Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Laplace and Cayley transforms – an approximation point of view

Ville Havu
Insitute of Mathematics, TKK
Box 1100 FI-02015
Helsinki University of Technology, Finland
Ville.Havu@tkk.fi

Jarmo Malinen
Insitute of Mathematics, TKK
Box 1100 FI-02015
Helsinki University of Technology, Finland
Jarmo.Malinen@tkk.fi

***Abstract*— We interpret the Cayley transform of linear (finite- or infinite-dimensional) state space systems as a numerical integration scheme of Crank–Nicolson type. If such a scheme is applied to a *conservative* system, then the resulting discrete time system is conservative in the discrete time sense. We show that the convergence of this integration scheme is equivalent to an approximation of the Laplace transform.**

I. INTRODUCTION

In this paper, we consider convergence results for the time discretization scheme of type (2) for a linear (finite- or infinite-dimensional) state space dynamical systems. In finite-dimensional case, such systems are described by (1) but it is necessary to use more general equations (7) and (8) in infinite dimensions. In the infinite-dimensional case, also discretization (2) has to be generalized.

We show below how discretization (2) is induced by the Cayley transform (in the sense of linear system theory). Hence it has the the following important property: if the original continuous time dynamics is conservative (as defined in Subsection I-B), then the resulting discrete time dynamics satisfies a similar energy balance law. Since this is not a typical property of an arbitrary time discretization scheme, it is well-motivated to study the generalization of scheme (2) in the context of infinite-dimensional conservative linear systems. The presented techniques can be used for simulation of conservative systems governed by PDEs arising from applications in physics and engineering.

For approaches parallel to our work, see e.g. [4], [5].

A. Finite Dimensional Motivation

We consider first the finite-dimensional state space with scalar signals. Then the system S is described by the dynamical equations

$$S: \begin{cases} x'(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t) + Du(t), \quad t \geq 0, \\ x(0) = x_0, \end{cases} \tag{1}$$

where $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times 1}$, $C \in \mathbb{C}^{1\times n}$, and $D \in \mathbb{C}$. Given a *discretization parameter* $h > 0$, a slightly non-standard time discretization of (1) of Crank–Nicolson type is given by

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

for $j \geq 1$. This induces the discrete time dynamics

$$\begin{cases} \frac{x_j^{(h)}-x_{j-1}^{(h)}}{h} &= A\frac{x_j^{(h)}+x_{j-1}^{(h)}}{2} + B\frac{u_j^{(h)}}{\sqrt{h}}, \\ \frac{y_j^{(h)}}{\sqrt{h}} &= C\frac{x_j^{(h)}+x_{j-1}^{(h)}}{2} + D\frac{u_j^{(h)}}{\sqrt{h}}, \quad j \geq 1, \\ x_0^{(h)} &= x_0, \end{cases} \tag{3}$$

where $u_j^{(h)}/\sqrt{h}$ is an approximation to $u(jh)$. The purpose of this paper is to show under rather general assumptions that $y_j^{(h)}/\sqrt{h}$ converges to $y(jh)$ as $h \to 0$. After some computations, equations (3) take the form

$$\phi_\sigma : \begin{cases} x_j^{(h)} = \mathbf{A}_\sigma x_{j-1}^{(h)} + \mathbf{B}_\sigma u_j^{(h)}, \\ y_j^{(h)} = \mathbf{C}_\sigma x_{j-1}^{(h)} + \mathbf{D}_\sigma u_j^{(h)}, \quad j \geq 1, \\ x_0^{(h)} = x_0, \end{cases} \tag{4}$$

where $\sigma := 2/h$, and the operators $\mathbf{A}_\sigma$, $\mathbf{B}_\sigma$, $\mathbf{C}_\sigma$ and $\mathbf{D}_\sigma$ comprise *the discrete time linear system* (henceforth, DLS)

$$\begin{aligned} \phi_\sigma &= \begin{bmatrix} \mathbf{A}_\sigma & \mathbf{B}_\sigma \\ \mathbf{C}_\sigma & \mathbf{D}_\sigma \end{bmatrix} \\ &= \begin{bmatrix} (\sigma+A)(\sigma-A)^{-1} & \sqrt{2\sigma}(\sigma-A)^{-1}B \\ \sqrt{2\sigma}C(\sigma-A)^{-1} & \mathbf{G}(\sigma) \end{bmatrix}. \end{aligned} \tag{5}$$

Here $\mathbf{G}(\cdot)$ denotes the transfer function of system S in (1), and it is defined by $\mathbf{G}(s) = C(s-A)^{-1}B + D$ for all $s \in \rho(A)$. Then the transfer function $\mathbf{D}_\sigma(\cdot)$ of ϕ_σ satisfies

$$\begin{aligned} \mathbf{D}_\sigma(z) :=& \mathbf{D}_\sigma + z\mathbf{C}_\sigma(I - z\mathbf{A}_\sigma)^{-1}\mathbf{B}_\sigma \\ =& \mathbf{G}\left(\frac{1-z}{1+z}\sigma\right) \end{aligned} \tag{6}$$

for all $z \in \rho(\mathbf{A}_\sigma)$. The mapping $S \mapsto \phi_\sigma$ described above is called *the Cayley transform* of continuous time systems to discrete time systems. As described above, ϕ_σ can always be regarded as a time discretization of S.

0-7803-9568-9/05/$20.00 ©2005 IEEE

B. Infinite Dimensional Systems

Even though we have considered above only matrix systems (1), the Cayley transform can be defined similarly to (5) for any *conservative system node* S. Let us state first what we mean by such S.

Let $S = \left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ be a system node on the separable Hilbert spaces (U, X, U) in the sense of [1, Definition 2.2] with domain denoted by $\mathrm{dom}\,(S)$. By A_{-1} denote the usual extension of the main operator A of S. Then, as it is well-known, the Cauchy problem associated to S

$$\begin{cases} x'(t) = A_{-1}x(t) + Bu(t), \quad t \geq 0, \\ x(0) = x_0 \end{cases} \tag{7}$$

is uniquely solvable for any input $u \in C^2(\mathbb{R}_+; U)$ and any initial state $x_0 \in X$ for which the compatibility condition $\left[\begin{smallmatrix} x_0 \\ u(0) \end{smallmatrix}\right] \in \mathrm{dom}\,(S)$ holds. Moreover, $\left[\begin{smallmatrix} x(\cdot) \\ u(\cdot) \end{smallmatrix}\right] \in C(\mathbb{R}_+; \mathrm{dom}\,(S))$ and because $C\&D \in \mathcal{L}(\mathrm{dom}\,(S); U)$, the output signal given by

$$y(t) = C\&D \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} \tag{8}$$

is well defined for all $t \geq 0$. These and many other facts can be found in [1, Section 2].

The system node S is *(scattering) energy preserving* if for all $T > 0$ *the energy balance*

$$\|x(T)\|_X^2 + \int_0^T \|y(t)\|_Y^2 dt = \|x_0\|_X^2 + \int_0^T \|u(t)\|_U^2 dt \tag{9}$$

holds, where u, x, y and x_0 are as in (7) and (8). 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 scattering energy preserving, then $\left[\begin{smallmatrix} A\&B \\ C\&D \end{smallmatrix}\right]$ is called *(scattering) conservative*; see [1, Definitions 3.1 and 4.1].

As is discussed in [1], the Cayley transform can be extended to energy preserving system nodes S. Indeed, we define for any $\sigma > 0$ the Cayley transform of S as the DLS given by

$$\phi_\sigma = \begin{bmatrix} (\sigma + A)(\sigma - A)^{-1} & \sqrt{2\sigma}(\sigma - A_{-1})^{-1}B \\ \sqrt{2\sigma}C(\sigma - A)^{-1} & \mathbf{G}(\sigma) \end{bmatrix}. \tag{10}$$

When comparing to the matrix formula (5), we see that A has been replaced by its extension A_{-1}. Also the definition of the transfer function $\mathbf{G}(\cdot)$ must be generalized, and it is now given by $\mathbf{G}(s) = C\&D \begin{bmatrix} (s - A_{-1})^{-1}B & I \end{bmatrix}^T$ for all $s \in \mathbb{C}_+$. The relation between $\mathbf{G}(\cdot)$ and $\mathbf{D}_\sigma(\cdot)$ is described by (6) without change.

C. Conservativity is preserved

The motivation for the study of the discretization scheme (3) lies in the fact that conservative characteristics of the system are preserved.

We say that the DLS $\phi = \left[\begin{smallmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{smallmatrix}\right]$ on Hilbert spaces (U, X, U) is *energy preserving* if the block matrix $\left[\begin{smallmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{smallmatrix}\right]$ is isometric on $\left[\begin{smallmatrix} X \\ U \end{smallmatrix}\right]$. Then, and only then, the discrete time balance equation

$$\|x_N\|^2 - \|x_0\|^2 = \sum_{j=1}^N \|u_{j-1}\|^2 - \sum_{j=1}^N \|y_{j-1}\|^2$$

is satisfied for all $N \geq 1$, all initial values $x_0 \in X$ and all sequences $\{u_j\}$, $\{x_j\}$ and $\{y_j\}$ satisfying

$$\begin{cases} x_{j+1} &= \mathbf{A}x_j + \mathbf{B}u_j, \\ y_{j+1} &= \mathbf{C}x_j + \mathbf{D}u_j, \quad j \geq 0. \end{cases}$$

The DLS ϕ is *conservative* if both ϕ and the dual DLS $\phi^d := \left[\begin{smallmatrix} \mathbf{A}^* & \mathbf{C}^* \\ \mathbf{B}^* & \mathbf{D}^* \end{smallmatrix}\right]$ are energy preserving. Equivalently, ϕ is conservative if and only if $\left[\begin{smallmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{smallmatrix}\right]$ is unitary on $\left[\begin{smallmatrix} X \\ U \end{smallmatrix}\right]$.

Proposition 1: The Cayley transform ϕ_σ of an energy preserving system node S is an energy preserving DLS. Moreover, such ϕ_σ is (discrete time) conservative if and only if S is conservative.

Proof: See [1, Theorem 3.2(v) and Theorem 4.2(iii)]. ∎

II. APPROXIMATION OF THE INPUT/OUTPUT MAPPING

In this section, we describe the discretization (4) of dynamical system (7) and (8) in the language of operator theory.

A. Spaces and transforms

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

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

As usual, the Laplace transform is defined

$$(\mathcal{L}f)(s) = \int_0^\infty e^{-st} f(t)\, dt \quad \text{for all} \quad s \in \mathbb{C}_+, \tag{11}$$

and it maps $L^2(\mathbb{R}_+) \to H^2(\mathbb{C}_+)$ unitarily. The norm of $H^2(\mathbb{D})$ is given by $\|\phi\|_{H^2(\mathbb{D})}^2 = \sum_{j\geq 0} |\phi_j|^2$ 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 (11), 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 $s \in \mathbb{C}_+$. It follows immediately that (6) takes the form of the similarity transformation

$$\mathbf{G} = \mathcal{C}_\sigma^{-1} \mathbf{D}_\sigma \mathcal{C}_\sigma, \tag{12}$$

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 $\left(\mathcal{C}_\sigma^{-1} f\right)(s) := f(\frac{s-\sigma}{s+\sigma})$ for all $s \in \mathbb{C}_+$ and $f : \mathbb{D} \to \mathbb{C}$. In addition, we have

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 since 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 $H^2(\mathbb{C}_+)$. ∎

B. 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 by

$$\begin{aligned}(T_\sigma u)(z) =& \sum_{j\geq 1} u_j^{(h)} z^j \quad \text{where} \\ \frac{u_j^{(h)}}{\sqrt{h}} =& \frac{1}{h}\int_{(j-1)h}^{jh} u(t)\,dt,\end{aligned} \tag{13}$$

with $h = 2/\sigma$; see (3) and (4). Then the adjoint T_σ^* is given by

$$\left(T_\sigma^* \tilde{v}\right)(t) = \frac{1}{\sqrt{h}} \sum_{j\geq 1} v_j \chi_{[(j-1)h,jh]}(t) \tag{14}$$

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 should be noted that the definition of T_σ is not unique and other operators can also be considered.

It is also 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:

$$\begin{aligned}\|T_\sigma^* \tilde{v}\|^2_{L^2(\mathbb{R}_+)} &= \frac{1}{h}\int_0^\infty |\sum_{j\geq 1} v_j \chi_{[(j-1)h,jh]}|^2\,dt \\ &= \frac{1}{h}\int_0^\infty \sum_{j\geq 1} |v_j|^2 \chi_{[(j-1)h,jh]}\,dt = \|\tilde{v}\|^2_{H^2(\mathbb{D})}.\end{aligned} \tag{15}$$

C. Approximation of the Laplace transform.

Let us now use the discrete time trajectories of (4) to approximate the continuous time dynamics in (1).

Let $u \in L^2(\mathbb{R}_+)$ be arbitrary. In the operator notation, the output of the discretized dynamics (4) (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) 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

$$\|T_\sigma^* \mathbf{D}_\sigma T_\sigma u - \mathcal{L}^* \mathbf{G} \mathcal{L} u\|_{L^2([0,T])} \to 0 \tag{16}$$

at some rate as $\sigma \to \infty$. By Proposition 2 and equation (12) 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 (16), 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.

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 (13) we get

$$\begin{aligned}(L_\sigma u)(s) =& \frac{\sqrt{2/\sigma}}{1+s/\sigma} \times \\ & \sum_{j\geq 1} \left(\frac{1}{h}\int_{(j-1)h}^{jh} u(t)\,dt\right) \left(\frac{\sigma - s}{\sigma + s}\right)^j \\ =& \frac{1}{1+s/\sigma} \times \\ & \sum_{j\geq 1} \left(\int_0^\infty \chi_{[(j-1)h,jh]}(t) \left(\frac{\sigma - s}{\sigma + s}\right)^j u(t)\,dt\right) \\ =& \int_0^\infty K_{s,\sigma}(t) u(t)\,dt,\end{aligned} \tag{17}$$

where $\sigma = 2/h$. Now, if j is such that $t \in [(j-1)h, jh]$, then we obtain from the previous

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

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

$$|K_{s,\sigma}(t)| \leq \left(e\sqrt{3}\right)^{|s|t}.$$

The proposition now follows from the Lebesgue dominated convergence theorem, as the integrand in (17) is has a compact support. ∎

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

III. 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.

A. The main result

Before stating the main theorem 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

$$(I_{h,s}u)(t) = \bar{u}_{j,h} + \frac{c_j(h,s)}{h}(t - t_{j-1/2}), \quad t \in I_j, \tag{18}$$

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 $\mathrm{supp}(u) := \{t \in \mathbb{R} : u(t) \neq 0\} \subset [0, T]$.

1) 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 $s \in \overline{\mathbb{C}_+}$.
2) For any such choice of the sequence $\{c_j(h,s)\}_{j\geq 1}$, we have

$$|(L_\sigma u)(s) - (\mathcal{L}u)(s)| \leq \frac{hT^{1/2}|s|}{\pi} \times \left(\|I_{h,s}u - P_h u\|_{L^2([0,T])} + \frac{h}{\pi}|u|_{H^1([0,T])} \right) \tag{19}$$

for all $s \in \overline{\mathbb{C}_+}$.
3) The sequence $\{c_j(h,s)\}_{j\geq 1}$ in claim (1) 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

$$|(L_\sigma u)(s) - (\mathcal{L}u)(s)| \leq \frac{3h^{1/2}|s|}{100}\|u\|_{L^2([0,T])} + \frac{2hT^{1/2}|s|^2}{1000}\|u\|_{L^2([0,T])} + \frac{h^2T^{1/2}|s|}{10}|u|_{H^1([0,T])}. \tag{20}$$

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 $t \in I_j$

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

and it follows that

$$\|I_{h,s}u - P_h u\|^2_{L^2([0,T])} = \frac{h}{12}\sum_{j=1}^{J} c_j(h,s)^2. \tag{21}$$

In claim (1) 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

$$\sum_{j=1}^{J}\bar{u}_{j,h}I_j^{(0)}(h,s) + \sum_{j=1}^{J} c_j(h,s)J_j(h,s) = 0, \tag{22}$$

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

$$I_j^{(0)}(h,s) := \int_{I_j}\left[\frac{1}{1+s/\sigma}\left(\frac{\sigma - s}{\sigma + s}\right)^j - e^{-st}\right] dt \tag{23}$$
$$= \frac{2}{\sigma + s}\left(\frac{\sigma - s}{\sigma + s}\right)^j + \frac{1}{s}\left[e^{-sjh} - e^{-s(j-1)h}\right],$$

and

$$J_j(h,s) := I_j^{(1)}(h,s) - (j - 1/2)h \cdot I_j^{(0)}(h,s) \tag{24}$$
$$= \frac{1}{s^2}\left[e^{-sjh} - e^{-s(j-1)h}\right] + \frac{h}{2s}\left[e^{-sjh} + e^{-s(j-1)h}\right],$$

together with

$$I_j^{(1)}(h,s) := \int_{I_j}\left[\frac{1}{1+s/\sigma}\left(\frac{\sigma - s}{\sigma + s}\right)^j - e^{-st}\right] t\, dt$$
$$= \frac{(2j-1)h}{\sigma + s}\left(\frac{\sigma - s}{\sigma + s}\right)^j + \left(\frac{jh}{s} + \frac{1}{s^2}\right)\left[e^{-sjh} - e^{-s(j-1)h}\right] + \frac{h}{s}e^{-s(j-1)h}. \tag{25}$$

It is clear that (22) 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 (2) is to be treated next. Recalling (17) and (18)

$$(L_\sigma u)(s) - (\mathcal{L}u)(s) = \int_0^T (K_{s,\sigma}(t) - e^{-st})u(t)\, dt$$
$$= \int_0^T (K_{s,\sigma}(t) - e^{-st})(u(t) - (I_{h,s}u)(t))\, dt$$
$$= \sum_{j=1}^{J}\int_{t_{j-1}}^{t_j}(K_{s,\sigma}(t) - e^{-st})(u(t) - \bar{u}_{j,h})\, dt \tag{26}$$
$$- \sum_{j=1}^{J}\frac{c_j(h,s)}{h}\int_{t_{j-1}}^{t_j}(K_{s,\sigma}(t) - e^{-st})(t - t_{j-1/2})\, dt$$
$$= (I) - (II).$$

Let us first give an estimate to the term (II). By the Poincaré inequality (see e.g. [6, Theorem 1.7]) we obtain for all $j = 1, \ldots, J$

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

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 $0 \leq a < b < \infty$,

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

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

$$\|(I - P_h)(K_{s,\sigma} - e^{-s(\cdot)})\|_{L^2(I_j)} \leq \frac{h^{3/2}|s|}{\pi} \tag{27}$$

for all $s \in \overline{\mathbb{C}_+}$. Using (27) we have

$$\begin{aligned}(II) =& \sum_{j=1}^{J} \int_{t_{j-1}}^{t_j} (K_{s,\sigma}(t) - e^{-st}) \times \\ & \frac{c_j(h,s)}{h}(t - t_{j-1/2})\, dt \\ =& \sum_{j=1}^{J} \int_{t_{j-1}}^{t_j} \left((I - P_h)\left(K_{s,\sigma} - e^{-s(\cdot)}\right)\right)(t) \times \\ & \frac{c_j(h,s)}{h}(t - t_{j-1/2})\, dt \\ \leq& \sum_{j=1}^{J} \frac{h^{3/2}|s|}{\pi} \cdot \left[\frac{c_j(h,s)^2}{h^2} \int_{t_{j-1}}^{t_j} (t - t_{j-1/2})^2\, dt\right]^{1/2} \\ \leq& \frac{h^{3/2}|s|}{\pi} J^{1/2} \cdot \|I_{h,s}u - P_h u\|_{L^2([0,T])} \\ =& \frac{hT^{1/2}|s|}{\pi} \|I_{h,s}u - P_h u\|_{L^2([0,T])}\end{aligned} \tag{28}$$

where the Schwarz inequality has been used twice, and the second to last step is by (21). It remains to estimate term (I) in (26). 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 Poincaré, together with (27)

$$\begin{aligned}(I) \leq& \sum_{j=1}^{J} \int_{t_{j-1}}^{t_j} \left((I - P_h)\left(K_{s,\sigma} - e^{-s(\cdot)}\right)\right)(t) \times \\ & (u(t) - \bar{u}_{j,h})\, dt \\ \leq& \sum_{j=1}^{J} \frac{h^{3/2}|s|}{\pi} \cdot \frac{h}{\pi} |u|_{H^1(I_j)} \\ \leq& \frac{h^{5/2}|s|}{\pi^2} \left(\sum_{j=1}^{J} 1\right)^{1/2} \left(\sum_{j=1}^{J} |u|^2_{H^1(I_j)}\right)^{1/2} \\ =& \frac{h^2 T^{1/2}|s|}{\pi^2} |u|_{H^1([0,T])}.\end{aligned} \tag{29}$$

Estimate (19) follows from combining (28) and (29) with (26).

To prove claim (3), we shall minimize $\frac{h}{12}\sum_{j\geq 1} c_j(h,s)^2$ under the constraint (22), see (21) for motivation. We obtain the minimizing sequence

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

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

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

Hence, choosing the operator $I_{h,s}$ in (21) optimally gives

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

since $\|P_h u\|_{L^2([0,T])} = \left(h\sum_{j=1}^{J} \bar{u}_{j,h}^2\right)^{1/2}$. To estimate the required two square sums in (23) and (24) long computations are required. As a final result, we get by Propositions 4 and 5, see [3] for their proofs.

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

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

$$h^{1/2}|s|^2 T^{1/2} \leq \frac{|s|}{3} \cdot |s| T e^{-\frac{2}{3}|s|T} \leq \frac{|s|}{2e},$$

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

B. Some auxiliary results

In this section we give some auxiliary results that were used above. For the proofs of these results, see [3].

Proposition 4: Let $J_j(h,s)$ be defined through (24). 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

$$\|\{J_j(h,s)\}_{j=1}^{J}\|_{\ell^2} \geq \frac{5}{109} T h^2 |s|. \tag{30}$$

Proposition 5: Let $I_j^{(0)}(h,s)$ be defined through (23). 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

$$\begin{aligned}\left\|\left\{I_j^{(0)}(h,s)\right\}_{j=1}^{J}\right\|_{\ell^2} \leq& \frac{1}{2} h^{5/2}|s|^3 T^{3/2} \\ &+ \frac{3}{2} h^{3/2}|s|T^{1/2}.\end{aligned} \tag{31}$$

IV. 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.

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

$$\langle v, L_\sigma u\rangle_{K^2(i\mathbb{R})} \to \langle v, \mathcal{L}u\rangle_{H^2(i\mathbb{R})} \text{ as } \sigma \to \infty \tag{32}$$

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 (32) 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 6: 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.

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 (32) shows that $L_\sigma^* v \to \mathcal{L}^* v$ weakly. Since L_σ is a coisometry by Proposition 2 and (15), 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 6 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

$$\|T_\sigma u\|^2_{H^2(\mathbb{D})} = h\sum_{j\geq 1} |\overline{u}_{j,h}|^2 = \|u\|^2_{L^2(\mathbb{R}_+)}.$$

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 $u \in L^2(\mathbb{R}_+)$

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

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 6 and (32). ∎

Using Theorem 2 we can now show that the output of integration scheme (4) converges to the output of continuous time dynamics (1) 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 II.

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

$$\|T_\sigma^* \mathbf{D}_\sigma T_\sigma u - \mathcal{L}^* \mathbf{G} \mathcal{L} u\|_{L^2(\mathbb{R}_+)} \to 0 \tag{33}$$

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 $\sigma > 0$

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

Now (33) follows by Theorem 2. ∎

V. CONCLUSIONS

The operators L_σ for $\sigma > 0$ have been introduced just before Proposition 3 with the aid of the Cayley transform (6). It is shown in Theorem 2 that the operators L_σ provide an approximation to the 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 (4) for the dynamical equations (1) (or for their infinite-dimensional analogue in [1, Proposition 2.5] as well).

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

$$\mathbf{G}(s) \to 0 \text{ as } |s| \to \infty \tag{34}$$

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 possibly show up in (33) in some form. Unfortunately, (34) 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 (33), one could assume either

1) that the high frequencies are damped by the linear system itself (e.g. by a property like (34)), or
2) 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).

We finally remark that the approximation of the state trajectory $x(\cdot)$ by the discrete trajectories $\{x_j^{(h)}\}_{j\geq 0}$ solving (4) has not been studied here. This will be carried out in a future paper on the state space approximation for conservative systems.

REFERENCES

[1] J. Malinen, O. Staffans, and G. Weiss, When is a linear system conservative? *To appear in: Quarterly of Applied Mathematics*, 2005.

[2] J. Malinen, Conservativity of time-flow invertible and boundary control systems. *Proc. of the Joint 44th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC'05)*, 2005.

[3] V. Havu and J. Malinen, Approximation of the Laplace transform by the Cayley transform. *Helsinki Univ. of Tech. Inst. of Math. Research Reports A480*, 2004.

[4] D.Z. Arov, I.P. Gavrilyuk, A Method for Solving Initial Value Problems for Linear Differential Equations in Hilbert Space Based on the Cayley Transform, *Numer. Funct. Anal. and Optimiz.*, vol. 14, 1993, pp 459-473.

[5] I.P. Gavrilyuk and V.L. Makarov, Explicit and Approximate Solutions of Second-Order Evolution Differential Equations in Hilbert Space, *Numer. Methods Partial Differential Eq.*, vol. 15, 1999, pp 111-131.

[6] B. Dacorogna, *Direct Methods in the Calculus of Variations*, Springer-Verlag, Berlin; 1989.