongly in } \mathcal{H}^s_A \textit{ as } t \to 0^+.$$

Proof. Since $\beta(t) \in \mathcal{L}(\mathcal{H}^s_B, \mathcal{H}^{s+2}_B)$ and thanks to Lemma 3.1, we see that $\beta(t) \in \mathcal{L}(\mathcal{H}^s_A, \mathcal{H}^s_A)$.

We know already that $\beta(t)$ converges strongly in $\mathcal{H}^0$ to $\mathbf{1}$, i.e. for all $v \in \mathcal{H}^0$, $\beta(t)v \to v$ in $\mathcal{H}^0$ and that $\mathcal{H}^0$ is dense in $\mathcal{H}^s_A$; to conclude by density it suffices to prove that $\|\beta(t)\|_{\mathcal{L}(\mathcal{H}^s_A)}$ is bounded. This is clearly true, since by virtue of Lemma 3.2,

$$\|\beta(t)\|_{\mathcal{L}(\mathcal{H}^s_A)} \le C_n \|\beta(t)\|_{\mathcal{L}(\mathcal{H}^s_B)};$$

the operator norm of $\beta(t)$ in $\mathcal{L}(\mathcal{H}^s_B)$ is equal to its operator norm in $\mathcal{L}(\mathcal{H}^0)$, giving therefore the conclusion. □

4. Stability of RSS

Let us prove the stability of the Residual Smoothing Scheme defined in Eq. (1.4).

We will systematically write $a = A^{1/2}$ and $b = B^{1/2}$.

Define the energy norm by

$$|x|_A = (x^* A x)^{1/2}.$$

This norm is the above defined norm over $\mathcal{H}^1_A = D(a)$. The corresponding operator norm is

$$\|L\|^2_A = \sup_{x \in \mathcal{H}^1_A} \frac{x^* L^* A L x}{x^* A x}.$$

It is clear that the energy operator norm of L is bounded iff the ordinary operator norm $\|a^{-1}L^*ALa^{-1}\|$ is bounded, where the double norm $\| \quad \|$ denotes from now on the operator norm $\mathcal{L}(H)$.

We remark that unlike the operator P, the operator $a^{-1}P^*APa^{-1}$ is self-adjoint, which simplifies a lot the proof.

We have the following stability result on (1.4):

Theorem 4.1. *Let A and B be positive definite self-adjoint operators in H satisfying $A \precsim B$ and let $P(t)$ be defined by* (1.5)*. Then, for τ larger than $[A\!:\!B]/2$, the energy norm of $P(t)$ is at most equal to* 1.

Proof. The energy operator norm of $P(t)$ is

$$\|P(t)\|_A = \|a^{-1}P(t)^*AP(t)a^{-1}\|^{1/2}$$

and a straightforward computation gives

$$a^{-1}P(t)^*AP(t)a^{-1} = \mathbf{1} - 2ta\beta a + t^2\left(a\beta a\right)^2. \tag{4.1}$$

It is convenient to let

$$Q(t) = a^{-1}P(t)^*AP(t)a^{-1}. \tag{4.2}$$

It is clear that $Q(t)$ is semi-definite positive. We see that $Q(t) \prec \mathbf{1}$ iff

$$2ta\beta a \succ t^2\left(a\beta a\right)^2$$

and this will be true if

$$2 \times \mathbf{1} \succ ta\beta a. \tag{4.3}$$

Let us check that for all $t > 0$ the following inequality holds:

$$ta\beta a \prec \frac{[A\!:\!B]}{\tau}\mathbf{1}. \tag{4.4}$$

We have indeed the inequalities

$$tA \prec [A\!:\!B]tB \prec (\mathbf{1} + t\tau B)\frac{[A\!:\!B]}{\tau};$$

if we apply (2.5), we find that

$$\frac{\tau}{[A\!:\!B]}\beta \prec t^{-1}A^{-1}. \tag{4.5}$$

We multiply (4.5) on the left and on the right by a and we find immediately that (4.4) holds. Therefore, if

$$\tau \geq [A\!:\!B]/2,$$

the inequality $Q(t) \prec \mathbf{1}$ will be satisfied, proving thus the stability of $P(t)$ in energy norm. $\square$

5. The General Proof of Stability in Energy Norm of the Extrapolation of RSS

Let us now extend the result of the previous section to the Richardon's extrapolations of RSS; we need before to show some algebraic lemmas.

5.1. *A preliminary inequality*

Given k distinct strictly positive integers $1 \leq n_1 < n_2 < \cdots < n_j < \cdots < n_k$, we define the coefficients of Richardson's extrapolation as follows: let ℓ_j^k be the Lagrange basis relative to the nodes $1/n_j$, $1 \leq j \leq k$:

$$\ell_j^k(t) = \prod_{\{i:i\neq j\}} \frac{t - 1/n_i}{(1/n_j) - (1/n_i)}. \tag{5.1}$$

Some well-known choices for these nodes are

- the harmonic sequence $n_j = j$,
- the Romberg sequence $n_j = 2^j$,
- the Bulirsch sequence

$$1, 2, 3, 4, 6, 8, 12, 16, \ldots, 2^j, \frac{3}{2}2^j, 2^{j+1}, \frac{3}{2}2^{j+1}, \ldots.$$

By interpolation of $1, t, \ldots, t^{k-1}$, we have the equalities:

$$\begin{aligned} &\sum_{j=1}^{k} \ell_j^k(t) = 1, \\ \forall\, p = 1, \ldots, k-1, \quad &\sum_{j=1}^{k} \ell_j^k(t) \frac{1}{n_j^p} = t^p, \end{aligned} \tag{5.2}$$

from which we infer taking $t = 0$

$$\forall\, p = 0, \ldots, k-1, \quad \sum_{j=1}^{k} \frac{\ell_j^k(0)}{n_j^p} = \delta_{0p}. \tag{5.3}$$

The following function

$$\phi_k(s,t) = \sum_{j=1}^{k} \frac{\ell_j^k(t)}{1 + s/n_j}$$

will play an essential role in our analysis; $\phi_k(s,\cdot)$ interpolates the function $f_s : t \mapsto 1/(1+st)$ at the points $1/n_j$, $1 \leq j \leq k$.

Lemma 5.1. *The function $s \mapsto \phi_k(s,0)$ is strictly positive over $\mathbb{R}^+$.*

Proof. For any function g, denote by $g[x_1, \ldots, x_n]$ the divided difference of the function g at the knots $x_1, \ldots, x_n$.

We use Newton's form of interpolation:

$$\phi_k(s,t) = f_s(1/n_1) + f_s[1/n_1, 1/n_2](t - 1/n_1) + \cdots \\ + f_s[1/n_1, 1/n_2, \ldots, 1/n_k](t - 1/n_1)(t - 1/n_2)\cdots(t - 1/n_{k-1}).$$

In particular,

$$\phi_k(s,0) = f_s(1/n_1) - \frac{f_s[1/n_1, 1/n_2]}{n_1} + \frac{f_s[1/n_1, 1/n_2, 1/n_3]}{n_1 n_2} + \cdots \\ + \frac{(-1)^{k-1} f_s[1/n_1, 1/n_2, \ldots, 1/n_k]}{n_1 n_2 \cdots n_{k-1}}. \tag{5.4}$$

If S^j is the simplex

$$S^j = \{x \in (\mathbb{R}^+)^j : x_1 + \cdots + x_j \leq 1\},$$

the divided differences are given by the integral representation

$$f_s[a_1, \ldots, a_{j+1}] = \int_{S^j} f_s^{(j)}(a_1 + t_1(a_2 - a_1) + \cdots + t_j(a_{j+1} - a_j)). \tag{5.5}$$

But in our case,

$$f_s^{(j)}(t) = \frac{j!(-s)^j}{(1+st)^{j+1}}. \tag{5.6}$$

The term $(-1)^{j-1} f_s[1/n_1, \ldots, 1/n_j]$ involves $f_s^{(j-1)}$ in the integral representation; it is therefore positive, and the lemma is proved. □

We need another algebraic fact:

Lemma 5.2. *For all $k = 1, 2, \ldots$ the following identity holds*:

$$\sum_{j=1}^{k} n_j \ell_j^k(0) = \sum_{j=1}^{k} n_j. \tag{5.7}$$

Proof. Write

$$T_k = \sum_{j=1}^{k} n_j \ell_j^k(0).$$

We infer from formula (5.1) that $\ell_j^k(0)$ is given as

$$\ell_j^k(0) = (-n_j)^{k-1} \prod_{\{i:1\leq i\leq k, i\neq j\}} \frac{1}{n_i - n_j}. \tag{5.8}$$

Therefore, we have the relation

$$T_k - T_{k-1} = n_k \ell_k^k(0) + \sum_{j=1}^{k-1} A_j$$

with

$$A_j = (-n_j)^{k-1}\left(n_j \prod_{\{i:1\le i\le k,\, i\ne j\}} \frac{1}{n_i - n_j} + \prod_{\{i:1\le i\le k-1,\, i\ne j\}} \frac{1}{n_i - n_j}\right).$$

We remark that the factor of $(-n_j)^{k-1}$ in A_j contains $k-2$ common factors; therefore, it is equal to

$$\left(\frac{n_j}{n_k - n_j} + 1\right) \prod_{\{i:1\le i\le k-1,\, i\ne j\}} \frac{1}{n_i - n_j},$$

and therefore

$$A_j = (-n_j)^{k-1} n_k \prod_{\{i:1\le i\le k,\, i\ne j\}} \frac{1}{n_i - n_j}$$

and with the help of (5.8),

$$A_j = n_k \ell_j^k(0);$$

therefore,

$$T_k - T_{k-1} = n_k \sum_{j=1}^{k} \ell_j^k(0)$$

which concludes the proof, thanks to (5.3). □

Lemma 5.1 says that the function $\psi_k(s) = \phi_k(s,0)$ is non-negative on $\mathbb{R}^+$; Lemma 5.2 enables us to find an equivalent of ψ_k at infinity:

$$\psi_k(s) \sim \sum_{j=1}^{k} \frac{n_j \ell_j^k(0)}{s} = \frac{\sum_{j=1}^{k} n_j}{s},$$

and therefore, the following corollary holds:

Corollary 5.1. *There exist $c_k > 0$ and $C_k > 0$ such that for all $s > 0$:*

$$\frac{c_k}{1 + s/n_k} \le \sum_{j=1}^{k} \frac{\ell_j^k(0)}{1 + s/n_j} \le \frac{C_k}{1 + s/n_k}. \tag{5.9}$$

5.2. *Proof of the stability of the extrapolation*

We introduce the notation

$$\beta_j(t) = \beta(t/n_j).$$

The purpose of this section is to show that the extrapolation of RSS is unconditionally stable in energy norm for large enough values of τ.

Theorem 5.1. *Let A and B be positive definite self-adjoint operators in H satisfying $A \precsim B$. For all $k \in \mathbb{N}$ and for any choice of integers $1 \leq n_1 < n_2 < \cdots < n_k$, there exists $\tau_0 > 0$ such that for all $\tau \geq \tau_0$ the following estimate holds*:

$$\forall\, t > 0, \quad \|P_k(t)\|_A \leq 1.$$

Proof. We expand $P(t/n_j)^{n_j}$ according to the binomial formula. Therefore, if $\binom{n}{p}$ is set equal to zero for $p < 0$ or $p > n$, we have

$$P_k(t) = \sum_{j=1}^{k} \ell_j^k(0) \sum_{i=0}^{n_k} (-1)^i \binom{n_j}{i} \left(\frac{t}{n_j}\beta_j A\right)^i.$$

If we define

$$p_0(t) = \mathbf{1}, \tag{5.10a}$$

$$p_1(t) = \sum_{j=1}^{k} \ell_j^k(0)\beta_j A, \tag{5.10b}$$

and for all $i = 2, \ldots, n_k$

$$p_i(t) = \sum_{\{j:1\leq j\leq k, n_j\geq i\}} \binom{n_j}{i} \ell_j^k(0) \frac{(\beta_j A)^i}{n_j^i}, \tag{5.10c}$$

the expression of P_k can be rewritten

$$P_k(t) = \sum_{i=0}^{n_k} (-t)^i p_i(t). \tag{5.11}$$

The energy norm of the operator P_k is equal to

$$\left\| a^{-1} \sum_{i=0}^{n_k} (-t)^i p_i(t)^* A \sum_{l=0}^{n_k} (-t)^l p_l(t) a^{-1} \right\|;$$

the operator inside the norm symbol can be rewritten as

$$\mathbf{1} - 2\sum_{j=1}^{k} t\ell_j^k(0) a\beta_j a + \sum_{i+l\geq 2} (-t)^{i+l} a^{-1} p_i(t)^* A p_l(t) a^{-1}.$$

As in the proof of Theorem 4.1, this operator is semi-definite positive. Therefore, the stability in energy norm will hold if

$$0 \prec 2\sum_{j=1}^{k} t\ell_j^k(0) a\beta_j a - \sum_{i+l\geq 2} (-t)^{i+l} a^{-1} p_i(t)^* A p_l(t) a^{-1}.$$

We can deduce from Eq. (5.9) that

$$c_k a\beta_k a \prec \sum_{j=1}^{k} \ell_j^k(0) a\beta_j a \prec C_k a\beta_k a.$$

Therefore, it suffices to find values of τ for which

$$2tc_k a\beta_k a \succ \sum_{i+l\geq 2} (-t)^{i+l} a^{-1} p_i(t)^* A p_l(t) a^{-1}. \tag{5.12}$$

Condition (5.12) holds if

$$2c_k \times \mathbf{1} \succ -\sum_{i+l\geq 2} (-t)^{i+l-1} \beta_k^{-1/2} A^{-1} p_i(t)^* A p_l(t) A^{-1} \beta_k^{-1/2}. \tag{5.13}$$

Therefore, we have to estimate the operator norm of L_{iljr} given by

$$L_{iljr} = C_{iljr} M_{iljr}$$

with

$$C_{iljr} = \frac{\ell_j^k(0)\ell_r^k(0)}{n_j^i n_r^l} \binom{n_j}{i} \binom{n_r}{l},$$

$$M_{iljr} = (-t)^{i+l-1} \beta_k^{-1/2} A^{-1} (A\beta_j)^i A (\beta_r A)^l A^{-1} \beta_k^{-1/2}.$$

The terms M_{iljr} can be rewritten for $\min(i,l) \geq 1$

$$M_{iljr} = (-t)^{i+l-1} \beta_k^{-1/2} \beta_j^{1/2} (\beta_j^{1/2} A \beta_j^{1/2})^{i-1} \beta_j^{1/2} A \beta_r^{1/2} (\beta_r^{1/2} A \beta_r^{1/2})^{l-1} \times \beta_r^{1/2} \beta_k^{-1/2}.$$

We infer from the obvious inequality

$$t\tau A \prec [A\!:\!B](\mathbf{1} + t\tau B)$$

that

$$t\tau \beta^{1/2} A \beta^{1/2} \prec [A\!:\!B]\mathbf{1}.$$

Therefore, we have the estimate

$$\beta_j^{1/2} A \beta_j^{1/2} \prec \frac{[A\!:\!B] n_j}{\tau t} \mathbf{1}; \tag{5.14}$$

on the other hand, by the spectral theorem,

$$\|\beta_k^{-1/2} \beta_j^{1/2}\| \leq 1. \tag{5.15}$$

We deduce from the inequality

$$\|\beta_j^{1/2} A \beta_r^{1/2}\| \leq \|\beta_j^{1/2} A \beta_j^{1/2}\|^{1/2} \|\beta_r^{1/2} A \beta_r^{1/2}\|^{1/2}$$

the following estimate

$$\|\beta_j^{1/2} A \beta_r^{1/2}\| \leq \frac{[A\!:\!B]\sqrt{n_j n_r}}{\tau t}. \tag{5.16}$$

We put together the estimates (5.14), (5.15) and (5.16) and we find that

$$\|M_{iljr}\| \leq \left(\frac{[A\!:\!B]}{\tau} \right)^{i+l-1} n_j^{i-1/2} n_r^{l-1/2}$$

and therefore that

$$\|L_{iljr}\| \le \left(\frac{[A:B]}{\tau}\right)^{i+l-1} n_j^{i-1/2} n_r^{l-1/2} |C_{iljr}|.$$

If i vanishes, the expression for M_{iljr} is even simpler:

$$M_{0l0r} = (-t)^{l-1} \beta_k^{-1/2} (\beta_r A)^l A^{-1} \beta_k^{-1/2},$$

and the norm of L_{0l0r} can be estimated by

$$\|L_{0l0r}\| \le \left(\frac{[A:B]}{\tau}\right)^{l-1} n_r^{l-1} |C_{0l0r}|.$$

Let us write

$$\nu_{il} = \sum_{\substack{\{j:n_j \ge i\}\\ \{r:n_r \ge l\}}} n_j^{i-1/2} n_r^{l-1/2} |C_{iljr}| \quad \text{and} \quad \nu_{0l} = \nu_{l0} = \sum_{\{r:n_r \ge l\}} n_r^{l-1} |C_{0l0r}|.$$

There is a finite number of terms to estimate, and therefore, a sufficient condition for (5.12) to hold is

$$2c_k \ge \sum_{\substack{0 \le i \le k\\ 0 \le l \le k\\ i+l \ge 2}} \nu_{il} \left(\frac{[A:B]}{\tau}\right)^{i+l-1}.$$

It is clear that for large enough values of τ, (5.12) is satisfied, which completes the proof of the theorem. $\square$

6. Conditional Stability

If the operator A is bounded and in particular in the finite dimension case, we may be interested by conditional stability results. We start with an improvement of Theorem 4.1.

Lemma 6.1. *Let A and B be self-adjoint positive definite operators such that $A \precsim B$ and A is bounded. Then for all $\tau > 0$, $\|P(t)\|_A \le 1$ if*

$$t\|A\| \left(1 - \frac{2\tau}{[A:B]}\right) \le 2. \tag{6.1}$$

Proof. As in (4.3), it suffices to find a condition under which

$$ta\beta a \prec 2 \times \mathbf{1};$$

using Lemma 2.1 and Eq. (2.4), it is realized provided that

$$tA \prec 2(\mathbf{1} + t\tau B). \tag{6.2}$$

Observe that

$$A \prec [A:B]B,$$

and therefore (6.2) will hold provided that

$$tA \prec 2(\mathbf{1} + t\tau[A:B]^{-1}A)$$

which is true if

$$t(1 - 2\tau[A:B]^{-1})A \prec 2 \times \mathbf{1}$$

and the conclusion is clear. □

Relation (6.1) is typical of a conditional stability condition, since it could have been obtained for the fully explicit scheme, corresponding to $\tau = 0$,

$$u_{n+1} = u_n - tAu_n,$$

where the stability condition reads

$$\|\mathbf{1} - tA\| \le 1;$$

this inequality is satisfied under condition (6.1).

Let us prove now the conditional stability for the Richardson's extrapolation of RSS.

Lemma 6.2. *Under the assumption of Lemma 6.1, for all sequence of distinct positive integers $n_1 < n_2 < \cdots < n_k$, if A and B are positive semi-definite operators, A is bounded and $A \precsim B$, there exists $\varepsilon_k > 0$ such that*

$$t\|A\|\left(1 - \frac{\tau\varepsilon_k}{n_k[A:B]}\right) \le \varepsilon_k$$

implies $\|P_k(t)\|_A \le 1$.

Proof. As in the proof of Theorem 5.1, and with the same notations, it suffices to prove, using (5.13),

$$\sum_{\substack{0\le i\le k\\0\le l\le k\\i+l\ge 2}} \sum_{\substack{\{j:n_j\ge i\}\\\{r:n_r\ge l\}}} |C_{iljr}|\|M_{iljr}\| \le 2c_k.$$

We observe that

$$\left(1 + \frac{t\tau\|A\|}{[A:B]}\right)A \prec \|A\|(\mathbf{1} + t\tau B)$$

and therefore

$$A \prec \frac{\|A\|}{1 + t\tau\|A\|[A:B]^{-1}}(\mathbf{1} + t\tau B)$$

which implies immediately

$$\beta^{1/2}A\beta^{1/2} \prec \frac{\|A\|}{1 + t\tau\|A\|[A:B]^{-1}}\mathbf{1}.$$

Therefore, we have now the estimates

$$\|M_{iljr}\| \leq \left(\frac{t\|A\|}{1+t\tau\|A\|n_j^{-1}[A:B]^{-1}}\right)^{i-1/2}\left(\frac{t\|A\|}{1+t\tau\|A\|n_r^{-1}[A:B]^{-1}}\right)^{l-1/2}$$

and

$$\|M_{0l0r}\| \leq \left(\frac{t\|A\|}{1+t\tau\|A\|n_r^{-1}[A:B]^{-1}}\right)^{l-1}.$$

We write

$$\nu_{i,l} = \sum_{\substack{\{j:n_j\geq i\}\\ \{r:n_r\geq l\}}} |C_{iljr}| \quad \text{and} \quad \nu_{l,0} = \nu_{0,l} = \sum_{\{r:n_r\geq l\}} |C_{0l0r}|.$$

Therefore it suffices to have the estimate

$$\sum_{i+l\geq 2} \nu_{il}\left(\frac{t\|A\|}{1+t\tau\|A\|n_k^{-1}[A:B]^{-1}}\right)^{i+l-1} \leq 2c_k.$$

The polynomial

$$\sum_{i+l\geq 2} \nu_{il}x^{i+l-1}$$

vanishes at 0; if we denote by ε_k the smallest positive real for which it takes the value $2c_k$, we see that $\|P_k(t)\|_A$ is at most equal to 1 provided that

$$t\|A\| \leq \varepsilon_k\left(1+\frac{t\tau\|A\|}{n_k[A:B]}\right). \qquad \square$$

7. Convergence of RSS in Energy Norm and General Proof of the Convergence of the Extrapolation of RSS

We first prove the convergence of the residual smoothing scheme:

Theorem 7.1. *Assume $A \sim B$. Then for $\tau \geq [A:B]/2$, $P(t_n)^n$ converges strongly to e^{-tA} in energy norm, as n tends to $+\infty$, t_n tends to 0 and nt_n tends to t.*

Proof. We use Theorem IX.3.6 of Kato [17] which describes the theory of approximation of continuous semigroups by discrete semigroups.

Define indeed

$$A_n = \frac{1}{t_n}(\mathbf{1}-P(t_n)).$$

Theorem 4.1 implies that

$$U_n^k = (\mathbf{1}-t_nA_n)^k = P(t_n)^k$$

is norm bounded by 1 for all integers k and n. Therefore, it suffices to find a complex number ζ such that

$$(A_n+\zeta)^{-1} \xrightarrow{s} (A+\zeta)^{-1} \text{ in } \mathcal{H}_A^1$$

to conclude the proof of the theorem.

We choose $\zeta = 1$ and we prove first that for $f \in \mathcal{H}_A^1 = \mathcal{H}_B^1$, it is possible to find a solution $u(t)$ of

$$\left(\mathbf{1} + \frac{\mathbf{1} - P(t)}{t}\right) u(t) = f.$$

By definition,

$$\mathbf{1} + \frac{\mathbf{1} - P(t)}{t} = \mathbf{1} + \beta(t)A,$$

and therefore it suffices to find a solution of

$$u(t) + (t\tau B + A)u(t) = f + t\tau Bf. \tag{7.1}$$

Thanks to our assumptions on f, A and B, $f + t\tau Bf$ belongs to $\mathcal{H}_A^{-1} = \mathcal{H}_B^{-1}$, and Lax–Milgram's lemma gives a unique solution of (7.1); moreover, this solution belongs to $\mathcal{H}_A^1$.

We may rewrite (7.1) as

$$(\mathbf{1} + t\tau B + A)(u(t) - f) = -Af. \tag{7.2}$$

If we multiply scalarly (7.2) by $u(t) - f$, we obtain the estimate

$$|u(t) - f|^2 + |u(t) - f|_A^2 \le |f|_A |u(t) - f|_A$$

which implies immediately

$$|u(t)|_A \le 2|f|_A.$$

For any sequence t_n decreasing toward 0, we select a subsequence, still denoted by t_n, such that

$$u(t_n) \rightharpoonup u_0 \quad \text{weakly in } \mathcal{H}_A^1.$$

Clearly, $t_n \tau B u(t_n)$ tends to zero weakly in $\mathcal{H}_A^{-1}$ and $t_n \tau Bf$ tends to zero strongly in $\mathcal{H}_A^{-1}$. Therefore, in the limit, we must have

$$u_0 + Au_0 = f \tag{7.3}$$

since $Au(t_n)$ converges to Au_0 weakly in $\mathcal{H}_A^{-1}$.

Lax–Milgram's lemma shows that there exists a unique u_0 satisfying (7.3). In order to show the strong convergence of $u(t_n)$ to u_0 in $\mathcal{H}_A^1$, we multiply (7.1) by $u(t_n)^*$, getting thus the identity

$$|u(t_n)|^2 + t_n \tau u(t_n)^* B u(t_n) + u(t_n)^* A u(t_n) = u(t_n)^* f + t_n \tau u(t_n)^* Bf. \tag{7.4}$$

On the one hand, we infer from (7.4)

$$\begin{aligned}\overline{\lim_{n\to+\infty}}|u(t_n)|^2 + u(t_n)^*Au(t_n) &\le \overline{\lim_{n\to+\infty}}\, u(t_n)^*f + t_n\tau u(t_n)^*Bf\\ &\le u_0^*f = |u_0|^2 + u_0^*Au_0.\end{aligned}$$

On the other hand, general theorems imply

$$\underline{\lim_{n\to+\infty}}\,|u(t_n)|^2 + u(t_n)^*Au(t_n) \ge |u_0|^2 + u_0^*Au_0.$$

Therefore,

$$\lim_{n\to+\infty}|u(t_n)|^2 + u(t_n)^*Au(t_n) = |u_0|^2 + u_0^*Au_0,$$

proving the desired strong convergence.

Moreover, since the sequence t_n was arbitrary and u_0 is unique we have the stronger result

$$\lim_{t\to 0}|u(t) - u_0|_A = 0. \qquad \square$$

We will prove now the convergence of the extrapolation P_k of RSS.

Theorem 7.2. *If $A \sim B$, for all $k \in \mathbb{N}$, for any choice of integers $1 \le n_1 < n_2 < \cdots < n_k$, there exists $\tau_0 > 0$ such that for all $\tau \ge \tau_0$, $P_k(t_n)^n$ converges strongly to e^{-tA} in energy norm, as n tends to $+\infty$, t_n tends to 0 and nt_n tends to t.*

Proof. As in Theorem 7.1, we will use Theorem IX.3.6 of [17]. Theorem 5.1 yields that $\{\|P_k(t_n)\|_A\}_n$ is bounded uniformly by 1 and we have to prove that for all f in $\mathcal{H}_A^1$,

$$\left(\mathbf{1} + \frac{\mathbf{1} - P_k(t)}{t}\right)^{-1} f \xrightarrow[t\to 0]{s} (\mathbf{1} + A)^{-1} f.$$

We show first that for $f \in \mathcal{H}_A^1$, we can find a solution $u(t)$ of

$$\left(\mathbf{1} + \frac{\mathbf{1} - P_k(t)}{t}\right) u(t) = f. \tag{7.5}$$

Using Eq. (5.11), we find that

$$\mathbf{1} + \frac{\mathbf{1} - P_k(t)}{t} = \mathbf{1} + \sum_{i=1}^{n_k}(-t)^{i-1} p_i(t)$$

and therefore after applying A to Eq. (7.5), we can rewrite this equation as

$$Au(t) + Ap_1(t)u(t) + \sum_{i=2}^{n_k}(-t)^{i-1} Ap_i(t)u(t) = Af. \tag{7.6}$$

In order to apply Lax–Milgram's lemma, let us show first that there exists $\tau_0 > 0$ such that for all $\tau \ge \tau_0$, for all $t > 0$,

$$Ap_1(t) + \sum_{i=2}^{n_k}(-t)^{i-1} Ap_i(t) \succ 0, \tag{7.7}$$

that is to say, using Eq. (5.10) and multiplying Eq. (7.7) on the left and on the right by A^{-1}, that, for τ large enough,

$$\sum_{j=1}^{k} \ell_j^k(0)\beta_j \succ -\sum_{i=2}^{n_k}(-t)^{i-1} \sum_{\{j:n_j\geq i\}} \binom{n_j}{i} \frac{\ell_j^k(0)}{n_j^i} \beta_j^{1/2}(\beta_j^{1/2}A\beta_j^{1/2})^{i-1}\beta_j^{1/2}.$$

Since, from Eq. (5.9),

$$\sum_{j=1}^{k} \ell_j^k(0)\beta_j \succ c_k\beta_k,$$

it suffices to show that for τ large enough,

$$c_k \times \mathbf{1} \succ -\sum_{i=2}^{n_k}(-t)^{i-1} \sum_{\{j:n_j\geq i\}} \binom{n_j}{i} \frac{\ell_j^k(0)}{n_j^i} \beta_k^{-1/2}\beta_j^{1/2}(\beta_j^{1/2}A\beta_j^{1/2})^{i-1}\beta_j^{1/2}\beta_k^{-1/2}. \tag{7.8}$$

Let us write

$$\mu_i = \sum_{\{j:n_j\geq i\}} \binom{n_j}{i} \frac{|\ell_j^k(0)|}{n_j},$$

which is positive and

$$Q(x) = \sum_{i=2}^{n_k} \mu_i x^{i-1},$$

which is strictly increasing on $\mathbb{R}^+$ with $Q(0) = 0$. Using estimates (5.14) and (5.15), Eq. (7.8) and therefore Eq. (7.7) are satisfied if

$$c_k \geq \sum_{i=2}^{n_k} \mu_i \left(\frac{[A:B]}{\tau}\right)^{i-1} = Q\left(\frac{[A:B]}{\tau}\right),$$

which is true if $\tau \geq [A:B]/\varepsilon_k$, where ε_k is the positive real such that $Q(\varepsilon_k) = c_k$.

Therefore we have proved Eq. (7.7) for τ large enough and Lax–Milgram's lemma yields the existence and the uniqueness of a solution $u(t)$ of (7.6), which belongs to $\mathcal{H}_A^1$.

We multiply now Eq. (7.6) by $u(t)$ and we obtain

$$u(t)^*Au(t) + u(t)^*Ap_1(t)u(t) + \sum_{i=2}^{n_k}(-t)^{i-1}u(t)^*Ap_i(t)u(t) = u(t)^*Af. \tag{7.9}$$

Using Eq. (7.7), we find that

$$u(t)^*Au(t) \leq u(t)^*Af$$

and therefore that

$$|u(t)|_A \leq |f|_A. \tag{7.10}$$

Thus, for any subsequence t_n, decreasing toward 0, we extract a subsequence, still denoted by t_n, such that

$$u(t_n) \rightharpoonup u_0 \text{ weakly in } \mathcal{H}_A^1. \tag{7.11}$$

We pass to the limit in equation

$$u(t_n) + p_1(t_n)u(t_n) + \sum_{i=2}^{n_k}(-t_n)^{i-1}p_i(t_n)u(t_n) = f. \tag{7.12}$$

Since

$$p_1(t_n)u(t_n) = \sum_{j=1}^{k} \ell_j^k(0)\beta_j Au(t_n)$$

and $Au(t_n) \rightharpoonup Au_0$ weakly in $\mathcal{H}_A^{-1}$, we deduce from Lemma 3.3 that

$$p_1(t_n)u(t_n) \rightharpoonup \sum_{j=1}^{k} \ell_j^k(0)Au_0 = Au_0 \text{ weakly in } \mathcal{H}_A^{-1}. \tag{7.13}$$

Moreover, since for all j, $1 \leq j \leq k$,

$$\|\beta_j\|_{\mathcal{L}(\mathcal{H}_B^{-1},\mathcal{H}_B^1)} \leq 1, \tag{7.14}$$

there exists $C > 0$ such that

$$|\beta_j Au(t_n)|_A \leq C|u(t_n)|_A \tag{7.15}$$

and therefore the term $p_i(t_n)u(t_n)$ is bounded by $C|f|_A$ in $\mathcal{H}_A^1$, where C is a positive constant. Thus, the following limit holds true:

$$(-t_n)^{i-1}p_i(t_n)u(t_n) \to 0 \text{ strongly in } \mathcal{H}_A^1. \tag{7.16}$$

Thus, from limits (7.11), (7.13) and (7.16), Eq. (7.12) yields

$$u_0 + Au_0 = f.$$

To conclude, as in the proof of Theorem 7.1, that $u(t)$ converges strongly to u_0 in $\mathcal{H}_A^1$, we have to prove that

$$|u(t_n)|^2 + u(t_n)^*Au(t_n) \to |u_0|^2 + u_0^*Au_0. \tag{7.17}$$

We deduce from Eq. (7.12) multiplied on the left by $u(t_n)^*$ that

$$\begin{aligned} |u(t_n)|^2 + u(t_n)^*Au(t_n) &= u(t_n)^*Au(t_n) - u(t_n)^*p_1(t_n)u(t_n) \\ &\quad - \sum_{i=2}^{n_k}(-t_n)^{i-1}u(t_n)^*p_i(t_n)u(t_n) + u(t_n)^*f. \end{aligned} \tag{7.18}$$

The last term on the right-hand side of (7.18) converges:

$$u(t_n)^*f \to u_0^*f = |u_0|^2 + u_0^*Au_0. \tag{7.19}$$

Now let us prove that the sum of the other terms of the right-hand side of (7.18) converges to zero. For that purpose, we first prove that $u(t_n)^* p_i(t_n)u(t_n)$ is bounded. We remark that, thanks to hypothesis (1.2),

$$|u(t_n)^* p_i(t_n)u(t_n)| \leq |u(t_n)||p_i(t_n)u(t_n)| \leq |u(t_n)|_A |p_i(t_n)u(t_n)|_A;$$

the $\mathcal{H}_A^1$ norm of $p_i(t_n)u(t_n)$ is bounded by $C|f|_A$ from Eq. (7.15), as explained above, and $|u(t_n)|_A$ is bounded by $|f|_A$ thanks to Eq. (7.10). Therefore,

$$(-t_n)^{i-1} u(t_n)^* p_i(t_n)u(t_n) \to 0. \tag{7.20}$$

Now let us compute $A - p_1(t)$ in order to factorize it by t. Hence, a simple computation leads to

$$\begin{aligned}
&\mathbf{1} - \sum_{j=1}^{k} \ell_j^k(0)\beta_j \\
&\quad = \prod_{i=1}^{k} \beta_i \left(\prod_{l=1}^{k} \left(\mathbf{1} + \frac{t\tau}{n_l} B \right) - \sum_{j=1}^{k} \ell_j^k(0) \prod_{\{l:1\leq l\leq k, l\neq j\}} \left(\mathbf{1} + \frac{t\tau}{n_l} B \right) \right) \\
&\quad = \prod_{i=1}^{k} \beta_i \left(\mathbf{1} - \sum_{j=1}^{k} \ell_j^k(0)\mathbf{1} + tS(B) \right),
\end{aligned}$$

where S is a polynomial of degree k, with coefficients depending continuously on t. Therefore, using Eq. (5.2), we find that

$$A - p_1(t) = t \prod_{i=1}^{k} \beta_i S(B) A.$$

Thus, we can estimate for l, $1 \leq l \leq k$ the term $u(t_n)^* \prod_i \beta_i B^l A u(t_n)$ as follows:

$$\left| u(t_n)^* \prod_{i=1}^{k} \beta_i B^l A u(t_n) \right| \leq \left| B^l \prod_{i=1}^{k} \beta_i u(t_n) \right|_A |u(t_n)|_A;$$

using the fact that $|\beta_j B u(t_n)|_A$ is bounded by $C|u(t_n)|_A$ by Eq. (7.14) and using Eq. (7.10), we conclude that

$$u(t_n)^* \prod_{i=1}^{k} \beta_i S(B) A u(t_n)$$

is bounded and therefore that the term

$$u(t_n)^* A u(t_n) - u(t_n)^* p_1(t_n) u(t_n) = t_n u(t_n)^* \prod_{i=1}^{k} \beta_i S(B) A u(t_n)$$

converges to 0. Using the two other limits (7.19) and (7.20), we can conclude that (7.17) holds true and that $u(t)$ converges strongly to u_0 in $\mathcal{H}_A^1$; the proof of Theorem 7.2 is complete. $\square$

8. Order of the Residual Smoothing Scheme and of Its Extrapolations

In the following, we will denote by $C(|u|_{\mathcal{H}_A^n})$ a constant depending only on $|u|_{\mathcal{H}_A^n}$ and by $C(|u|_{\mathcal{H}_A^n}, \tau)$ a constant depending on $|u|_{\mathcal{H}_A^n}$ and τ.

Let us prove that the residual smoothing scheme is of order one in time:

Theorem 8.1. *Suppose that $\mathcal{H}_B^2$ and $\mathcal{H}_A^2$ are isomorphic. There exists $t_0 > 0$ such that for all $u \in \mathcal{H}_A^5$, for all τ larger than $[A:B]/2$ and for all $T > 0$, there exists $C(|u|_{\mathcal{H}_A^5}, T, \tau)$ such that for all $t \in (0, t_0]$ and for all n such that $nt \leq T$,*

$$|P(t)^n u - e^{-ntA} u|_A \leq C(|u|_{\mathcal{H}_A^5}, T, \tau) t.$$

Proof. Let $u \in \mathcal{H}_A^5$,

$$|e^{-tA}u - u + tAu|_A \leq Ct^2 |A^2 u|_A = Ct^2 |A^{5/2} u|. \tag{8.1}$$

We also have the following equality:

$$\beta(t) = \mathbf{1} - t\tau\beta(t)B;$$

thus $P(t)u$ can be expressed as follows:

$$P(t)u = u - t\beta(t)Au = u - tAu + t^2\tau\beta(t)BAu. \tag{8.2}$$

Equations (8.1) and (8.2) lead to the following estimate:

$$|P(t)u - e^{-tA}u|_A \leq Ct^2(|A^2u|_A + \tau|BAu|_A) \tag{8.3}$$

and as $\mathcal{H}_B^2$ and $\mathcal{H}_A^2$ are isomorphic,

$$\leq Ct^2\tau|A^2u|_A.$$

Let us now consider n iteration steps. Using the triangle inequality, we obtain

$$|P(t)^n u - e^{-ntA}u|_A \leq \sum_{j=0}^{n-1} |P(t)^{n-j-1}(P(t) - e^{-tA})e^{-jtA}u|_A$$

using Theorem 4.1 and estimate (8.3),

$$\begin{aligned} &\leq \sum_{j=0}^{n-1} C\tau t^2 |A^2 e^{-jtA} u|_A \\ &\leq C\tau n t^2 |u|_{\mathcal{H}_A^5} \\ &\leq CTt\tau |u|_{\mathcal{H}_A^5} \end{aligned}$$

and the proof is complete. □

Now let us prove that the extrapolation P_k of P is of order k in time.

Theorem 8.2. *Suppose that $\mathcal{H}_B^{2k+2}$ and $\mathcal{H}_A^{2k+2}$ are isomorphic. There exist $p \in \mathbb{N}$, $\tau_0 > 0$ and $t_0 > 0$ such that for all $u \in \mathcal{H}_A^p$, for all $\tau \geq \tau_0$ and for all $T > 0$, there exists $C(|u|_{\mathcal{H}_A^p}, T, \tau)$ such that for all $t \in (0, t_0]$ and for all n such that $nt \leq T$,*

$$|P_k(t)^n u - e^{-ntA} u|_A \leq C(|u|_{\mathcal{H}_A^p}, T, \tau) t^k.$$

Proof. We will use Theorem 3.1 of [11]. A is an operator with bounded inverse and $Y = \cap_{k \in \mathbb{Z}} D(A^k)$ is dense in $\mathcal{H}^0$. As in [11], we will denote by $\mathcal{Z}_k$ the set of operators L such that $Y \subset D(L)$, $L : Y \to Y$ and

$$\text{for all } m \in \mathbb{Z}, \quad |L|_{m,k} = \sup_{u \in Y \setminus \{0\}} \frac{|A^{m-k} L u|_0}{|A^m u|_0} < \infty.$$

$\mathcal{Z} = \cup_{k \in \mathbb{Z}} \mathcal{Z}_k$ is a subalgebra of the algebra of linear operators from Y to itself. We can remark that for all $k \in \mathbb{N}$, $A^k \in \mathcal{Z}_k \subset \mathcal{Z}$.

A is also the generator of a strongly continuous semigroup $\exp(-tA)$ which satisfies the following estimate:

$$\text{for all } m \in \mathbb{Z}, \quad \left| \exp(-tA) - \sum_{j=0}^{k} \frac{(-tA)^j}{j!} \right|_{m,k+1} = O(t^{k+1}). \tag{8.4}$$

The formula

$$\beta(t) = \sum_{l=0}^{k} (-t\tau)^l B^l + (-t\tau)^{k+1} \beta(t) B^{k+1}$$

enables us to develop $P(t)u$ as follows:

$$P(t)u = u + \sum_{l=1}^{k+1} (-t)^l \tau^{l-1} B^{l-1} A u + (-t)^{k+2} \tau^{k+1} \beta(t) B^{k+1} A u. \tag{8.5}$$

Let us define

$$f_l = (-1)^l \tau^{l-1} B^{l-1} A \quad \text{and} \quad \varepsilon_{k+1}(t) = (-1)^{k+2} \tau^{k+1} t \beta(t) B^{k+1} A.$$

Then,

$$P(t) = \mathbf{1} + \sum_{l=1}^{k+1} t^l f_l + t^{k+1} \varepsilon_{k+1}(t).$$

We can remark that $f_1 = -A$ as required.

As $\mathcal{H}_B^{2k+2}$ and $\mathcal{H}_A^{2k+2}$ are isomorphic, $f_l \in \mathcal{Z}_l$ and $\varepsilon_{k+1}(t) \in \mathcal{Z}_{k+2}$. We also have

$$\text{for all } m \in \mathbb{Z}, \quad \lim_{t \to 0} |\varepsilon_{k+1}(t)|_{m,k+2} = 0.$$

Finally, thanks to estimate (8.4) and Eq. (8.5), we obtain

$$\text{for all } m \in \mathbb{Z}, \quad t^{-2} |P(t) - \exp(-tA)|_{m,2} = C(|u|_{\mathcal{H}_A^4}).$$

Finally, we can adapt the proof of Theorem 3.1 of [11] and we find that there exists i large enough such that, for all $m \in \mathbb{Z}$,

$$t^{-(k+1)} \left| \sum_{j=1}^{k} \ell_j^k(0) P(t/n_j)^{n_j} - \exp(-tA) \right|_{m,i} = O(1).$$

And in particular, for $m = i + 1$, we obtain, using Eq. (1.2), that

$$\left| \left(\sum_{j=1}^{k} \ell_j^k(0) P(t/n_j)^{n_j} - \exp(-tA) \right) u \right|_A \leq C t^{k+1} |u|_{\mathcal{H}_A^{2i+2}}.$$

For n time steps, the end of the proof is similar to Theorem 8.1. □

9. Conclusion

As a conclusion, we have proved stability, convergence and order properties for the extrapolations of the Residual Smoothing Scheme. This scheme has already been applied to domain decomposition method and fictitious domain method in [18] and to image processing in [1]. We also applied it and analyzed it for some spectral discretizations of Laplacian preconditioned by finite elements methods [28]. Finite volume methods or Discontinuous Galerkin methods may also be considered in the future as accurate preconditioners of spectral discretizations for more complex problems.

From a theoretical and practical point of view, another interesting perspective would be the proof of stability, convergence and order of preconditioned Runge–Kutta schemes as those presented in [30]. This would be an important generalization of this article, especially with a focus on the preconditioning of nonlinear problems by linear operators, getting inspired by Newton's method for example.

Acknowledgments

In memoriam of Michelle Schatzman. With my gratitude for all she taught me. M.R.

References

1. A. Averbuch, A. Cohen and M. Israeli, A stable and accurate explicit scheme for parabolic evolution equations, 1998, http://www.ann.jussieu.fr/∼cohen/para.ps.gz.
2. A. V. Balakrishnan, Fractional powers of closed operators and the semigroups generated by them, *Pacific J. Math.* **10** (1960) 419–437.
3. F. A. Bornemann, An adaptive multilevel approach to parabolic equations, II. Variable-order time discretization based on a multiplicative error correction, *Impact Comput. Sci. Engrg.* **3** (1991) 93–122.
4. F. A. Bornemann, An adaptive multilevel approach to parabolic equations, III. 2D error estimation and multilevel preconditioning, *Impact Comput. Sci. Engrg.* **4** (1992) 1–45.

5. S. H. Brill and G. F. Pinder, Analysis of a block red-black preconditioner applied to the Hermite collocation discretization of a model parabolic equation, *Numer. Methods Partial Differential Equations* **17** (2001) 584–606.
6. P. N. Brown and C. S. Woodward, Preconditioning strategies for fully implicit radiation diffusion with material-energy transfer, *SIAM J. Sci. Comput.* **23** (2001) 499–516 (electronic), Copper Mountain Conference (2000).
7. C. Canuto and A. Quarteroni, Preconditioned minimal residual methods for Chebyshev spectral calculations, *J. Comput. Phys.* **60** (1985) 315–337.
8. B. Costa, L. Dettori, D. Gottlieb and R. Temam, Time marching multilevel techniques for evolutionary dissipative problems, *SIAM J. Sci. Comput.* **23** (2001) 46–65 (electronic).
9. M. Deville and E. Mund, Chebyshev pseudospectral solution of second-order elliptic equations with finite element preconditioning, *J. Comput. Phys.* **60** (1985) 517–533.
10. M. O. Deville and E. H. Mund, Finite-element preconditioning for pseudospectral solutions of elliptic problems, *SIAM J. Sci. Statist. Comput.* **11** (1990) 311–342.
11. B. Dia and M. Schatzman, On the order of extrapolation of integration formulae, Technical Report 275, Équipe d'Analyse Numérique de Lyon, 1998, http://maply.univ-lyon1.fr/publis/publiv/1998/dia190298/texte.ps.
12. P. Dutt, P. Biswas and G. Naga Raju, Preconditioners for spectral element methods for elliptic and parabolic problems, *J. Comput. Appl. Math.* **215** (2008) 152–166.
13. I. Gerace, P. Pucci, N. Ceccarelli, M. Discepoli and R. Mariani, A preconditioned finite element method for the p-Laplacian parabolic equation, *Appl. Numer. Anal. Comput. Math.* **1** (2004) 155–164.
14. P. Haldenwang, G. Labrosse, S. Abboudi and M. Deville, Chebyshev 3D spectral and 2D pseudospectral solvers for the Helmholtz equation, *J. Comput. Phys.* **55** (1984) 115–128.
15. X.-Q. Jin, A note on circulant preconditioners for hyperbolic and parabolic equations, *Chinese J. Math.* **21** (1993) 129–142.
16. X.-Q. Jin and R. H. Chan, Circulant preconditioners for second order hyperbolic equations, *BIT* **32** (1992) 650–664.
17. T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, 1995), reprint of the 1980 edition.
18. Y. M. Laevsky, Preconditioning operators for grid parabolic problems, *Russ. J. Numer. Anal. Math. Model.* **11** (1996) 497–515.
19. J.-L. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications*, Vol. I (Springer-Verlag, 1972), translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.
20. K.-A. Mardal, T. K. Nilssen and G. A. Staff, Order-optimal preconditioners for implicit Runge-Kutta schemes applied to parabolic PDEs, *SIAM J. Sci. Comput.* **29** (2007) 361–375 (electronic).
21. P. K. Moore and R. H. Dillon, A comparison of preconditioners in the solution of parabolic systems in three space dimensions using DASPK and a high order finite element method, *Appl. Numer. Math.* **20** (1996) 117–128, *Workshop on the Method of Lines for Time-Dependent Problems*, Lexington, KY, 1995.
22. L. S. Mulholland and D. M. Sloan, The role of preconditioning in the solution of evolutionary partial differential equations by implicit Fourier pseudospectral methods, *J. Comput. Appl. Math.* **42** (1992) 157–174.
23. S. A. Orszag, Spectral methods for problems in complex geometries, *J. Comput. Phys.* **37** (1980) 70–92.

24. S. V. Parter, Preconditioning Legendre special collocation methods for elliptic problems, I. Finite difference operators, *SIAM J. Numer. Anal.* **39** (2001) 330–347 (electronic).
25. S. V. Parter, Preconditioning Legendre spectral collocation methods for elliptic problems, II. Finite element operators, *SIAM J. Numer. Anal.* **39** (2001) 348–362 (electronic).
26. M. Reed and B. Simon, *Methods of Modern Mathematical Physics. II. Fourier Analysis, Self-Adjointness* (Academic Press, 1975).
27. M. Ribot, Asymptotics of some ultra-spherical polynomials and their extrema, *Methods Appl. Anal.* **12** (2005) 43–74.
28. M. Ribot, Etude théorique de méthodes numériques pour les systèmes de réaction-diffusion; application à des équations paraboliques non linéaires et non locales, Ph.D. Thesis, December 2003, http://math.unice.fr/∼ribot/recherche/these.pdf.
29. M. Ribot and M. Schatzman, Equivalence between the spectral and the finite elements matrices, in *Computational Fluid and Solid Mechanics, Proc. of the Second MIT Conference on Computational Fluid and Solid Mechanics*, ed. K. J. Bathe, Cambridge, Massachusetts, USA, June 17–20, 2003 (Elsevier, 2003), http://math.unice.fr/∼ribot/recherche/RibotSchatzmanMIT2.pdf.
30. M. Schatzman, Toward noncommutative numerical analysis: High order integration in time, in *Proc. of the Fifth Int. Conf. on Spectral and High Order Methods* (ICOSAHOM-01) (Uppsala), Vol. 17, 2002, pp. 99–116.
31. G. A. Staff, K.-A. Mardal and T. K. Nilssen, Preconditioning of fully implicit Runge–Kutta schemes for parabolic PDEs, *Model. Identif. Control* **27** (2006) 109–123.
32. K. Yosida, *Functional Analysis* (Springer-Verlag, 1995), reprint of the sixth 1980 edition.