Helsinki University of Technology, Institute of Mathematics, Research Reports
Teknillisen korkeakoulun matematiikan laitoksen tutkimusraporttisarja
Espoo 2004 A480

APPROXIMATION OF THE LAPLACE TRANSFORM BY THE CAYLEY TRANSFORM

Ville Havu Jarmo Malinen

TEKNILLINEN KORKEAKOULU
TEKNISKA HÖGSKOLAN
HELSINKI UNIVERSITY OF TECHNOLOGY
TECHNISCHE UNIVERSITÄT HELSINKI
UNIVERSITE DE TECHNOLOGIE D'HELSINKI

Helsinki University of Technology, Institute of Mathematics, Research Reports
Teknillisen korkeakoulun matematiikan laitoksen tutkimusraporttisarja
Espoo 2004 A480

APPROXIMATION OF THE LAPLACE TRANSFORM BY THE CAYLEY TRANSFORM

Ville Havu **Jarmo Malinen**

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Institute of Mathematics

V. Havu[1] and J. Malinen: *Approximation of the Laplace transform by the Cayley transform*; Helsinki University of Technology, Institute of Mathematics, Research Reports A480 (2004).

Abstract: *We interpret the usual Cayley transform of linear (infinite-dimensional) state space systems as a numerical integration scheme of Crank–Nicholson type. This turns out to be equivalent to an approximation procedure of the Laplace transform. The convergence properties of such an approximation are investigated.*

AMS subject classifications: 47A48, 65J10, 93C25, (34G10, 47N70, 65L70)

Keywords: Conservative system, Crank–Nicholson scheme

Correspondence

Ville.Havu@tkk.fi, Jarmo.Malinen@tkk.fi

ISBN 951-22-7453-1
ISSN 0748-3143
Espoo, 2004

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Institute of Mathematics
P.O. Box 1100, 02015 HUT, Finland
email:math@hut.fi http://www.math.hut.fi/

[1]This work is supported by the European Commission's 5th Framework Programme: Smart Systems; New materials, adaptive systems and their nonlinearities, HPRN-CT-2002-00284.

1 Introduction and motivation

Let U and X be separable Hilbert spaces. Let $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ be a system node in the sense of [8], whose input and output space are U, and the state space is X. An additional space $V := \left\{ \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in \left[\begin{smallmatrix} X \\ U \end{smallmatrix}\right] : A_{-1}x + Bu \in X \right\}$ is defined as usual, and it is equipped with the natural norm making it a Hilbert space. Then, as is well-known, the Cauchy problem

$$\begin{cases} x'(t) = A_{-1}x(t) + Bu(t), \quad t \geq 0, \\ x(0) = x_0 \end{cases} \tag{1.1}$$

is uniquely solvable for any input $u \in C^2(\mathbb{R}_+; U)$ and initial state $x_0 \in X$ for which the compatibility condition $\left[\begin{smallmatrix} x_0 \\ u(0) \end{smallmatrix}\right] \in V$ holds. Moreover, then also $\left[\begin{smallmatrix} x(\cdot) \\ u(\cdot) \end{smallmatrix}\right] \in C(\mathbb{R}_+; V)$, and hence the output relation $y(t) = C\&D \left[\begin{smallmatrix} x(t) \\ u(t) \end{smallmatrix}\right]$ is well defined for all $t \geq 0$ as $C\&D \in \mathcal{L}(V; U)$. These and many other facts can be found in [8, Section 2].

The system node $\left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ is *energy preserving* if the following energy balance holds for all $T > 0$

$$\langle x(T), x(T)\rangle_X^2 + \int_0^T \langle y(t), y(t)\rangle_Y^2 dt = \langle x_0, x_0\rangle_X^2 + \int_0^T \langle u(t), u(t)\rangle_U^2 dt, \tag{1.2}$$

where u, x, y and x_0 are as in (1.1). For any energy preserving S, the semigroup generator A is maximally dissipative and $\mathbb{C}_+ \subset \rho(A)$. If both $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ and its *dual node* $S^d = \left[\begin{smallmatrix} [A\&B]^d \\ [C\&D]^d \end{smallmatrix}\right]$ are energy-preserving, then $\left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ is called *conservative*; see [8, Definitions 3.1 and 4.1]. Conservative system nodes are known in classical operator theory as *operator colligations* or *Livšic – Brodskiĭ nodes*. A wide classical literature exists for them but the practical linear systems content might sometimes be hard to understand. See e.g. Brodskiĭ [4, 6, 5], Livšic [12], Livšic and Yantsevich [11], Sz.-Nagy and Foiaş [15], Smuljan [13], and Helton [3]. An up-to-date, comprehensive reference for operator nodes is Staffans [14]. The general conservative case is treated in Malinen, Staffans and Weiss [8], and the special case of *boundary control* systems are described in [7, 9].

For simplicity, it will be henceforth assumed that all system nodes treated in this paper are conservative, even though most of the results could be given in a more general setting. For the same reason, we assume that $U = \mathbb{C}$, i.e. the signals $u(\cdot)$ and $y(\cdot)$ in (1.1) are scalar valued, even though everything would still remain true (with similar proofs) even if U was a separable Hilbert space.

Let us assume, for a moment, that we are treating the matrix case. Then the dynamical equations take the usual form

$$\begin{cases} x'(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t) + Du(t), \quad t \geq 0, \\ x(0) = x_0. \end{cases} \tag{1.3}$$

where $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times 1}$, $C \in \mathbb{C}^{1\times n}$, and $D \in \mathbb{C}$. Let $h > 0$ be a *discretization parameter*. We can carry out a slightly nonstandard time discretization of (1.3) and obtain an approximation of Crank–Nicholson type

$$\begin{cases} \frac{x(jh)-x((j-1)h)}{h} & \approx A\frac{x(jh)+x((j-1)h)}{2} + Bu(jh), \\ y(jh) & \approx C\frac{x(jh)+x((j-1)h)}{2} + Du((j-1)h), \quad j \geq 1, \\ x(0) & = x_0. \end{cases}$$

Clearly, this induces the discrete time dynamics

$$\begin{cases} \frac{x_j^{(h)}-x_{j-1}^{(h)}}{h} & = A\frac{x_j^{(h)}+x_{j-1}^{(h)}}{2} + B\frac{u_j^{(h)}}{\sqrt{h}}, \\ \frac{y_j^{(h)}}{\sqrt{h}} & = C\frac{x_j^{(h)}+x_{j-1}^{(h)}}{2} + D\frac{u_j^{(h)}}{\sqrt{h}}, \quad j \geq 1, \\ x_0^{(h)} & = x_0, \end{cases} \tag{1.4}$$

where loosely speaking $u_j^{(h)}/\sqrt{h}$ is an approximation of $u(jh)$. We hope very much that $y_j^{(h)}/\sqrt{h}$ would be close to $y(jh)$ — at least under some exceptionally happy circumstances. After some easy computations, equations (1.4) take the form

$$\begin{cases} x_j^{(h)} & = A_\sigma x_{j-1}^{(h)} + B_\sigma u_j^{(h)}, \\ y_j^{(h)} & = C_\sigma x_{j-1}^{(h)} + D_\sigma u_j^{(h)}, \quad j \geq 1, \\ x_0^{(h)} & = x_0, \end{cases} \tag{1.5}$$

where $A_\sigma := (\sigma+A)(\sigma-A)^{-1}$, $B_\sigma := \sqrt{2\sigma}(\sigma-A)^{-1}B$, $C_\sigma := \sqrt{2\sigma}C(\sigma-A)^{-1}$ and $D_\sigma := D + C(\sigma - A)^{-1}B$ with $\sigma := 2/h$.

Even though the computation leading to (1.5) was carried out in the matrix setting, exactly the same transformation can be done for any system node $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$. We simply define the *discrete time linear system* (henceforth, DLS) described by the operator quadruple

$$\phi_\sigma = \begin{bmatrix} A_\sigma & B_\sigma \\ C_\sigma & D_\sigma \end{bmatrix} = \begin{bmatrix} (\sigma + A)(\sigma - A)^{-1} & \sqrt{2\sigma}(\sigma - A_{-1})^{-1}B \\ \sqrt{2\sigma}C(\sigma - A)^{-1} & \mathbf{G}(\sigma) \end{bmatrix} \tag{1.6}$$

for any $\sigma > 0$ (or even for any $\sigma \in \mathbb{D}$, $\mathbb{D}$ being the unit disk, but we shall not use this in this paper). Here $\mathbf{G}(\cdot)$ denotes the transfer function of S, and it is defined by $\mathbf{G}(s) = C\&D \left[\begin{smallmatrix} (s-A_{-1})^{-1}B & I \end{smallmatrix}\right]^T$ for all $s \in \mathbb{C}_+$.

In system theory, the transformation $S \mapsto \phi_\sigma$ is called *Cayley transform* of continuous time systems to discrete time systems. By some computations, it can be checked that the discrete time transfer function $\mathbf{D}_\sigma(\cdot)$ of ϕ_σ satisfies

$$\mathbf{D}_\sigma(z) := D_\sigma + zC_\sigma(I - zA_\sigma)^{-1}B_\sigma = \mathbf{G}\left(\frac{1-z}{1+z}\sigma\right). \tag{1.7}$$

We say that the DLS ϕ_σ of type (1.5) is *conservative* if the defining block matrix $\left[\begin{smallmatrix} A_\sigma & B_\sigma \\ C_\sigma & D_\sigma \end{smallmatrix}\right]$ is unitary. Then the discrete time balance equation

$$\sum_{j=1}^{N} \|x_j\|^2 - \sum_{j=1}^{N} \|x_{j-1}\|^2 = \sum_{j=1}^{N} \|u_{j-1}\|^2 - \sum_{j=1}^{N} \|y_{j-1}\|^2$$

is satisfied for all $N \geq 1$, where the sequences $\{u_j\}$, $\{x_j\}$ and $\{y_j\}$ satisfy (1.5). Studying the approximation scheme (1.4) might not be well motivated, unless the following proposition did not hold:

Proposition 1. *Let the system node* $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix} \right]$ *and the DLS* $\phi_\sigma = \left[\begin{smallmatrix} A_\sigma & B_\sigma \\ C_\sigma & D_\sigma \end{smallmatrix} \right]$ *be connected by* (1.6). *Then S is (continuous time) conservative (passive) if and only if ϕ_σ is (discrete time) conservative (resp., passive).*

There exists an extensive literature on the Cayley transform of systems, and we shall not try to make a full account of it here. See e.g. Ober and Montgomery-Smith [10]. A nice piece of work, parallelling our approach, is Arov and Gavrilyuk [1].

2 Approximation of the input/output mapping

In this section, we describe the discretization (1.5) of dynamical system (1.1) in operator theory language.

2.1 Spaces and transforms.

The norm of the usual Hardy space $H^2(\mathbb{C}_+)$ is given by

$$\|\Phi\|^2_{H^2(\mathbb{C}_+)} = \sup_{x>0} \frac{1}{2\pi} \int_{-\infty}^{\infty} |\Phi(x+yi)|^2 \, dy.$$

As usual, the Laplace transform is defined

$$(\mathcal{L}f)(s) = \int_0^\infty e^{-st} f(t) \, dt \quad \text{for all} \quad s \in \mathbb{C}_+, \tag{2.1}$$

and it maps $L^2(\mathbb{R}_+) \to H^2(\mathbb{C}_+)$ unitarily. The norm of $H^2(\mathbb{D})$ is given by $\|\phi\|^2_{H^2(\mathbb{D})} = \sum_{j\geq 0} |\phi_j|^2$ if $\phi(z) = \sum_{j \geq 0} \phi_j z^j$, which makes the Z-transform unitary from $\ell^2(\mathbb{Z}_+) \to H^2(\mathbb{D})$. If, say, $f \in C_c(\mathbb{R})$ in (2.1), then $(\mathcal{L}f)(s)$ is well defined for all $s \in i\mathbb{R}$, too. We then call the function $i\omega \mapsto (\mathcal{L}f)(i\omega)$ the Fourier transform of f.

From now on, denote by $\mathbf{D}_\sigma : H^2(\mathbb{D}) \to H^2(\mathbb{D})$ the multiplication operator defined by $(\mathbf{D}_\sigma \tilde{u})(z) = \mathbf{D}_\sigma(z)\tilde{u}(z)$ for all $z \in \mathbb{D}$ and $\sigma > 0$. Similarly, denote by $\mathbf{G} : H^2(\mathbb{C}_+) \to H^2(\mathbb{C}_+)$ the multiplication operator satisfying $(\mathbf{G}\hat{u})(s) = \mathbf{G}(s)\hat{u}(s)$ for all $s \in \mathbb{C}_+$[2]. It follows immediately that (1.7) takes the form of the similarity transformation

$$\mathbf{G} = \mathcal{C}_\sigma^{-1} \mathbf{D}_\sigma \mathcal{C}_\sigma, \tag{2.2}$$

where the *composition operator* is defined by $(\mathcal{C}_\sigma F)(z) := F(\frac{1-z}{1-z}\sigma)$ for all $z \in \mathbb{D}$ and $F : \mathbb{C}_+ \to \mathbb{C}$. Trivially $(\mathcal{C}_\sigma^{-1} f)(s) := f(\frac{s-\sigma}{s+\sigma})$ for all $s \in \mathbb{C}_+$ and $f : \mathbb{D} \to \mathbb{C}$.

[2]Then $\mathbf{D}_\sigma$ and $\mathbf{G}$ are unitarily equivalent to the input/output mappings of ϕ_σ and S, respectively.

Proposition 2. *The mapping $f \mapsto F$ given by $F(s) = \frac{\sqrt{2/\sigma}}{1+s/\sigma} f(\frac{s-\sigma}{s+\sigma})$ is unitary from $H^2(\mathbb{D})$ onto $H^2(\mathbb{C}_+)$. In particular, the operator $\mathcal{M}_\sigma \mathcal{C}_\sigma^{-1} : H^2(\mathbb{D}) \to H^2(\mathbb{C}_+)$ is unitary, where $\mathcal{M}_\sigma : H(\mathbb{C}_+) \to H(\mathbb{C}_+)$ denotes the multiplication operator by $\frac{\sqrt{2/\sigma}}{1+s/\sigma}$.*

Proof. This follows as soon as it is shown that for each $\sigma > 0$, the sequence $\left\{ \frac{\sqrt{2/\sigma}}{1+s/\sigma} \left(\frac{s-\sigma}{s+\sigma}\right)^j \right\}_{j\geq 0}$ is an orthonormal basis for $H^2(\mathbb{C}_+)$. □

2.2 Discretizing operators.

By T_σ we denote a discretizing (or sampling) bounded linear operator $T_\sigma : L^2(\mathbb{R}_+) \to H^2(\mathbb{D})$. The adjoint T_σ^* of T_σ maps then $H^2(\mathbb{D}) \to L^2(\mathbb{R}_+)$, and it is typically an interpolating operator. In this paper, we define T_σ is by

$$(T_\sigma u)(z) = \sum_{j\geq 1} u_j^{(h)} z^j \quad \text{where} \quad \frac{u_j^{(h)}}{\sqrt{h}} = \frac{1}{h}\int_{(j-1)h}^{jh} u(t)\, dt, \tag{2.3}$$

with $h = 2/\sigma$; see (1.4) and (1.5). Then the adjoint T_σ^* is given by

$$\left(T_\sigma^* \tilde{v}\right)(t) = \frac{1}{\sqrt{h}} \sum_{j\geq 1} v_j \chi_{[(j-1)h, jh]}(t) \tag{2.4}$$

where $\tilde{v}(z) = \sum_{j\geq 0} v_j z^j \in H^2(\mathbb{D})$ and $\chi_I(\cdot)$ denotes the characteristic function of the interval I.

It is worth noticing that the operator $T_\sigma : L^2(\mathbb{R}_+) \to H^2(\mathbb{D})$ is a coisometry. This can be seen as follows:

$$\begin{aligned} \|T_\sigma^* \tilde{v}\|^2_{L^2(\mathbb{R}_+)} &= \frac{1}{h}\int_0^\infty |\sum_{j\geq 1} v_j \chi_{[(j-1)h,jh]}|^2\, dt = \frac{1}{h}\int_0^\infty \sum_{j\geq 1} |v_j|^2 \chi_{[(j-1)h,jh]}\, dt \\ &= \frac{1}{h}\sum_{j\geq 1} |v_j|^2 \int_0^\infty \chi_{[(j-1)h,jh]}\, dt = \sum_{j\geq 1} |v_j|^2 = \|\tilde{v}\|^2_{H^2(\mathbb{D})}. \end{aligned} \tag{2.5}$$

2.3 Approximation of the Laplace transform.

Let us now use the discrete time trajectories of (1.5) to approximate the continuous time dynamics in (1.3).

Let $u \in L^2(\mathbb{R}_+)$ be arbitrary. In the operator notation, the output of the discretized dynamics (1.5) (after interpolation by T_σ^* back to a continuous time signal) is given by $T_\sigma^* \mathbf{D}_\sigma T_\sigma u$. The output of continuous time dynamics (1.3) is given by $\mathcal{L}^* \mathbf{G} \mathcal{L} u$. Our first task is to show that at least for some nice $u \in L^2(\mathbb{R}_+)$ and $T > 0$ we have convergence

$$\|T_\sigma^* \mathbf{D}_\sigma T_\sigma u - \mathcal{L}^* \mathbf{G} \mathcal{L} u\|_{L^2([0,T])} \to 0 \tag{2.6}$$

at some speed as $\sigma \to \infty$. By Proposition 2 and equation (2.2) we see that

$$\begin{aligned} T_\sigma^* \mathbf{D}_\sigma T_\sigma &= T_\sigma^* \left(\mathcal{C}_\sigma \mathcal{M}_\sigma^{-1}\right) \cdot \mathbf{G} \cdot \left(\mathcal{M}_\sigma \mathcal{C}_\sigma^{-1}\right) T_\sigma \\ &= T_\sigma^* \left(\mathcal{M}_\sigma \mathcal{C}_\sigma^{-1}\right)^{-1} \cdot \mathbf{G} \cdot \left(\mathcal{M}_\sigma \mathcal{C}_\sigma^{-1}\right) T_\sigma = \left(\mathcal{M}_\sigma \mathcal{C}_\sigma^{-1} T_\sigma\right)^* \cdot \mathbf{G} \cdot \left(\mathcal{M}_\sigma \mathcal{C}_\sigma^{-1} T_\sigma\right) \end{aligned}$$

since the multiplication operator $\mathcal{M}_\sigma$ commutes with $\mathbf{G}$. Hence by (2.6), we are led to inquire whether the operators $L_\sigma := \mathcal{M}_\sigma \mathcal{C}_\sigma^{-1} T_\sigma$ are close (on compact intervals) to the Laplace transform $\mathcal{L}$ when σ is large. This, indeed, appears to be true to some extent [3].

Proposition 3. *For any $u \in C_c(\mathbb{R}_+)$ and $s \in \mathbb{C}_+$, we have $(\mathcal{L}u)(s) = \lim_{\sigma\to\infty} (L_\sigma u)(s)$ where L_σ is defined as above.*

Proof. Defining T_σ by (2.3) we get

$$\begin{aligned} (L_\sigma u)(s) &= \frac{\sqrt{2/\sigma}}{1+s/\sigma} \sum_{j\geq 1} \left(\frac{1}{h}\int_{(j-1)h}^{jh} u(t)\,dt\right)\left(\frac{\sigma-s}{\sigma+s}\right)^j \qquad (2.7)\\ &= \frac{1}{1+s/\sigma} \sum_{j\geq 1} \left(\int_0^\infty \chi_{[(j-1)h,jh]}(t) \left(\frac{\sigma-s}{\sigma+s}\right)^j u(t)\,dt\right) \\ &= \int_0^\infty K_{s,\sigma}(t) u(t)\,dt, \end{aligned}$$

where $\sigma = 2/h$ and

$$K_{s,\sigma}(t) = \frac{1}{1+s/\sigma} \sum_{j\geq 1} \chi_{[(j-1)h,jh]}(t) \left(1 - \frac{2s}{s+\sigma}\right)^j. \qquad (2.8)$$

Now, if j is such that $t \in [(j-1)h, jh]$, then we obtain from the previous

$$K_{s,\sigma}(t) \approx \frac{1}{1+s/\sigma}\left(1 - \frac{s}{s/2+\sigma/2}\right)^{(\sigma/2)\cdot t} \to e^{-st} \text{ as } \sigma \to \infty.$$

We conclude that $\lim_{\sigma\to\infty} K_{s,\sigma}(t) = e^{-st}$ for all $s \in \mathbb{C}_+$ and $t \geq 0$. Moreover, for each fixed $s \in \mathbb{C}_+$ and $\sigma \geq 2|s|$ we have

$$\begin{aligned} |K_{s,\sigma}(t)| &\leq 2 \cdot \left(1 + \frac{2|s|}{\sigma - |s|}\right)^{(\sigma/2)\cdot t} \\ &\leq 2 \cdot \left(1 + \frac{2|s|}{\sigma-|s|}\right)^{(\sigma-|s|)t/2} \cdot \left(1 + \frac{2|s|}{\sigma-|s|}\right)^{|s|t/2} \leq 2\left(e\sqrt{3}\right)^{|s|t}. \end{aligned}$$

The proposition now follows from the Lebesgue dominated convergence theorem, as the integrand in (2.7) is has a compact support. ☐

The purpose of this paper is to give stronger versions of Proposition 3.

[3] Note that by Proposition 2 and equality (2.5), we see that each $L_\sigma : L^2(\mathbb{R}_+) \to H^2(\mathbb{C}_+)$ is a coisometry. The Laplace transform, in its turn, is an unitary mapping between the same spaces. Hence, the convergence of $L_\sigma \to \mathcal{L}$ must be rather weak.

3 A pointwise convergence estimate

Our main result will be given in this section. Theorem 1 provides a uniform speed estimate for the convergence of $(L_\sigma u)(i\omega) \to (\mathcal{L}u)(i\omega)$ for $i\omega \in K$ where $K \subset i\mathbb{R}$ is compact.

Before that some new definitions and notations must be given: Let $I_j = ((j-1)h, jh] = (t_{j-1}, t_j]$ and $t_{j-1/2} = \frac{1}{2}(t_{j-1} + t_j)$. For $u \in L^2(\mathbb{R}_+)$, let $I_{h,s}u$ be the piecewise constant interpolating function, defined by

$$(I_{h,s}u)(t) = \bar{u}_{j,h} + \frac{c_j(h,s)}{h}(t - t_{j-1/2}), \quad t \in I_j, \tag{3.1}$$

where $\bar{u}_{j,h} = \frac{1}{h}\int_{I_j} u(t)\,dt$ and the defining sequence $\{c_j(h,s)\}_{j\geq 1}$ (depending on two parameters h and s) will be later chosen in a particular way. Let P_h denote the orthogonal projection in $L^2(\mathbb{R}_+)$ onto the subspace of functions that are constant on each interval I_j. Then clearly for all $u \in L^2(\mathbb{R}_+)$, $j \geq 1$ and $t \in I_j$ we have $(P_h u)(t) = \bar{u}_{j,h}$.

Theorem 1. *Let $h > 0$, $\sigma = 2/h$, $T = Jh$ for some $J \in \mathbb{N}$, $u \in C_c(\mathbb{R}_+) \cap H^1(\mathbb{R}_+)$, and assume that* $\operatorname{supp}(u) := \{t \in \mathbb{R} : u(t) \neq 0\} \subset [0, T]$.

(i) *Then the sequence $\{c_j(h,s)\}_{j\geq 1}$ can be chosen so that $(L_\sigma - \mathcal{L})(I_{h,s}u)(s) = 0$ for all $s \in \overline{\mathbb{C}_+}$.*

(ii) *For any such choice of the sequence $\{c_j(h,s)\}_{j\geq 1}$, we have*

$$\begin{aligned} &|(L_\sigma u)(s) - (\mathcal{L}u)(s)| \\ &\leq \frac{hT^{1/2}|s|}{\pi}\left(\|I_{h,s}u - P_h u\|_{L^2([0,T])} + \frac{h}{\pi}|u|_{H^1([0,T])}\right) \end{aligned} \tag{3.2}$$

for all $s \in \overline{\mathbb{C}_+}$.

(iii) *The sequence $\{c_j(h,s)\}_{j\geq 1}$ in claim* (i) *can be chosen* optimally *so that*

$$\|I_{h,s}u - P_h u\|_{L^2([0,T])} \leq \frac{15}{218}\left(h^{-1/2}T^{-1/2} + \frac{|s|}{6e}\right)\|P_h u\|_{L^2([0,T])}$$

for a given $s \in i\mathbb{R}$, $T \geq 1$ if $9h \leq T^{2/3}e^{-\frac{4}{3}|s|T}$. Furthermore, then

$$\begin{aligned} &|(L_\sigma u)(s) - (\mathcal{L}u)(s)| \\ &\leq \frac{3h^{1/2}|s|}{100}\|u\|_{L^2([0,T])} + \frac{2hT^{1/2}|s|^2}{1000}\|u\|_{L^2([0,T])} \\ &+ \frac{h^2T^{1/2}|s|}{10}|u|_{H^1([0,T])}. \end{aligned} \tag{3.3}$$

Proof. Let us first make some general observations. By a simple argument, $\|P_h u\|^2_{L^2(\mathbb{R}_+)} = h\sum_{j\geq 1}\bar{u}^2_{j,h}$. Clearly for all $t \in I_j$

$$(I_{h,s}u - P_h u)(t) = \frac{c_j(h,s)}{h}(t - t_{j-1/2}).$$

Since for any $b > a$ we have

$$\frac{1}{(b-a)^2}\int_a^b \left(t - \frac{b+a}{2}\right)^2 = \frac{b-a}{12},$$

it follows that

$$\|I_{h,s}u - P_h u\|^2_{L^2([0,T])} = \sum_{j=1}^{J} \frac{c_j(h,s)^2}{h^2}\int_{t_{j-1}}^{t_j} (t - t_{j-1/2})^2\, dt \tag{3.4}$$
$$= \frac{h}{12}\sum_{j=1}^{J} c_j(h,s)^2.$$

In claim (i) we want to determine the sequence $\{c_j(h,s)\}_{j\geq 1}$ so as to satisfy $(L_\sigma - \mathcal{L})(I_{h,s}u)(s) = 0$ for given h and s. After some computations, we see that this is equivalent to requiring that $\{c_j(h,s)\}_{j\geq 1}$ satisfies

$$\sum_{j=1}^{J} \bar{u}_{j,h} I_j^{(0)}(h,s) + \sum_{j=1}^{J} c_j(h,s) J_j(h,s) = 0, \tag{3.5}$$

where for $s \in \overline{\mathbb{C}_+} \setminus \{0\}$

$$I_j^{(0)}(h,s) := \int_{I_j} \left[\frac{1}{1+s/\sigma}\left(\frac{\sigma - s}{\sigma + s}\right)^j - e^{-st}\right] dt \tag{3.6}$$
$$= \frac{2}{\sigma+s}\left(\frac{\sigma-s}{\sigma+s}\right)^j + \frac{1}{s}\left[e^{-sjh} - e^{-s(j-1)h}\right],$$

and

$$J_j(h,s) := I_j^{(1)}(h,s) - (j-1/2)h \cdot I_j^{(0)}(h,s) \tag{3.7}$$
$$= \frac{1}{s^2}\left[e^{-sjh} - e^{-s(j-1)h}\right] + \frac{h}{2s}\left[e^{-sjh} + e^{-s(j-1)h}\right],$$

together with

$$I_j^{(1)}(h,s) := \int_{I_j} \left[\frac{1}{1+s/\sigma}\left(\frac{\sigma - s}{\sigma + s}\right)^j - e^{-st}\right] t\, dt$$
$$= \frac{(2j-1)h}{\sigma+s}\left(\frac{\sigma-s}{\sigma+s}\right)^j + \left(\frac{jh}{s} + \frac{1}{s^2}\right)\left[e^{-sjh} - e^{-s(j-1)h}\right] + \frac{h}{s}e^{-s(j-1)h}.$$

It is clear that (3.5) has a huge number of solutions $\{c_j(h,s)\}_{j=1}^{J}$ for any fixed s and h, and most of the functions $(h,s) \mapsto c_j(h,s)$ need not even be continuous.

Claim (ii) is to be treated next. Recalling (2.7), (2.8) and (3.1)

$$
\begin{aligned}
&(L_\sigma u)(s) - (\mathcal{L}u)(s) = \int_0^T (K_{s,\sigma}(t) - e^{-st})u(t)\,dt \\
&= \int_0^T (K_{s,\sigma}(t) - e^{-st})(u(t) - (I_{h,s}u)(t))\,dt \\
&= \sum_{j=1}^J \int_{t_{j-1}}^{t_j} (K_{s,\sigma}(t) - e^{-st})(u(t) - \bar{u}_{j,h})\,dt \\
&- \sum_{j=1}^J \frac{c_j(h,s)}{h} \int_{t_{j-1}}^{t_j} (K_{s,\sigma}(t) - e^{-st})(t - t_{j-1/2})\,dt = I - II.
\end{aligned}
\tag{3.8}
$$

Let us first give an estimate to the term II. By the Poincare inequality, Proposition 6, we obtain for all $j = 1, \dots, J$

$$
\|(I - P_h)(K_{s,\sigma} - e^{-s(\cdot)})\|_{L^2(I_j)} \le \frac{h}{\pi}|K_{s,\sigma} - e^{-s(\cdot)}|_{H^1(I_j)} = \frac{h}{\pi}|e^{-s(\cdot)}|_{H^1(I_j)},
$$

where the equality follows because the function $K_{s,\sigma}$ is constant on each interval I_j. By the mean value theorem we get for $s \in \mathbb{C}_+$ and $0 \le a < b < \infty$,

$$
\begin{aligned}
&|e^{-s(\cdot)}|^2_{H^1([a,b])} = \int_a^b |\frac{d}{dt}e^{-st}|^2\,dt = \frac{|s|^2}{2\mathrm{Re}\, s}\left(e^{-2a\mathrm{Re}\, s} - e^{-2b\mathrm{Re}\, s}\right) \\
&\le \frac{|s|^2}{2\mathrm{Re}\, s} \cdot 2\mathrm{Re}\, s e^{-2\xi\mathrm{Re}\, s}(b-a) \le (b-a)|s|^2 e^{-2a\mathrm{Re}\, s}.
\end{aligned}
$$

Hence $|e^{-s(\cdot)}|_{H^1(I_j)} \le h^{1/2}|s|e^{-(j-1)h\mathrm{Re}\, s}$ and this estimate is seen to hold also for all $s \in \overline{\mathbb{C}_+}$. We now conclude that $|e^{-s(\cdot)}|_{H^1([0,T])} \le T^{1/2}|s|$ and

$$
\|(I - P_h)(K_{s,\sigma} - e^{-s(\cdot)})\|_{L^2(I_j)} \le \frac{h^{3/2}|s|}{\pi}
\tag{3.9}
$$

for all $s \in \overline{\mathbb{C}_+}$. Using (3.9) we have

$$
\begin{aligned}
II &= \sum_{j=1}^J \int_{t_{j-1}}^{t_j} (K_{s,\sigma}(t) - e^{-st}) \cdot \frac{c_j(h,s)}{h}(t - t_{j-1/2})\,dt \qquad (3.10)\\
&= \sum_{j=1}^J \int_{t_{j-1}}^{t_j} \left((I - P_h)\left(K_{s,\sigma} - e^{-s(\cdot)}\right)\right)(t) \cdot \frac{c_j(h,s)}{h}(t - t_{j-1/2})\,dt \\
&\le \sum_{j=1}^J \frac{h^{3/2}|s|}{\pi} \cdot \left[\frac{c_j(h,s)^2}{h^2}\int_{t_{j-1}}^{t_j}(t - t_{j-1/2})^2\,dt\right]^{1/2} \\
&\le \left(\sum_{j=1}^J \frac{h^3|s|^2}{\pi^2}\right)^{1/2} \cdot \left(\sum_{j=1}^J \frac{c_j(h,s)^2}{h^2}\int_{t_{j-1}}^{t_j}(t - t_{j-1/2})^2\,dt\right)^{1/2} \\
&\le \frac{h^{3/2}|s|}{\pi} J^{1/2} \cdot \|I_{h,s}u - P_h u\|_{L^2([0,T])} = \frac{hT^{1/2}|s|}{\pi}\|I_{h,s}u - P_h u\|_{L^2([0,T])}
\end{aligned}
$$

where the Schwarz inequality has been used twice, and the second to last step is by (3.4).

It remains to estimate term I in (3.8). In this case, since P_h maps on piecewise constant functions and each $u(t)-\bar{u}_{j,h}$ has zero mean on subintervals I_j, we obtain by the inequalities of Schwarz and Poincare, together with (3.9)

$$
\begin{aligned}
II \leq & \sum_{j=1}^{J} \int_{t_{j-1}}^{t_j} \left((I-P_h)\left(K_{s,\sigma}-e^{-s(\cdot)}\right)\right)(t)(u(t)-\bar{u}_{j,h})\,dt \\
\leq & \sum_{j=1}^{J} \frac{h^{3/2}|s|}{\pi} \cdot \frac{h}{\pi}|u|_{H^1(I_j)} \leq \frac{h^{5/2}|s|}{\pi^2} \sum_{j=1}^{J} |u|_{H^1(I_j)} \qquad (3.11)\\
\leq & \frac{h^{5/2}|s|}{\pi^2} \left(\sum_{j=1}^{J} 1\right)^{1/2} \left(\sum_{j=1}^{J} |u|^2_{H^1(I_j)}\right)^{1/2} = \frac{h^2 T^{1/2}|s|}{\pi^2} |u|_{H^1([0,T])}.
\end{aligned}
$$

Estimate (3.2) follows from combining (3.10) and (3.11) with (3.8).

To prove claim (iii), we shall minimise $\frac{h}{12}\sum_{j\geq 1} c_j(h,s)^2$ under the constraint (3.5), see (3.4) for motivation. We form the Langrange function

$$
\begin{aligned}
& L(c_1,\ldots,c_k\ldots,c_J,\lambda) \\
& = \frac{h}{12}\sum_{j=1}^{J} c_j^2 + \lambda\left(\sum_{j=1}^{J} \bar{u}_{j,h} I_j^{(0)}(h,s) + \sum_{j=1}^{J} c_j J_j(h,s)\right),
\end{aligned}
$$

and compute its (unique) critical point giving the minimum. We obtain

$$
\begin{cases}
\frac{\partial L}{\partial c_k} = \frac{h}{6} c_k + \lambda J_k(h,s) = 0 \quad \text{for } 1 \leq k \leq J, \\
\sum_{j=1}^{J} \bar{u}_{j,h} I_j^{(0)}(h,s) + \sum_{j=1}^{J} c_j J_j(h,s) = 0.
\end{cases}
$$

Solving this gives the minimising sequence

$$
c_k = c_k(h,s) = -\frac{6\lambda}{h} J_k(h,s) = -\frac{\sum_{j=1}^{J} \bar{u}_{j,h} I_j^{(0)}(h,s)}{\sum_{j=1}^{J} J_j(h,s)^2} J_k(h,s),
$$

for all $1 \leq k \leq J$, and then for the minimum value

$$
\begin{aligned}
& \frac{h}{12}\sum_{j=1}^{J} c_j(h,s)^2 = \frac{h}{12}\left(\frac{\sum_{j=1}^{J} \bar{u}_{j,h} I_j^{(0)}(h,s)}{\sum_{j=1}^{J} J_j(h,s)^2}\right)^2 \sum_{k=1}^{J} J_k(h,s)^2 \\
& = \frac{h}{12}\frac{\left(\sum_{j=1}^{J} \bar{u}_{j,h} I_j^{(0)}(h,s)\right)^2}{\sum_{j=1}^{J} J_j(h,s)^2}.
\end{aligned}
$$

Hence, choosing the operator $I_{h,s}$ in (3.4) optimally gives

$$
\|I_{h,s}u - P_h u\|_{L^2([0,T])} \leq \frac{\left(\sum_{j=1}^{J} I_j^{(0)}(h,s)^2\right)^{1/2}}{\left(\sum_{j=1}^{J} J_j(h,s)^2\right)^{1/2}} \frac{\|P_h u\|_{L^2([0],)}}{2\sqrt{3}}
$$

since $\|P_h u\|_{L^2([0,T])} = \left(h \sum_{j=1}^J \bar{u}_{j,h}^2\right)^{1/2}$. We must now attack (3.6) and (3.7) to estimate the required two square sums, and the resulting long computations will be done in separate subsections 3.1 and 3.2. As a final result, we get by Propositions 4 and 5

$$\frac{\left(\sum_{j=1}^J I_j^{(0)}(h,s)^2\right)^{1/2}}{\left(\sum_{j=1}^J J_j(h,s)^2\right)^{1/2}} \le \frac{5}{218}\left(3h^{-1/2}T^{-1/2} + h^{1/2}|s|^2T^{1/2}\right)$$

assuming that $9h \le T^{2/3}e^{-\frac{4}{3}|s|T}$. But then

$$h^{1/2}|s|^2T^{1/2} \le \frac{|s|}{3}\cdot|s|T^{5/6}e^{-\frac{2}{3}|s|T} \le \frac{|s|}{3}\cdot|s|Te^{-\frac{2}{3}|s|T} \le \frac{|s|}{2e},$$

since $\max_{r\ge 0} re^{-\frac{2}{3}r} = 3/(2e)$. Noting that the norm of the orthogonal projection P_h is 1, the proof of 1 is now complete. $\square$

3.1 Estimation of (3.7)

In this subsection, we shall estimate the square sum of

$$J_j(h,s) = \frac{1}{s^2}\left[e^{-sjh} - e^{-s(j-1)h}\right] + \frac{h}{2s}\left[e^{-sjh} + e^{-s(j-1)h}\right] \tag{3.12}$$

from below and above. For the first term on the left of (3.12) we obtain

$$\begin{aligned}
\frac{1}{s^2}\left[e^{-sjh} - e^{-s(j-1)h}\right] &= \frac{1}{s^2}\left[\sum_{k\ge 0}\frac{(-sjh)^k}{k!} - \sum_{k\ge 0}\frac{(-s(j-1)h)^k}{k!}\right] \\
&= \frac{1}{s^2}\left[-sh + \sum_{k\ge 2}\frac{(-sh)^k(j^k-(j-1)^k)}{k!}\right] \\
&= -\frac{h}{s} + \sum_{k\ge 2}\frac{(j^k-(j-1)^k)}{k!}(-s)^{k-2}h^k.
\end{aligned}$$

For the latter term in (3.12) we get

$$\begin{aligned}
\frac{h}{2s}\left[e^{-sjh} + e^{-s(j-1)h}\right] &= \frac{h}{s}\sum_{k\ge 0}\frac{(-s)^k(j^k+(j-1)^k)}{2k!}h^k \\
&= \frac{h}{s} - \sum_{k\ge 2}\frac{(j^{k-1}+(j-1)^{k-1})}{2(k-1)!}(-s)^{k-2}h^k.
\end{aligned}$$

Hence, for all $s \in \overline{\mathbb{C}_+}\setminus\{0\}$

$$J_j(h,s) = \sum_{k\ge 2}\frac{d_k(j)}{2k!}(-s)^{k-2}h^k$$

where the coefficient polynomials satisfy (by the binomial theorem)

$$\begin{aligned} d_k(j) &= 2\left(j^k - (j-1)^k\right) - k\left(j^{k-1} + (j-1)^{k-1}\right) \\ &= \sum_{m=0}^{k-3} \binom{k}{m} (k-m-2)(-1)^{k-m} j^m \quad \text{for} \quad k \geq 3 \end{aligned}$$

and $d_2(j) = 0$. Hence $d_k(j)$ is a polynomial of degree $k-3$ in variable j. Finally, we get

$$J_j(h,s) = \sum_{k\geq 3} \sum_{m=0}^{k-3} \frac{k-m-2}{2m!(k-m)!} (-j)^m s^{k-2} h^k.$$

Let us compute an upper estimate for

$$\|\{J_j(h,s)\}_j\|_{\ell^2} := \left(\sum_{j=1}^{J} J_j(h,s)^2\right)^{1/2}.$$

By the triangle inequality

$$\begin{aligned} &\|\{J_j(h,s)\}_j\|_{\ell^2} \\ &\leq |s^{-2}| \cdot \sum_{k\geq 3} \sum_{m=0}^{k-3} \frac{k-m-2}{2m!(k-m)!} |sh|^k \left(\sum_{j=1}^{J} j^{2m}\right)^{1/2} \\ &\leq |s^{-2}| \cdot \sum_{k\geq 3} \sum_{m=0}^{k-3} \frac{k-m-2}{2m!(k-m)!} |sh|^k \cdot \frac{J^{m+1/2}}{\sqrt{2m+1}} \\ &\leq \frac{1}{2}|s| T^{1/2} h^{5/2} \cdot \sum_{k\geq 3} \sum_{m=0}^{k-3} \frac{k-m-2}{2\sqrt{2m+1}\, m!(k-m)!} |s|^{k-3} T^m h^{k-m-3}. \end{aligned}$$

Noting that for $k-3 \geq m \geq 0$ we have $\frac{k-m-2}{\sqrt{2m+1}\, m!(k-m)!} \leq \frac{1}{m!(k-m-3)!}$ and $|s|^{k-3} T^m h^{k-m-3} = |sh|^{k-3} \cdot (T/h)^m$, we may estimate the sum term above

$$\begin{aligned} &\sum_{k\geq 3} \sum_{m=0}^{k-3} \frac{k-m-2}{2\sqrt{2m+1}\, m!(k-m)!} |s|^{k-3} T^m h^{k-m-3} \\ &\leq \sum_{k\geq 3} \left(\frac{|sh|^{k-3}}{(k-3)!} \sum_{m=0}^{k-3} \binom{k-3}{m} \left(\frac{T}{h}\right)^m \right) \\ &\leq \sum_{k\geq 3} \frac{|sh|^{k-3}}{(k-3)!} \left(1 + \frac{T}{h}\right)^{k-3} = e^{|s|(h+T)}. \end{aligned}$$

We now conclude for all $h, T > 0$ and $s \in \overline{\mathbb{C}_+} \setminus \{0\}$ that

$$\|\{J_j(h,s)\}_{j=1}^{J}\|_{\ell^2} \leq \frac{1}{2}|s| T^{1/2} h^{5/2} e^{|s|(h+T)}. \tag{3.13}$$

In addition to estimate (3.13) a lower bound can also be obtained: Decompose

$$
\begin{aligned}
J_j(h,s) &= \sum_{k=3}^{\infty}\sum_{m=0}^{k-3}\frac{k-m-2}{2m!(k-m)!}(-j)^m s^{k-2}h^k \\
=&\sum_{k=3}^{\infty}\left(\frac{1}{2(k-3)!3!}(-j)^{k-3}s^{k-2}h^k+\sum_{m=0}^{k-4}\frac{k-m-2}{2m!(k-m)!}(-j)^m s^{k-2}h^k\right) \\
=&\sum_{k=3}^{\infty}\frac{1}{2(k-3)!3!}(-j)^{k-3}s^{k-2}h^k+\sum_{k=4}^{\infty}\sum_{m=0}^{k-4}\frac{k-m-2}{2m!(k-m)!}(-j)^m s^{k-2}h^k
\end{aligned}
$$

so that by the triangle inequality

$$
\begin{aligned}
\left\|\{J_j(h,s)\}_{j=1}^J\right\|_{\ell^2} \geq& \left\|\left\{\sum_{k=3}^{\infty}\frac{1}{2(k-3)!3!}(-j)^{k-3}s^{k-2}h^k\right\}_{j=1}^J\right\|_{\ell^2} \\
&- \left\|\left\{\sum_{k=4}^{\infty}\sum_{m=0}^{k-4}\frac{k-m-2}{2m!(k-m)!}(-j)^m s^{k-2}h^k\right\}_{j=1}^J\right\|_{\ell^2}.
\end{aligned} \tag{3.14}
$$

For the first term in the right hand side of (3.14) we have

$$
\begin{aligned}
&\left\|\left\{\sum_{k=3}^{\infty}\frac{1}{2(k-3)!3!}(-j)^{k-3}s^{k-2}h^k\right\}_{j=1}^J\right\|_{\ell^2} \\
&= \left\|\left\{\frac{1}{12}sh^3\sum_{k=3}^{\infty}\frac{1}{(k-3)!}(-j)^{k-3}s^{k-3}h^{k-3}\right\}_{j=1}^J\right\|_{\ell^2} \\
&= \frac{1}{12}|s|h^3\cdot\left\|\{e^{-jsh}\}_{j=1}^J\right\|_{\ell^2}
\end{aligned} \tag{3.15}
$$

where

$$
\begin{aligned}
\left\|\{e^{-jsh}\}_{j=1}^J\right\|_{\ell^2} &= \sum_{j=1}^{J}|e^{-jsh}|^2 \\
&= \begin{cases} J=h^{-1}T, & \text{when } \operatorname{Re}s=0 \\ e^{-2h\operatorname{Re}s}\frac{1-e^{-2(J+1)h\operatorname{Re}s}}{1-e^{-2h\operatorname{Re}s}}, & \text{when } \operatorname{Re}s>0. \end{cases}
\end{aligned} \tag{3.16}
$$

For the latter term in (3.14) we have a similar upper estimate to (3.13). Indeed,

$$\begin{aligned}
&\left\|\left\{\sum_{k=4}^{\infty}\sum_{m=0}^{k-4}\frac{k-m-2}{2m!(k-m)!}(-j)^m s^{k-2}h^k\right\}_{j=1}^{J}\right\|_{\ell^2}\\
\leq&\sum_{k=4}^{\infty}\sum_{m=0}^{k-4}\frac{k-m-2}{2m!(k-m)!}|s|^{k-2}h^k\frac{J^{m+1/2}}{\sqrt{2m+1}}\\
=&\sum_{k=4}^{\infty}\sum_{m=0}^{k-4}\frac{k-m-2}{2m!(k-m)!}|s|^{k-2}h^k h^{-m-1/2}T^{m+1/2}\\
=&|s|^2h^{7/2}\sum_{k=4}^{\infty}\sum_{m=0}^{k-4}\frac{k-m-2}{2m!(k-m)!}|s|^{k-4}h^{k-m-4}T^m\\
\leq&|s|h^{7/2}e^{|s|(h+T)}.
\end{aligned}\tag{3.17}$$

As a conclusion we can now state

Proposition 4. *Let $J_j(h,s)$ be defined through* (3.12). *Then for any $s\in i\mathbb{R}$, $T,h>0$ satisfying $T=Jh$, $J\in\mathbb{N}$ and $9h\leq T^{2/3}e^{-\frac{4}{3}|s|T}$ we have*

$$\|\{J_j(h,s)\}_{j=1}^{J}\|_{\ell^2}\geq\frac{5}{109}Th^2|s|.\tag{3.18}$$

Proof. It is clear that (3.18) is satisfied for $s=0$. For $s\in i\mathbb{R}\setminus\{0\}$ it follows from (3.14) and (3.15) – (3.17) that for all $s\in i\mathbb{R}\setminus\{0\}$, $h,T>0$ satisfying $T=Jh$ for $J\in\mathbb{N}$ that the estimate

$$\left\|\{J_j(h,s)\}_{j=1}^{J}\right\|_{\ell^2}\geq\left(\frac{T}{12}-h^{3/2}e^{|s|(h+T)}\right)h^2|s|$$

holds. Since always $h\leq T$, we have $h^{3/2}e^{|s|(h+T)}\leq h^{3/2}e^{2|s|T}\leq\frac{T}{27}$ provided that $h\leq\frac{T^{2/3}}{9}e^{-\frac{4}{3}|s|T}$. The claim follows from this. □

3.2 Estimation of (3.6)

In this subsection, we compute an upper estimate for

$$\left\|\left\{I_j^{(0)}(h,s)\right\}_{j=1}^{J}\right\|_{\ell^2}:=\left(\sum_{j=1}^{J}I_j^{(0)}(h,s)^2\right)^{1/2}.$$

Writing $\tau=sh$ and recalling $\sigma=2/h$, we get for $s\in\overline{\mathbb{C}_+}$

$$\begin{aligned}
I_j^{(0)}(h,s)=&\frac{2}{\sigma+s}\left(\frac{\sigma-s}{\sigma+s}\right)^j+\frac{1}{s}\left(e^{-sjh}-e^{-s(j-1)h}\right)\\
=&\frac{2}{\sigma+s}\left(\left(\frac{\sigma-s}{\sigma+s}\right)^j-e^{-sjh}\right)+\left(\frac{2}{\sigma+s}-\frac{1}{s}(e^{sh}-1)\right)e^{-sjh}\\
=&\frac{2h}{2+\tau}\left(\left(\frac{2-\tau}{2+\tau}\right)^j-e^{-\tau j}\right)+\left(\frac{2h}{2+\tau}-\frac{h}{\tau}(e^{\tau}-1)\right)e^{-\tau j}.
\end{aligned}$$

Let $\Omega \subset \overline{\mathbb{C}_+}$ be any set. Then for any $\tau \in \Omega$ we have

$$
\begin{aligned}
|I_j^{(0)}(h,s)| \leq & \left|\frac{2h}{2+\tau}\right| \left| \left(\frac{2-\tau}{2+\tau}\right)^j - e^{-\tau j} \right| + \left| \frac{2h}{2+\tau} - \frac{h}{\tau}(e^\tau - 1) \right| \left|e^{-\tau j}\right| \\
\leq & \left|\frac{2h}{2+\tau}\right| \left| \left(\frac{2-\tau}{2+\tau}\right) - e^{-\tau} \right| \left| \sum_{k=1}^{j-1} \left(\frac{2-\tau}{2+\tau}\right)^k e^{-\tau(j-k-1)} \right| \\
& + \left| \frac{2h}{2+\tau} - \frac{h}{\tau}(e^\tau - 1) \right| \\
\leq & h|\tau| \left(C_\Omega \left| \frac{2j\tau^2}{2+\tau} \right| + C'_\Omega \right)
\end{aligned}
$$

where the constants are given by

$$
C_\Omega = \sup_{\tau\in\Omega} \left| \frac{1}{\tau^3} \left(\frac{2-\tau}{2+\tau} - e^{-\tau} \right) \right| \text{ and } C'_\Omega = \sup_{\tau\in\Omega} \left| \frac{1}{\tau} \left(\frac{2}{2+\tau} - \frac{1}{\tau}(e^\tau - 1) \right) \right|.
$$

This implies for all $h \geq 0$ and $\tau = sh \in \Omega$

$$
\begin{aligned}
\left\| \left\{ I_j^{(0)}(h,s) \right\}_{j=1}^J \right\|_{\ell^2} \leq & C_\Omega \frac{2h|\tau|^3}{|2+h|} \left(\sum_{j=1}^J j^2 \right)^{1/2} + C'_\Omega h|\tau| \left(\sum_{j=1}^J 1 \right)^{1/2} \\
\leq & C_\Omega h^4 |s|^3 \left(\frac{1}{3}J^3 + \frac{1}{2}J^2 + \frac{1}{6}J \right)^{1/2} + C'_\Omega h^2 |s| J^{1/2} \qquad (3.19) \\
\leq & C_\Omega h^{5/2} |s|^3 T^{3/2} + C'_\Omega h^{3/2} |s| T^{1/2}
\end{aligned}
$$

by the facts that $T = Jh$ and $J \geq 1$. We now have to choose the set Ω in a clever way, so that the resulting estimate is properly "fine tuned" according to Proposition 4.

Proposition 5. *Let $I_j^{(0)}(h,s)$ be defined through* (3.6)*. Then for any $s \in i\mathbb{R}$, $T \geq 1, h > 0$ satisfying $T = Jh$, $J \in \mathbb{N}$ and $9h \leq T^{2/3} e^{-\frac{4}{3}|s|T}$ we have*

$$
\left\| \left\{ I_j^{(0)}(h,s) \right\}_{j=1}^J \right\|_{\ell^2} \leq \frac{1}{2} h^{5/2} |s|^3 T^{3/2} + \frac{3}{2} h^{3/2} |s| T^{1/2} \qquad (3.20)
$$

Proof. Since we assume (motivated by Proposition 4) that $9h \leq T^{2/3} e^{-\frac{4}{3}|s|T}$, we have

$$
|\tau| = |s|h \leq \frac{|s|T^{2/3}}{9} e^{-\frac{4}{3}|s|T} \leq \frac{|s|T}{9} e^{-\frac{4}{3}|s|T} \leq \frac{1}{12e},
$$

since $\max_{r\geq 0} re^{-\frac{4}{3}r} = 3/(4e)$. Hence, we are invited to estimate the constants C_Ω and C'_Ω for the set $\Omega := [-i/(12e), i/(12e)]$. By computing the Taylor series, we see that

$$
\begin{aligned}
C_\Omega \leq & \sum_{j\geq 0} \left| \frac{1}{2^{j+2}} - \frac{1}{(j+3)!} \right| \cdot \left(\frac{1}{12e} \right)^j < \sum_{j\geq 0} \frac{1}{2^{j-1}} \cdot \left(\frac{1}{12e} \right)^j \\
= & \frac{6e}{24e-1} < \frac{1}{2}.
\end{aligned}
$$

Similarly

$$C'_\Omega \le \sum_{j\ge 0}\left|\left(-\frac{1}{2}\right)^{j+1} - \frac{1}{(j+2)!}\right| \cdot \left(\frac{1}{12e}\right)^j < \sum_{j\ge 0}\frac{1}{2^j}\cdot\left(\frac{1}{12e}\right)^j$$
$$= \frac{24e}{24e-1} < \frac{3}{2}.$$

But now (3.19) implies (3.20). □

3.3 Determination of the isoperimetric constant

In this section we give a basic interpolation estimate used several times in the proofs.

Proposition 6. *Assume that* $u \in H^1(I_j)$. *Then*

$$||u-\bar{u}||_{L^2(I_j)} \le \frac{h}{\pi}|u|_{H^1(I_j)}$$

Proof. Let $I_{ref} = (0,1]$ and define the bilinear forms $a(u,v) = \int_{I_{ref}} u'(v')^*\,dt$ and $b(u,v) = \int_{I_{ref}} uv^*\,dt$ where the asterisk denotes complex conjugation. Furthermore, let

$$V = \{v \in H^1(I_{ref}) \mid \int_{I_{ref}} v(t)\,dt = 0\}$$

and

$$\lambda_1 = \inf_{v\in V,\, v\neq 0}\frac{a(v,v)}{b(v,v)} \in \mathbb{R}^+$$

By Rayleigh's theorem, λ_1 is the smallest eigenvalue of the problem: Find $u \in V$ such that

$$a(u,v) = \lambda b(u,v) \quad \forall v \in V. \tag{3.21}$$

Solution to (3.21) can be sought for using the Euler equations for the eigenpair (λ, u). By standard calculus the first eigenpair is found to be $(\lambda_1, u_1) = (\pi^2, \cos(\pi t))$. It follows that $b(v,v) \le \frac{1}{\lambda_1}a(v,v)$, that is $||v||^2_{L^2(I_{ref})} \le \frac{1}{\pi^2}|v|^2_{H^1(I_{ref})}$ for any $v \in V$. Let now $u \in H^1(I_{ref})$ and set $v = u - \bar{u} \in V$ implying

$$||u-\bar{u}||^2_{L^2(I_{ref})} \le \frac{1}{\pi^2}|u-\bar{u}|^2_{H^1(I_{ref})} = \frac{1}{\pi^2}|u|^2_{H^1(I_{ref})} \tag{3.22}$$

For the general interval $I_j = (t_{j-1}, t_j]$ a standard scaling argument with $\hat{u}(\tau) = u((t-t_{j-1})/h)$ and $\tau = (t-t_{j-1})/h \in I_{ref}$ gives

$$||u-\bar{u}||^2_{L^2(I_j)} = h||\hat{u}-\bar{\hat{u}}||^2_{L^2(I_{ref})} \le \frac{1}{\pi^2}h|\hat{u}|^2_{H^1(I_{ref})} = \frac{1}{\pi^2}h^2|u|^2_{H^1(I_j)} \tag{3.23}$$

implying

$$||u-\bar{u}||_{L^2(I_j)} \le \frac{1}{\pi}h|u|_{H^1(I_j)}. \tag{3.24}$$

□

4 Weak and strong convergence

We first show that Theorem 1 implies that $L_\sigma \to \mathcal{L}$ in weak operator topology. Using this, it is then shown in Theorem 2 that the convergence is, in fact, strong.

Indeed, it follows from Theorem 1 that $(L_\sigma u)(i\omega) \to (\mathcal{L}u)(i\omega)$ uniformly in the compact subsets $i\omega \in K \subset i\mathbb{R}$ for any $u \in C_c(\mathbb{R}_+) \cap H^1(\mathbb{R}_+)$. Hence, for finite linear combinations s (also called simple functions) of characteristic functions χ_K of compact intervals $K \subset i\mathbb{R}$ we have $\langle s, L_\sigma u\rangle_{L^2(i\mathbb{R})} \to \langle s, \mathcal{L}u\rangle_{L^2(i\mathbb{R})}$. Since $\|L_\sigma\|_{\mathcal{L}(L^2(\mathbb{R}_+);H^2(\mathbb{C}_+))} \leq 1$ and simple functions are dense in $L^2(i\mathbb{R})$, it follows that

$$\langle v, L_\sigma u\rangle_{K^2(i\mathbb{R})} \to \langle v, \mathcal{L}u\rangle_{H^2(i\mathbb{R})} \text{ as } \sigma \to \infty \tag{4.1}$$

for all $u \in C_c(\mathbb{R}) \cap H^1(\mathbb{R}_+)$ and $v \in L^2(i\mathbb{R}_+)$. Another density argument implies finally that (4.1) holds even for all $u \in L^2(\mathbb{R}_+)$ and $v \in L^2(i\mathbb{R}_+)$.

We recall a result from elementary functional analysis:

Proposition 7. *Let H be a Hilbert space, and assume that $u_j \to u$ weakly in H. If $\|u_j\|_H \to \|u\|_H$, then $u_j \to u$ in the norm of H.*

Proof. $\langle u_j - u, u_j - u\rangle_H = \langle u_j, u_j\rangle_H - \langle u, u\rangle_H - \langle u, u_j - u\rangle_H - \langle u_j - u, u\rangle_H = \|u_j\|_H^2 - \|u\|_H^2 - 2\mathrm{Re}\,\langle u, u_j - u\rangle_H$. □

Theorem 2. *We have $\|L_\sigma u - \mathcal{L}u\|_{H^2(\mathbb{C}_+)} \to 0$ for any $u \in L^2(\mathbb{R}_+)$. Moreover, $\|L_\sigma^* v - \mathcal{L}^* v\|_{L^2(\mathbb{R}_+)} \to 0$ for any $v \in H^2(\mathbb{C}_+)$.*

Proof. Adjoining (4.1) shows that $L_\sigma^* v \to \mathcal{L}^* v$ weakly. Since L_σ is a coisometry by Proposition 2 and (2.5), we have

$$\|L_\sigma^* v\|^2_{L^2(\mathbb{R}_+)} = \langle L_\sigma L_\sigma^* v, v\rangle^2_{H^2(\mathbb{C}_+)} = \|v\|^2_{H^2(\mathbb{C}_+)}.$$

Now Proposition 7 implies the latter part of this Theorem.

To show the first part, we have to work a bit harder to verify that $\|L_\sigma u\|_{L^2(i\mathbb{R})} \to \|u\|_{L^2(\mathbb{R}_+)} = \|\mathcal{L}u\|_{L^2(i\mathbb{R})}$. Suppose that $h = 2/\sigma > 0$ and $u \in L^2(\mathbb{R}_+)$ is such that $u(t) = \overline{u}_{j,h} := \int_{((j-1)h,jh]} u(t)\,dt$ for all $t \in I_j := ((j-1)h, jh]$ — in other words, this is simply $u = P_h u$. For such u

$$\|u\|^2_{L^2(\mathbb{R}_+)} = \sum_{j\geq 1} \int_{I_j} |u(t)|^2\,dt = h\|\{\overline{u}_{j,h}\}_{j\geq 0}\|^2_{\ell^2}.$$

By the definition of the discretizing operator T_σ, we have

$$\|T_\sigma u\|^2_{H^2(\mathbb{D})} = \sum_{j\geq 1} \left(\frac{1}{\sqrt{h}} \int_{I_j} |u(t)|^2\,dt\right)^2 = h\sum_{j\geq 1} |\overline{u}_{j,h}|^2 = \|u\|^2_{L^2(\mathbb{R}_+)}.$$

Hence, we have $\|T_\sigma P_h u\|_{H^2(\mathbb{D})} = \|P_h u\|_{L^2(\mathbb{R}_+)}$ for all $u \in L^2(\mathbb{R}_+)$ where $\sigma = 2/h$. Also note that $T_\sigma u = T_\sigma P_h u$ for all $u \in L^2(\mathbb{R}_+)$ provided that $\sigma = 2/h$.

We now have for any $u \in L^2(\mathbb{R}_+)$

$$
\begin{aligned}
&\left| \|T_\sigma u\|_{H^2(\mathbb{D})} - \|u\|_{L^2(\mathbb{R}_+)} \right| \\
&\leq \left| \|T_\sigma u\|_{H^2(\mathbb{D})} - \|T_\sigma P_h u\|_{H^2(\mathbb{D})} \right| + \left| \|T_\sigma P_h u\|_{H^2(\mathbb{D})} - \|P_h u\|_{L^2(\mathbb{R}_+)} \right| \\
&+ \left| \|P_h u\|_{L^2(\mathbb{R}_+)} - \|u\|_{L^2(\mathbb{R}_+)} \right| = \left| \|P_h u\|_{L^2(\mathbb{R}_+)} - \|u\|_{L^2(\mathbb{R}_+)} \right|
\end{aligned}
$$

where again $\sigma = 2/h$. Since the projections $P_h \to I$ strongly in $L^2(\mathbb{R}_+)$ as $h \to 0$, we conclude that $\|T_\sigma u\|_{H^2(\mathbb{D})} \to \|u\|_{L^2(\mathbb{R}_+)}$ and hence $\|L_\sigma u\|_{H^2(\mathbb{C}_+)} \to \|u\|_{L^2(\mathbb{R}_+)}$ as $\sigma \to \infty$, see Proposition 2. The first claim of this theorem follows from this, Proposition 7 and (4.1). □

Using Theorem 2 we can show that the output of integration scheme (1.5) converges to the output of continuous time dynamics (1.3) for *input/output stable* systems S. These are systems for which $\mathbf{G}(\cdot) \in H^\infty(\mathbb{C}_+)$ or, equivalently, $\mathbf{G} \in \mathcal{L}(H^2(\mathbb{C}_+))$. To understand the formulation of the following theorem, we refer back to Section 2.

Theorem 3. *For any $u \in L^2(\mathbb{R}_+)$ and $\mathbf{G} \in H^\infty(\mathbb{C}_+)$, we have*

$$
\|T_\sigma^* \mathbf{D}_\sigma T_\sigma u - \mathcal{L}^* \mathbf{G} \mathcal{L} u\|_{L^2(\mathbb{R}_+)} \to 0 \tag{4.2}
$$

as $\sigma \to \infty$.

Proof. As noted just before Proposition 3, we have $T_\sigma^* \mathbf{D}_\sigma T_\sigma = L_\sigma^* \mathbf{G} L_\sigma$. Then we get for all $\sigma > 0$

$$
\begin{aligned}
&\|L_\sigma^* \mathbf{G} L_\sigma u - \mathcal{L}^* \mathbf{G} \mathcal{L} u\|_{L^2(\mathbb{R}_+)} \leq \|(L_\sigma^* - \mathcal{L}^*)\, \mathbf{G}\, (L_\sigma u - \mathcal{L} u)\|_{L^2(\mathbb{R}_+)} \\
&+ \|(L_\sigma^* - \mathcal{L}^*)\, \mathbf{G} \mathcal{L} u\|_{L^2(\mathbb{R}_+)} + \|\mathcal{L}^* \mathbf{G}\, (L_\sigma u - \mathcal{L} u)\|_{L^2(\mathbb{R}_+)}.
\end{aligned}
$$

Now (4.2) follows by Theorem 2. □

5 A counterexample

We complete this paper by reviewing estimate (2.6) in the special case when $\mathbf{G}(s) = I$ for all $s \in \mathbb{C}_+$. It indicates that Theorem 3 cannot be improved by a speed estimate for convergence.

In this special case it follows from the very definitions that $L_\sigma^* \mathbf{G} L_\sigma = T_\sigma^* T_\sigma = P_{2/\sigma}$ where the orthogonal projection P_h is defined as in Section 3. Since $\mathcal{L}^* \mathcal{L} = \mathcal{I}$ on all of $L^2(\mathbb{R}_+)$, we should give an estimate to

$$
\|u - P_h u\|_{L^2([0,T])} \quad \text{for a family of functions} \quad u \in L^2(\mathbb{R}_+).
$$

It is, of course, true that $P_h u \to u$ as $h \to 0$ for all $u \in L^2(\mathbb{R}_+)$. However, there cannot be a uniform speed estimate of type

$$
\|u - P_h u\|_{L^2([0,T])} \leq C_u h^\alpha \tag{5.1}
$$

where $C_u < \infty$ for *all* $u \in L^2([0,T])$. If it were so, then for any $0 < \beta < \alpha$ we would have $\|h^{-\beta}(I - P_h)u\|_{L^2([0,T])} \leq C_u h^{\alpha-\beta} \to 0$ as $h \to 0$, for all $u \in L^2([0,T])$. By the uniform boundedness principle,

$$\sup_{h>0} \|h^{-\beta}(I - P_h)\|_{L^2([0,T])} =: M < \infty$$

and hence $\|(I - P_h)\|_{\mathcal{L}(L^2([0,T]))} \leq M h^\beta$ for all $h > 0$.

Making now h small enough, we see that then the norm of the orthogonal projection $(I - P_h)|L^2([0,T])$ is strictly less than 1; this implies that $I|L^2([0,T]) = P_h|L^2([0,T])$. But $P_h|L^2([0,T])$ is a finite rank operator, and the uniform speed estimate (5.1) cannot hold by contradiction. The same conclusion holds, if h^α in (5.1) is replaced by *any* increasing continuous function $\phi(h)$ satisfying $\phi(0) = 0$.

It should also be noted that for functions $u \in L^2(\mathbb{R}_+)$ that possess certain smoothness properties such a speed estimate can be obtained. See [2] for a further discussion on what is obtainable and what is not.

6 Conclusions

The operators L_σ for $\sigma > 0$ have been introduced just before Proposition 3 with aid of the Cayley transformation (1.7). It is shown in Theorem 2 that the operators L_σ provide an approximation to Laplace transform for a wide class of functions. In addition, Theorem 3 shows that for I/O-stable linear systems, the convergence extends to the input/output relation of the system. All this can be anticipated since the Cayley transform actually corresponds to the slightly "unorthodox", conservativity-preserving discretization (1.5) of the dynamical equations (1.3) (or their infinite-dimensional analogue e.g. in [8, Proposition 2.5]).

Theorem 3 gives no estimate on the speed of the convergence with respect to the sampling parameter $h = 2/\sigma$. If we had some decay

$$\mathbf{G}(s) \to 0 \quad \text{as} \quad |s| \to \infty \tag{6.1}$$

at some speed, then we could effectively restrict our analysis to compact subsets of $i\mathbb{R}$. Then the speed estimate of Theorem 1 could show up in (4.2) in some form. Unfortunately, (6.1) is not a generic property of $\mathbf{G} \in H^\infty(\mathbb{C}_+)$ – hence it is not a generic property of the transfer functions of conservative systems either.

In the time domain, the same problem appears because the sampling operator T_σ cannot detect above a certain cutoff frequency: there are always high-frequency signals carrying substantial energy that a given discretized system cannot capture. To achieve a speed estimate in (4.2), one could assume either

(i) that the high frequencies are damped by the linear system itself (e.g. by a property like (6.1)), or

(ii) that the high frequencies have a small amplitude in the signal u (e.g. an assumption such as $u \in H^1(\mathbb{R}_+)$ in Theorem 1).

The approximation of the state trajectory $x(\cdot)$ by the discrete trajectories $\{x_j^{(h)}\}_{j\geq 0}$ solving (1.5) has not been studied here. This will be carried out in a future paper on the state space approximation for conservative systems.

Remark 1. *We remark that practically all of the results presented in this paper hold if the input space of the node S is a separable Hilbert space instead of $\mathbb{C}$.*

References

[1] D. Arov and I. Gavrilyuk. A method for solving initial value problems for linear differential equations in Hilbert space based on the Cayley transform. *Numerical Functional Analysis and Optimization*, 14(5&6):459–473, 1993.

[2] D. Braess. *Finite elements. Theory, fast solvers and applications in solid mechanics.* Cambridge University Press, 2 edition, 2001.

[3] J. W. Helton. Systems with infinite dimensional state space: The Hilbert space approach. *Proceedings of the IEEE*, 64(1):145–160, 1976.

[4] M. S. Brodskiĭ. On operator colligations and their characteristic functions. *Soviet Mat. Dokl.*, 12:696–700, 1971.

[5] M. S. Brodskiĭ. *Triangular and Jordan representations of linear operators*, volume 32. American Mathematical Society, Providence, Rhode Island, 1971.

[6] M. S. Brodskiĭ. Unitary operator colligations and their characteristic functions. *Russian Math. Surveys*, 33(4):159–191, 1978.

[7] J. Malinen. Conservativity of time-flow invertible and boundary control systems. *Helsinki University of Technology Institute of Mathematics Research Reports*, A479, 2004.

[8] J. Malinen, O. Staffans, and G. Weiss. How to characterize conservative systems? *Mittag–Leffler preprints 2003/46; Submitted*, 2003.

[9] J. Malinen and O. J. Staffans. Conservative boundary control systems. Manuscript, 40pp., 2004.

[10] R. Ober and S. Montgomery-Smith. Bilinear transformation of infinite-dimensional state-space systems and balanced realizations of nonrational transfer functions. *SIAM Journal of Control and Optimization*, 28(2):438–465, 1990.

[11] M. S. Livšic and A. A. Yantsevich. *Operator colligations in Hilbert space.* John Wiley & sons, Inc., 1977.

[12] M.S. Livšic. *Operators, vibrations, waves. Open systems.* Nauka, Moscow, 1966.

[13] Yu. L. Smuljan. Invariant subspaces of semigroups and Lax–Phillips scheme. *Dep. in VINITI, No. 8009-B86, Odessa (A private translation by Daniela Toshkova, 2001)*, 1986.

[14] O. J. Staffans. *Well-Posed Linear Systems.* Cambridge University Press, 2004.

[15] B. Sz.-Nagy and C. Foias. *Harmonic Analysis of Operators on Hilbert space.* North-Holland Publishing Company, Amsterdam, London, 1970.

(continued from the back cover)

HELSINKI UNIVERSITY OF TECHNOLOGY INSTITUTE OF MATHEMATICS
RESEARCH REPORTS

The list of reports is continued inside. Electronical versions of the reports are available at *http://www.math.hut.fi/reports/* .

A477 Tuomo T. Kuusi
Moser's Method for a Nonlinear Parabolic Equation
October 2004

A476 Dario Gasbarra , Esko Valkeila , Lioudmila Vostrikova
Enlargement of filtration and additional information in pricing models: a Bayesian approach
October 2004

A473 Carlo Lovadina , Rolf Stenberg
Energy norm a posteriori error estimates for mixed finite element methods
October 2004

A472 Carlo Lovadina , Rolf Stenberg
A posteriori error analysis of the linked interpolation technique for plate bending problems
September 2004

A471 Nuutti Hyvönen
Diffusive tomography methods: Special boundary conditions and characterization of inclusions
April 2004

ISBN 951-22-7453-1
ISSN 0748-3143
Espoo, 2004