On discretizing the semigroup of solution operators for linear time invariant - time delay systems

Dimitri Breda * **Stefano Maset** ** **Rossana Vermiglio** *

* *Dipartimento di Matematica e Informatica*
Università degli Studi di Udine
via delle Scienze 206 I-33100 Udine, Italy
{*dimitri.breda,rossana.vermiglio*}*@uniud.it*
** *Dipartimento di Matematica e Informatica*
Università degli Studi di Trieste
via Valerio 12 I-34127 Trieste, Italy
maset@units.it

Abstract: in this paper we give an account of the basic facts to be considered when one attempts to discretize the semigroup of solution operators for Linear Time Invariant - Time Delay Systems (LTI-TDS). Two main approaches are presented, namely *pseudospectral* and *spectral*, based respectively on *classic interpolation* when the state space is $\mathcal{C} = C(-\tau, 0; \mathbb{C})$ and *generalized Fourier projection* when the state space is $\mathcal{X} = \mathbb{C} \times L^2(-\tau, 0; \mathbb{C})$. Full discretization details for constructing the approximation matrices are given. Moreover, concise, yet fundamental, convergence results are discussed, with particular attention to their similarities and differences as well as pros and cons with regards to solution approximation and asymptotic stability detection.

Keywords: delay differential equations, semigroup of solution operators, pseudospectral methods, spectral methods

1. INTRODUCTION

In this work we consider, for simplicity and ease of notation, the prototypical Linear Time Invariant - Time Delay System (LTI-TDS)

$$x'(t) = ax(t) + bx(t-\tau) \tag{1}$$

where $\tau > 0$ and $a, b \in \mathbb{C}$. All the arguments developed in the sequel apply as well to more general cases with matrix coefficients and multiple discrete or distributed delays, the extension concerning only technicalities useless to the treatment proposed in the paper. Instead, generalization to the Linear Time Varying (LTV) case is subject of ongoing works by the authors, although commented whenever it may lead to useful contributions.

In the recent decades, LTI- and LTV-TDS have attracted the attention of diverse scientific communities, automatic control and mathematics above all. A central question from a dynamical point of view is that of asymptotic stability for the zero solution of (1). Despite the great effort, analytical results are rather lacking in generality and, at best, suitable for restricted sub-classes (e.g. single delay or second order systems). As a natural consequence (Hale, 1977, p.109), a number of approximation techniques have been proposed, mostly based on computing the characteristic values (read roots, multipliers, Lyapunov exponents) associated to the system (see e.g. Breda et al. (2005); Butcher et al. (2004); Engelborghs and Roose (2002); Engelborghs et al. (2002); Farmer (1982); Insperger and Stépán (2002); Jarlebring (2008); Verheyden et al. (2008); Vyhlídal and Zítek (2009)).

When investigating on stability (but not only), the *state space* description of (1) is advantageous, and the classic literature resorts to the Banach space of continuous functions $\mathcal{C} := C(-\tau, 0; \mathbb{C})$, Bellen and Zennaro (2003); Diekmann et al. (1995); Hale (1977); Wu (1996). This choice seems to be motivated by the fact that, for rather general selections of the space of initial data, the "smoothing effect" makes the (forward) solution be continuous anyway: "...if some other space than continuous functions is used for initial data, then the solution lies in C...Therefore, for the fundamental theory, the space of initial data does not play a role which is too significant." (Hale, 1977, p.33). Anyway, Hale continues his comment by adding "However, in the applications, it is sometimes convenient to take initial data with fewer or more restrictions." In this sense, an alternative which has been quite studied is represented by the Hilbert product space $\mathcal{X} := \mathbb{C} \times L^2(-\tau, 0; \mathbb{C})$, Bensoussan et al. (1992, 1993); Borisovič and Turbabin (1969); Delfour and Mitter (1972); Hadd et al. (2008); Peichl (1982). This second choice is often justified in the context of quadratic feedback control and linear filtering for retarded systems (Delfour (1977); Hadd et al. (2008); Vinter (1978)), for approximation reasons (Ito and Kappel (1991); Kappel (1986)), or when orthogonality is necessary (Breda (2010)).

Once the proper state space is chosen, say $\mathcal{S}$, the long-time behavior of the evolution can be determined through the knowledge of the spectrum of suitable infinite dimensional maps $\mathcal{S} \to \mathcal{S}$ such as the semigroup of solution operators, its generator, the monodromy operator for periodic prob-

lems and so forth. Part of this manuscript is then devoted to resume the basic features of the semigroup approach for model (1) both in $\mathcal{S}=\mathcal{C}$ and $\mathcal{S}=\mathcal{X}$, focusing on their similarities and differences. We refer the interested readers to Engel and Nagel (1999) for a comprehensive treatment of the theory of general one-parameter linear semigroups.

The reduction of such operators to finite dimension allows to consider standard eigenvalue problems which can be easily solved, hopefully providing accurate estimates for the stability indicators (e.g. the rightmost root or the dominant multiplier). We present two main approaches for discretizing the relevant semigroups, namely the *pseudospectral* one when $\mathcal{S}=\mathcal{C}$ and the *spectral* one when $\mathcal{S}=\mathcal{X}$, comparing computational issues and convergence as well as discussing to what extent they can be applied. The core of the psuedospectral approach is based on *classic polynomial interpolation* and consists in substituting the exact operation to be done on a given (continuous) function with the same operation as applied to the interpolating polynomial. The spectral approach, instead, consists in considering, in a suitable pre-Hilbert space, the *generalized Fourier projection* of the given function to a finite degree. The books Trefethen (2000) and Canuto et al. (2007) may serve as a guide for the above methodologies in a context much broader than TDS.

The paper is structured as follows. Section 2 resumes the semigroup theory for (1). After some preliminaries discussed in Section 3, Section 4 deals with the discretization in $\mathcal{C}$. The same role is played by Sections 5 and 6, respectively, for $\mathcal{X}$. Section 7 collects some results and discussion on convergence.

As a final introductory comment, we recall that elements in L^2 (or similar Lebesgue spaces) have to be intended as *equivalence classes* of functions rather than functions themselves, Davis (1975). As a consequence, the notion of "value of a function at a given point" is meaningless, contrary to the continuous case. Moreover, in $\mathcal{C}$ we use the standard maximum norm $\|\varphi\|_{\mathcal{C}} := \max_{\theta\in[-\tau,0]} |\varphi(\theta)|$ while in $\mathcal{X}$ we use the norm induced by the inner product $< (u,\varphi),(v,\psi) >_{\mathcal{X}} := v^H u + \int_{-\tau}^{0} \psi^H(\theta)\varphi(\theta)d\theta$, i.e. $\|(u,\varphi)\|^2_{\mathcal{X}} := |u|^2 + \|\varphi\|^2_{L^2(-\tau,0;\mathbb{C})}$.

2. SEMIGROUP THEORY IN $\mathcal{C}$ AND $\mathcal{X}$

For a well-posed Initial Value Problem (IVP) for (1) with a given, fixed $r \geq 0$, the classic choice of initial data in $\mathcal{S}=\mathcal{C}$ leads to

$$\begin{cases} x'(t) = ax(t) + bx(t-\tau), & t \in [0,r] \\ x(\theta) = \varphi(\theta), & \theta \in [-\tau,0] \end{cases} \tag{2}$$

for a given $\varphi \in \mathcal{C}$. Diversely, if $\mathcal{S}=\mathcal{X}$ is chosen, then the IVP reads

$$\begin{cases} x'(t) = ax(t) + bx(t-\tau), & \text{for a.a. } t \in [0,r] \\ x(0) = u \\ x(\theta) = \varphi(\theta), & \text{for a.a. } \theta \in [-\tau,0) \end{cases} \tag{3}$$

for a given $(u,\varphi) \in \mathcal{X}$. Both IVPs admit a unique solution which continuously depends on the initial data, see e.g. Hale (1977) for (2) and Breda (2010) for (3). This allows to introduce the *solution operator* as the one-parameter linear and bounded operator $T(r) : \mathcal{S} \to \mathcal{S}$ given by

$$T(r)\hat{x}(0) = \hat{x}(r) \tag{4}$$

where $\mathcal{S}$ is either $\mathcal{C}$ or $\mathcal{X}$, and $\hat{x}(r)$ denotes the *state* of the system at time $r \geq 0$. In particular, we have

$$\hat{x}(r) := x_r \in \mathcal{C} \;\text{ with }\; \hat{x}(0) = \varphi \in \mathcal{C} \tag{5}$$

if $\mathcal{S}=\mathcal{C}$ or

$$\hat{x}(r) := (x(r), x_r) \in \mathcal{X} \;\text{ with }\; \hat{x}(0) = (u,\varphi) \in \mathcal{X}$$

if $\mathcal{S}=\mathcal{X}$, where we adopt the standard Hale-Krasovskii notation for the function $[-\tau,0] \ni \theta \mapsto x_r(\theta) \in \mathbb{C}$ defined as $x_r(\theta) := x(r+\theta)$, Hale (1977); Krasovskii (1959). In $\mathcal{X}$ this latter holds for *a.a.* $\theta \in [-\tau,0]$.

For completeness, the corresponding *infinitesimal generator* has action and domain dependent on the choice of $\mathcal{S}$. In particular, the linear unbounded operator $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subseteq \mathcal{S} \to \mathcal{S}$ is

$$\begin{cases} \mathcal{D}(\mathcal{A}) = \left\{\psi \in \mathcal{C}^1 : \psi'(0) = a\psi(0) + b\psi(-\tau)\right\} \\ \mathcal{A}\psi = \psi', \end{cases} \tag{6}$$

when $\mathcal{S}=\mathcal{C}$, while for $\mathcal{S}=\mathcal{X}$

$$\begin{cases} \mathcal{D}(\mathcal{A}) = \left\{(v,\psi) \in \mathcal{X}^1 : \psi(0) = v\right\} \\ \mathcal{A}(v,\psi) = (av + b\psi(-\tau), \psi')\,. \end{cases} \tag{7}$$

Above we used $\mathcal{C}^1 := C^1(-\tau,0;\mathbb{C})$ and $\mathcal{X}^1 := \mathbb{C} \times H^1(-\tau,0;\mathbb{C})$ where H^1 is the Sobolev space of L^2 functions with first (weak) derivative in L^2.

Remark 1. Let us underline that as far as the LTV case is considered, (4) can be extended to

$$T(r,s)\hat{x}(s) = \hat{x}(r)$$

where s is the initial time for the relevant IVP (which inevitably matters), while (6) and (7) can be extended to

$$\begin{cases} \mathcal{D}(\mathcal{A}(t)) = \left\{\psi \in \mathcal{C}^1 : \psi'(0) = a(t)\psi(0) + b(t)\psi(-\tau)\right\} \\ \mathcal{A}(t)\psi = \psi' \end{cases} \tag{8}$$

for $\mathcal{S}=\mathcal{C}$ and

$$\begin{cases} \mathcal{D}(\mathcal{A}(t)) = \left\{(v,\psi) \in \mathcal{X}^1 : \psi(0) = v\right\} \\ \mathcal{A}(t)(v,\psi) = (a(t)v + b(t)\psi(-\tau), \psi') \end{cases} \tag{9}$$

for $\mathcal{S}=\mathcal{X}$. At this point it is necessary to stress that the time dependency is confined to the *domain* for (8) and to the *action* for (9). This makes (8) difficult to treat, see e.g. (Chicone and Latushkin, 1999, p.59), (Diekmann et al., 1995, p.341), (Hadd et al., 2008, p.4), practically leaving (9) as the only possible extension of the concept of generator to the time-varying case, Breda (2010).

3. PRELIMINARIES AND NOTATION IN $\mathcal{C}$

Depending on the role of the various mathematical objects, in general we use normal case for operators and functions (infinite dimension), bold case for matrices and vectors (finite dimension).

Set $\mathcal{C}^- := \mathcal{C} = C(-\tau,0;\mathbb{C})$, $\mathcal{C}^+ := C(0,r;\mathbb{C})$ and $\mathcal{C}^\pm := C(-\tau,r;\mathbb{C})$. Whenever required, functions are denoted as $f^- \in \mathcal{C}^-$, $f^+ \in \mathcal{C}^+$ and $f^\pm \in \mathcal{C}^\pm$, but simply $f \in \mathcal{C}$ $(= \mathcal{C}^-)$ when reference to the state space has to be stressed. Also, $f^\pm \in \mathcal{C}^\pm$ is tacitly intended as divided into $f^- := f^\pm_{|[-\tau,0]}$ and $f^+ := f^\pm_{|[0,r]}$. The same notation holds for spaces (and functions) other than $\mathcal{C}$. In particular, we denote Π^-_N and Π^+_N the spaces of polynomials of degree at most N, respectively on $[-\tau,0]$ and $[0,r]$.

Rewrite the IVP (2) as

$$\begin{cases} x'(t) = (Gx)(t), \ t \in [0,r], \\ x(\theta) = \varphi(\theta), \qquad \theta \in [-\tau, 0] \end{cases} \tag{10}$$

where the operator $G : \mathcal{C}^{\pm} \to \mathcal{C}^{+}$ is defined as

$$(Gx)(t) = ax(t) + bx(t-\tau) \tag{11}$$

for $t \in [0,r]$. For a given positive integer N, consider the grid of distinct nodes $\Omega_N^- := \{-\tau = \theta_N^- < \cdots < \theta_0^- = 0\}$ in $[-\tau, 0]$ and set $\mathcal{C}_N^- := \mathbb{C}^{N+1}$ as the discrete counterpart of $\mathcal{C}^-$, i.e. a function $f^- \in \mathcal{C}^-$ is discretized by the vector $\boldsymbol{f}_N^- = \mathcal{R}_N^- f^- = (f^-(\theta_0^-), \ldots, f^-(\theta_N^-))^T \in \mathcal{C}_N^-$, $\mathcal{R}_N^- : \mathcal{C} \to \mathcal{C}_N^-$ the restriction operator. Correspondingly, let $f_N^- = \mathcal{P}_N^- \boldsymbol{f}_N^- \in \Pi_N^-$, $\mathcal{P}_N^- : \mathcal{C}_N^- \to \Pi_N^- \subset \mathcal{C}^-$ the prolongation operator, be the polynomial of degree at most N interpolating the values $\boldsymbol{f}_N^-$ at the nodes Ω_N^-, i.e.

$$f_N^-(\theta) := \sum_{j=0}^{N} \ell_j^-(\theta) f^-(\theta_j^-), \ \theta \in [-\tau, 0],$$

with

$$\ell_j^- := \prod_{\substack{k=0 \\ k \neq j}}^{N} \frac{\theta - \theta_k^-}{\theta_j^- - \theta_k^-}, \ j = 0, \ldots, N,$$

the Lagrange basis polynomials relevant to the nodes Ω_N^-. Observe that $\mathcal{R}_N^- \mathcal{P}_N^- = \boldsymbol{I}_N^- : \mathcal{C}_N^- \to \mathcal{C}_N^-$, the identity in $\mathcal{C}_N^-$, while $\mathcal{P}_N^- \mathcal{R}_N^- = \mathcal{L}_N^- : \mathcal{C}^- \to \Pi_N^- \subset \mathcal{C}^-$, the Lagrange interpolation operator on Ω_N^-.

Similarly, let $\Omega_N^+ := \{0 < \theta_1^+ < \cdots < \theta_N^+ < r\}$ in $(0,r)$ together with the auxiliary node $\theta_0^+ := 0$ $(= \theta_0^-)$. Then, for a function $f^+ \in \mathcal{C}^+$, set $\boldsymbol{f}_N^+ = \mathcal{R}_{N,0}^+ f^+ = (f^+(\theta_0^+), \ldots, f^+(\theta_N^+))^T \in \mathcal{C}_N^+$ and $f_N^+ = \mathcal{P}_{N,0}^+ \boldsymbol{f}_N^+ \in \Pi_N^+$, with the analogous meaning for $\mathcal{R}_{N,0}^+ : \mathcal{C} \to \mathcal{C}_N^+$ and $\mathcal{P}_{N,0}^+ : \mathcal{C}_N^+ \to \Pi_N^+ \subset \mathcal{C}^+$ w.r.t. the nodes $\{\theta_0^+\} \cup \Omega_N^+$. When only the nodes in Ω_N^+ are considered, the suffix 0 is drop and $\mathcal{L}_N^* = \mathcal{P}_N^+ \mathcal{R}_N^+$ is the relevant Lagrange interpolation operator.

Finally, for a given state (according to (5)) $\hat{x}(t) \in \mathcal{C}$, $t \in [0,r]$, we consider $\hat{\boldsymbol{x}}_N(t) = \hat{\mathcal{R}}_N^- \hat{x}(t) \in \mathcal{C}_N^-$ as the vector of its values at the nodes Ω_N^- and $\hat{x}_N(t) = \hat{\mathcal{P}}_N^- \hat{\boldsymbol{x}}_N(t) \in \Pi_N^- \subset \mathcal{C}$ as the relevant interpolating polynomial.

4. SEMIGROUP DISCRETIZATION IN $\mathcal{C}$

We aim at finding a finite dimensional approximation $\boldsymbol{T}_N(r)$ of the solution operator $T(r)$ in (4) for $\mathcal{S} = \mathcal{C}$. We basically use *collocation* to advance from $[-\tau, 0]$ to $[0,r]$ together with *classic polynomial interpolation* for discrete representation as introduced in Section 3. Collocation approaches have been proposed in Breda (2004, 2006). Alternatives have been considered also in Engelborghs and Roose (2002); Insperger and Stépán (2002); Verheyden et al. (2008).

According to the notation set in Section 3, we first construct matrices $\boldsymbol{U}_N^- : \mathcal{C}_N^- \to \mathcal{C}_N$ and $\boldsymbol{U}_N^+ : \mathcal{C}_N^+ \to \mathcal{C}_N$ relating the discretized initial function to the discretized collocation polynomial:

$$\boldsymbol{U}_N^+ \boldsymbol{p}_N^+ = \boldsymbol{U}_N^- \boldsymbol{\varphi}_N^- \tag{12}$$

where $p_N \in \Pi_N^{\pm}$ is divided into $p_N^- = \varphi_N$ and p_N^+ determined by collocation of (10):

$$\begin{cases} (p_N^+)'(\theta_i^+) = (Gp_N)(\theta_i^+), \ i = 1, \ldots, N, \\ p_N^+(0) = \varphi_N(0). \end{cases} \tag{13}$$

It is not conceptually difficult (although rather technical) to check that the above matrices have entries, respectively,

$$[\boldsymbol{U}_N^+]_{ij} := \begin{cases} \ell_j^+(0), & i = 0 \\ ((\ell_j^+)' - a\ell_j^+)(\theta_i^+), & i = 1, \ldots, N^+ \\ ((\ell_j^+)' - a\ell_j^+)(\theta_i^+) & -b\ell_j^+(\theta_i^+ - \tau), \\ & i = N^+ + 1, \ldots, N \end{cases}$$

and

$$[\boldsymbol{U}_N^-]_{ij} := \begin{cases} \ell_j^-(0), & i = 0 \\ b\ell_j^-(\theta_i^+ - \tau), & i = 1, \ldots, N^+ \\ 0, & i = N^+ + 1, \ldots, N \end{cases}$$

for all $j = 0, \ldots, N$, where

$$N^+ = N^+(r,\tau) := \max_{j=1,\ldots,N} \{\theta_j^+ - \tau \leq 0\}.$$

Second, and independently of the model coefficients a and b, we construct matrices $\boldsymbol{V}_N^+ : \mathcal{C}_N^+ \to \mathcal{C}_N$ and $\boldsymbol{V}_N^- : \mathcal{C}_N^- \to \mathcal{C}_N$ such that

$$\hat{\boldsymbol{x}}_N(r) = \boldsymbol{V}_N^+ \boldsymbol{p}_N^+ + \boldsymbol{V}_N^- \boldsymbol{\varphi}_N^- \tag{14}$$

by restriction of p_N to $[r-\tau, r]$ when $r \geq \tau$, respectively prolongation by φ_N when $r < \tau$. In particular, it is sufficient to define the above matrices with entries, respectively,

$$[\boldsymbol{V}_N^+]_{ij} := \begin{cases} \ell_j^+(r + \theta_i^-), & i = 0, \ldots, N^- \\ 0, & i = N^- + 1, \ldots, N \end{cases}$$

and

$$[\boldsymbol{V}_N^-]_{ij} := \begin{cases} 0, & i = 0, \ldots, N^- \\ \ell_j^-(r + \theta_i^-), & i = N^- + 1, \ldots, N \end{cases}$$

for all $j = 0, \ldots, N$, where

$$N^- = N^-(r,\tau) := \max_{j=0,\ldots,N} \{r + \theta_j^- \geq 0\},$$

with the convention that $\boldsymbol{V}_N^+$ is full and $\boldsymbol{V}_N^-$ is empty when $N^- = N$, i.e. for $r \geq \tau$.

Eventually, by setting $\hat{\boldsymbol{x}}_N(0) = \boldsymbol{\varphi}_N^-$, it follows from (12) and (14) that

$$\hat{\boldsymbol{x}}_N(r) = \boldsymbol{T}_N(r) \hat{\boldsymbol{x}}_N(0) \tag{15}$$

is the sought discrete approximation of (4) with $\boldsymbol{T}_N(r) : \mathcal{C}_N \to \mathcal{C}_N$ given by

$$\boldsymbol{T}_N(r) = \boldsymbol{V}_N^+ (\boldsymbol{U}_N^+)^{-1} \boldsymbol{U}_N^- + \boldsymbol{V}_N^-$$

(standard approximation arguments ensure that $\boldsymbol{U}_N^+$ is invertible for sufficiently large N).

5. PRELIMINARIES AND NOTATION IN $\mathcal{X}$

Similarly to what done in Section 3, rewrite the IVP (2) as

$$\begin{cases} x'(t) = (Gx)(t), \ \text{for a.a. } t \in [0,r], \\ x(0) = u \\ x(\theta) = \varphi(\theta), \qquad \text{for a.a. } \theta \in [-\tau, 0) \end{cases} \tag{16}$$

where the operator $G : L^{\pm} \to L^+$ is defined formally as in (11) for *a.a.* $t \in [0,r]$ and the spaces L^-, L^+ and $L^{\pm}$ intended as for the continuous but case from $L := L^2$.

For a given positive integer N, let $\{\phi_i^-\}_{i=0}^{\infty}$ be a system of orthogonal algebraic polynomials spanning L^- such

that $\phi_N^- \in \Pi_N^-$ has zeros $-\tau < \theta_N^- < \cdots < \theta_1^- < 0$ and set $L_N^- := \mathbb{C}^{N+1}$ as the discrete counterpart of L^-, i.e. a function $f^- \in L^-$ with Fourier coefficients $\{f_i^-\}_{i=0}^\infty$ is discretized by the vector $\boldsymbol{f}_N^- = \mathcal{R}_N^- f^- = (f_0^-, \ldots, f_N^-)^T \in L_N^-$, $\mathcal{R}_N^- : L \to L_N^-$ the restriction operator. Correspondingly, let $f_N^- = \mathcal{P}_N^- \boldsymbol{f}_N^- \in \Pi_N^-$ be the projection polynomial of degree at most N for f^-, i.e.

$$f_N^- := \sum_{j=0}^{N} f_j^- \phi_j^-,$$

$\mathcal{P}_N^- : L_N^- \to \Pi_N^- \subset L^-$ the prolongation operator. Observe that $\mathcal{R}_N^- \mathcal{P}_N^- = \boldsymbol{I}_N^- : L_N^- \to L_N^-$, the identity in $\mathcal{L}_N^-$, while $\mathcal{P}_N^- \mathcal{R}_N^- = \mathcal{L}_N^- : L^- \to \Pi_N^- \subset L^-$, the Fourier projection operator on Ω_N^-.

Similarly, let $\{\phi_i^+\}_{i=0}^\infty$ be a system of orthogonal algebraic polynomials spanning L^+ such that $\phi_N^+ \in \Pi_N^+$ has zeros $0 < \theta_0^+ < \cdots < \theta_N^+ < r$. Then, for a function $f^+ \in L^+$, set $\boldsymbol{f}_N^+ = \mathcal{R}_N^+ f^+ = (f_1^+, \ldots, f_N^+)^T \in L_N^+$ denotes the vector of its first $N+1$ Fourier coefficients and $f_N^+ = \mathcal{P}_N^+ \boldsymbol{f}_N^+ \in \Pi_N^+$ the relevant projection polynomial, with the analogous meaning for $\mathcal{R}_N^+ : L \to L_N^+$ and $\mathcal{P}_N^+ : L_N^+ \to \Pi_N^+ \subset L^+$, and also for $\boldsymbol{I}_N^+$ and $\mathcal{L}_N^+$.

Finally, for a given state $(v, f) =: \hat{x} \in \mathcal{X}$, we consider $\hat{\boldsymbol{x}}_N = \hat{\mathcal{R}}_N \hat{x} = (v, f_0^-, \ldots, f_N^-)^T \in \mathcal{X}_N := \mathbb{C} \times L_N = \mathbb{C}^{N+2}$, $\hat{\mathcal{R}}_N : \mathcal{X} \to \mathcal{X}_N$ the restriction operator, and $\hat{x}_N = \hat{\mathcal{P}}_N \hat{\boldsymbol{x}}_N = (v, f_N^-) \in \mathcal{X}$ as the relevant projected state, $\hat{\mathcal{P}}_N : \mathcal{X}_N \to \mathcal{X}$ the prolongation operator. We will consider also $\mathcal{X}_N^+ := \mathbb{C} \times L_N^+$ and $\mathcal{X}_N^- := \mathbb{C} \times L_N^-$, all isomorphic to $\mathbb{C}^{N+2}$ since L_N, L_N^- and L_N^+ are all isomorphic to $\mathbb{C}^{N+1}$ but, again, we reserve to distinguish the notation for the relevant meaning. Moreover, the indexing of a vector in $\mathbb{C}^{N+2}$ will be $-1, 0, \ldots, N$ to take into account for the presence of the scalar element (index -1) and the discrete functional element (indices $0, \ldots, N$).

6. SEMIGROUP DISCRETIZATION IN $\mathcal{X}$

We aim at finding a finite dimensional approximation $\boldsymbol{T}_N(r)$ of the solution operator $T(r)$ in (4) for $\mathcal{S} = \mathcal{X}$. We basically use *collocation* again together with *generalized Fourier projection* as introduced in Section 5. The following approach is new.

According to the notation set in Section 5, we construct matrices $\boldsymbol{U}_N^- : \mathcal{X}_N^- \to \mathcal{X}_N$ and $\boldsymbol{U}_N^+ : \mathcal{X}_N^+ \to \mathcal{X}_N$ such that

$$\boldsymbol{U}_N^+(p_N^+(r), \boldsymbol{p}_N^+) = \boldsymbol{U}_N^-(u, \boldsymbol{\varphi}_N^-) \tag{17}$$

where $p_N \in \Pi_N^\pm$ is divided into $p_N^- = \varphi_N$ and p_N^+ determined by collocation of (16):

$$\begin{cases} (p_N^+)'(\theta_i^+) = (Gp_N)(\theta_i^+), \ i = 1, \ldots, N, \\ p_N^+(0) = u. \end{cases} \tag{18}$$

Again, it is not conceptually difficult (although rather technical) to check that the above matrices have entries, respectively,

$$[\boldsymbol{U}_N^+]_{ij} := \begin{cases} -\phi_j^+(r), & i = -1 \\ \phi_j^+(0), & i = 0 \\ ((\phi_j^+)' - a\phi_j^+)(\theta_i^+), & i = 1, \ldots, N^+ \\ ((\phi_j^+)' - a\phi_j^+)(\theta_i^+) & -b\phi_j^+(\theta_i^+ - \tau), \\ & i = N^+ + 1, \ldots, N \end{cases}$$

for all $j = 0, \ldots, N$ plus the first column ($j = -1$) as $(1, 0, \ldots, 0) \in \mathbb{C}^{N+2}$, and

$$[\boldsymbol{U}_N^-]_{ij} := \begin{cases} 0, & i = -1, 0, N^+ + 1, \ldots, N \\ b\phi_j^-(\theta_i^+ - \tau), & i = 1, \ldots, N^+ \end{cases}$$

for all $j = 0, \ldots, N$, plus the first column ($j = -1$) as $(0, 1, 0, \ldots, 0) \in \mathbb{C}^{N+2}$, where

$$N^+ = N^+(r, \tau) := \max_{j=1,\ldots,N} \{\theta_j^+ - \tau \le 0\}.$$

Second, and independently of the model coefficients a and b, we construct matrices $\boldsymbol{V}_N, \boldsymbol{V}_N^- : \mathcal{X}_N^- \to \mathcal{X}_N$ and $\boldsymbol{V}_N^+ : \mathcal{X}_N^+ \to \mathcal{X}_N$ such that

$$\boldsymbol{V}_N \hat{\boldsymbol{x}}_N(r) = \boldsymbol{V}_N^+(p_N^+(r), \boldsymbol{p}_N^+) + \boldsymbol{V}_N^-(u, \boldsymbol{\varphi}_N^-) \tag{19}$$

by restriction of p_N to $[r - \tau, r]$ when $r \ge \tau$, respectively prolongation by φ_N when $r < \tau$. In particular, it is sufficient to define the above matrices with entries, respectively,

$$[\boldsymbol{V}_N]_{ij} := \begin{cases} \phi_j^-(0), & i = 0 \\ \phi_j^-(\theta_i^-), & i = 1, \ldots, N \end{cases}$$

for all $j = 0, \ldots, N$, plus $[\boldsymbol{V}_N]_{-1,-1} = 1$ and 0 elsewhere,

$$[\boldsymbol{V}_N^+]_{ij} := \begin{cases} \phi_j^+(r), & i = 0 \\ \phi_j^+(r + \theta_i^-), & i = 1, \ldots, N^- \\ 0, & i = N^- + 1, \ldots, N \end{cases}$$

for all $j = 0, \ldots, N$, plus $[\boldsymbol{V}_N^+]_{-1,-1} = 1$ and 0 elsewhere, and

$$[\boldsymbol{V}_N^-]_{ij} := \begin{cases} 0, & i = 0, \ldots, N^- \\ \phi_j^-(r + \theta_i^-), & i = N^- + 1, \ldots, N \end{cases}$$

for all $j = 0, \ldots, N$ and 0 elsewhere, where

$$N^- = N^-(r, \tau) := \max_{j=0,\ldots,N} \{r + \theta_j^- \ge 0\},$$

with the convention that $\boldsymbol{V}_N^+$ is full and $\boldsymbol{V}_N^-$ is empty when $N^- = N$, i.e. for $r \ge \tau$.

Eventually, by setting $\hat{\boldsymbol{x}}_N(0) = (u, \boldsymbol{\varphi}_N^-)$, it follows from (17) and (19), that

$$\hat{\boldsymbol{x}}_N(r) = \boldsymbol{T}_N(r) \hat{\boldsymbol{x}}_N(0) \tag{20}$$

is the sought discrete approximation of (4) with $\boldsymbol{T}_N(r) : \mathcal{X}_N \to \mathcal{X}_N$ given by

$$\boldsymbol{T}_N(r) = (\boldsymbol{V}_N)^{-1} [\boldsymbol{V}_N^+ (\boldsymbol{U}_N^+)^{-1} \boldsymbol{U}_N^- + \boldsymbol{V}_N^-]$$

(standard approximation arguments ensure that $\boldsymbol{V}_N^+$ and $\boldsymbol{U}_N^+$ are invertible for sufficiently large N).

7. CONVERGENCE ANALYSIS

The solution operator $T(r)$ in (4) is an infinite dimensional map, say $T(r) : \mathcal{S} \to \mathcal{S}$ with either $\mathcal{S} = \mathcal{C}$ or $\mathcal{S} = \mathcal{X}$, contrary to its matrix discretization $\boldsymbol{T}_N(r) : \mathcal{S}_N \to \mathcal{S}_N$ with either $\mathcal{S}_N = \mathcal{C}_N$ in (15) or $\mathcal{S}_N = \mathcal{X}_N$ in (20). For comparison, it is therefore necessary to introduce an intermediate infinite dimensional, but finite-rank, map $T_N(r) : \mathcal{S} \to \mathcal{S}$ as detailed later on.

This Section is then devoted to provide suitable error bounds for the remainder $T(r) - T_N(r)$ in the state space $\mathcal{S}$. Such errors will be measured in a *pointwise* sense in general (i.e. as applied to *a given* function in $\mathcal{S}$), reserving to comment on the convergence in *norm* when possible (i.e. as applied to *all* functions in $\mathcal{S}$). It is worthy to warn the reader that it is out of the scope of the present manuscript

to develop a systematic and complete error analysis for the exact stability indicators (roots, multipliers, exponents, etc.), i.e. spectral elements of $T(r)$ as approximated by the eigenvalues of $\boldsymbol{T}_N(r)$. However, let us stress that, according to the theory developed in Chatelin (1983), the pointwise convergence, i.e. $\|(T(r) - T_N(r))\hat{x}\|_{\mathcal{S}} \to 0$ as $N \to \infty$ for all $\hat{x} \in \mathcal{S}$, is a *mandatory* requirement to the goal, with all the consequences that the forthcoming analyses bring. This (and other) aspect(s) are planned to be fully addressed by the authors in a forthcoming work, Breda et al. (2010).

7.1 Convergence $T_N(r) \to T(r)$ in $\mathcal{C}$

Set

$$T_N(r)\varphi = (q_N)_r \tag{21}$$

where q_N is the collocation solution for the exact problem (10), i.e. given by (13) with φ instead of φ_N: note that in general p_N and q_N are different. It is not difficult to see that

$$\boldsymbol{T}_N(r) = \hat{\mathcal{R}}_N^- T_N(r) \hat{\mathcal{P}}_N^- \tag{22}$$

holds, giving a link for the spectral properties of $\boldsymbol{T}_N(r)$ and $T_N(r)$ as detailed at the end.

Assumption 2. Assume the nodes in Ω_N^- to be Chebyshev *extrema* in $[-\tau, 0]$, i.e. $\theta_i^- = \tau(\cos(i\pi/N) - 1)/2$, $i = 0, \dots, N$, and the nodes in Ω_N^+ to be Chebyshev *zeros* in $(0, r)$, i.e. $\theta_i^+ = r(1 - \cos((2i-1)\pi/2N))/2$, $i = 1, \dots, N$.

Theorem 3. Let x solve (10) with $r \geq 0$ and $\varphi \in \mathcal{C}$ and set $T(r)$ as in (4). Then, under Assumption 2 and for sufficiently large N, $T_N(r)$ in (21) is uniquely defined and

$$\|(T(r) - T_N(r))\varphi\|_{\mathcal{C}} \leq K \|(I - \mathcal{L}_N^*)Gx\|_{\mathcal{C}^+}$$

holds with $K = K(r, \tau, |a|, |b|)$ constant independent of N and φ.

Sketch of proof. The proof is based on considering the two functional equations in $\mathcal{C}^{\pm}$

$$\begin{cases} x = u_\varphi + VGx \\ q_N = u_\varphi + V\mathcal{L}_N^* G q_N \end{cases}$$

respectively for (10) and its collocation problem, both with initial function φ, with $u_\varphi \in \mathcal{C}^{\pm}$ as $u_\varphi(t) = \varphi(t)$ for $t \in [-\tau, 0]$ and $u_\varphi(t) = \varphi(0)$ elsewhere and $V : \mathcal{C}^+ \to \mathcal{C}^{\pm}$ as the integral operator $(Vx)(t) := \int_0^t x(\sigma)d\sigma$ for $t \in [0, r]$ and 0 elsewhere. The thesis follows by applying to the error function $e_N := x - q_N = V(Gx - \mathcal{L}_N^* G q_N)$ standard approximation results such as those used in the proofs of Breda et al. (2005), e.g. Natanson and Jackson Theorems, Natanson (1965); Davis (1975), but also Krylov (1956). ■

Let us observe that the above result does not ensure neither convergence in norm, i.e. for $\|T(r) - T_N(r)\|_{\mathcal{C}}$ nor pointwise, i.e. for $\|(T(r) - T_N(r))\varphi\|_{\mathcal{C}}$ for all $\varphi \in \mathcal{C}$, basically because *there is no choice of nodes making classic polynomial interpolation converge in all* $\mathcal{C}$ (Faber's Theorem, Davis (1975); Faber (1914)). Indeed, for pointwise convergence, more regularity of the initial function φ is required as stated in the following.

Corollary 4. If φ in Theorem 3 is *absolutely continuous*, then

$$\lim_{N \to \infty} \|(T(r) - T_N(r))\varphi\|_{\mathcal{C}} = 0.$$

This explains why the collocation approach is not used, in general, for approximating the solution (continuity of φ is not enough), while being extremely efficient for approximating the characteristic values: in this latter case the underlying solution (= eigenfunction) is analytic Hale (1977), leading to spectral convergence (see Breda et al. (2005); Breda (2006); Trefethen (2000)) since the error is controlled by the interpolation error over Ω_N^+ for the derivative $x' = Gx$ of the exact solution, analytic as well.

Finally, for a fixed N it can be proven that the matrix $\boldsymbol{T}_N(r)$ and the operator $\mathcal{P}_N^- \boldsymbol{T}_N(r) \mathcal{R}_N^- = \mathcal{L}_N^- T_N(r) \mathcal{L}_N^-$ have the same nonzero eigenvalues (and relevant multiplicities). Moreover, when $r \geq \tau$ we have $\mathcal{L}_N^- T_N(r) \mathcal{L}_N^- = T_N(r)\mathcal{L}_N^-$ and $\|(T_N(r)\mathcal{L}_N^- - T_N(r))\varphi\|_{\mathcal{C}} \to 0$ as $N \to \infty$ whenever φ is absolutely continuous. This latter is a consequence of Corollary 4 together with the Banach-Steinhaus Theorem. By resuming the backbone of the procedure, asymptotically with N, the spectrum of $T(r)$ is approximated by that of $T_N(r)$ which can be effectively computed by that of the matrix $\boldsymbol{T}_N(r)$ in (15) for any fixed N.

7.2 Convergence $T_N(r) \to T(r)$ in $\mathcal{X}$

The analysis is much similar to that for $\mathcal{S} = \mathcal{C}$, but some subtle details must be clarified. First of all and following (21), an operator

$$T_N(r)(u, \varphi) = (p_N(r), (p_N)_r) \tag{23}$$

cannot be defined, precisely it is not defined in all the state space $\mathcal{X}$ since a collocation scheme requires φ to be continuous. Nevertheless, (23) holds in $\mathcal{D}(\mathcal{A})$, which is dense in $\mathcal{X}$, Breda (2010) and here (22) is valid. Now, functions in $\mathcal{D}(\mathcal{A})$ are actually absolutely continuous by the Sobolev Embedding Theorem and so Theorem 3 holds applied with the necessary modifications as follows.

Assumption 5. Assume the orthogonal systems $\{\phi_i^-\}_{i=0}^{\infty}$ and $\{\phi_i^+\}_{i=0}^{\infty}$ to be of Gauss-Legendre type.

Theorem 6. Let x solve (16) with $r \geq 0$ and $(u, \varphi) \in \mathcal{D}(\mathcal{A})$ and set $T(r)$ as in (4). Then, under Assumption 5 and for sufficiently large N, $T_N(r)$ in (23) is uniquely defined. Moreover,

$$\lim_{N \to \infty} \|(T(r) - T_N(r))(u, \varphi)\|_{\mathcal{X}} = 0.$$

Sketch of proof. It follows exactly the same line of the proof sketched for Theorem 3, completed by Corollary 4. ■

Now, a similar comment about norm-convergence done in Section 7.1 holds, as well as about spectral convergence for the characteristic values and their computation through the matrix $\boldsymbol{T}_N(r)$ in (20).

8. CONCLUSIONS

We presented two discretization schemes for the semigroup of solution operators of LTI-TDS, namely pseudospectral in $\mathcal{S} = \mathcal{C}$ and spectral in $\mathcal{S} = \mathcal{X}$. The aim, beyond construction, was to stress some critical facts about the theoretical (i.e. not experimental) convergence of the approximations, mostly concerned with the final target of using these methods (in either one state space or the other) for stability purposes through the determination of

suitable spectral bounds via standard matrix eigenvalue problems. Future works of the authors will be aimed at addressing all the relevant (and numerous) mathematical details, out of the scope of the present manuscript.

REFERENCES

Bellen, A. and Zennaro, M. (2003). *Numerical methods for delay differential equations*. Numerical Mathemathics and Scientifing Computing series. Oxford University Press.

Bensoussan, A., Da Prato, G., Delfour, M.C., and Mitter, S.K. (1992, 1993). *Representation and control of infinite dimensional systems*, volume I and II. Birkhäuser.

Borisovič, J.G. and Turbabin, A.S. (1969). On the Cauchy problem for linear nonhomogeneous differential equations with retarded arguments. *Dokl. Akad. Nauk SSSR*, 185(4), 741–744. English transl. Soviet Math. Dokl., 10(2):401-405, 1969.

Breda, D. (2004). *Numerical computation of characteristic roots for delay differential equations*. Ph.D. thesis, PhD in Computational Mathematics, Università di Padova.

Breda, D. (2006). Solution operator approximation for characteristic roots of delay differential equations. *Appl. Numer. Math.*, 56(3-4), 305–317.

Breda, D. (2010). Nonautonomous delay differential equations in Hilbert spaces and Lyapunov exponents. *Diff. Int. Equations*. In print.

Breda, D., Maset, S., and Vermiglio, R. (2005). Pseudospectral differencing methods for characteristic roots of delay differential equations. *SIAM J. Sci. Comput.*, 27(2), 482–495.

Breda, D., Maset, S., and Vermiglio, R. (2010). Approximation of the spectrum of evolution operators for periodic linear Retarded Functional Differential Equations. In preparation.

Butcher, E.A., Ma, H.T., Bueler, E., Averina, V., and Szabo, Z. (2004). Stability of linear time-periodic delay-differential equations via chebyshev polynomials. *Int. J. Numer. Meth. Engng*, 59, 895–922.

Canuto, C., Hussaini, M.Y., Quarteroni, A., and Zang, T. (2007). *Spectral Methods. Evolution to Complex Geometries and Applications to Fluid Dynamics*. Scientific Computation Series. Springer, Berlin, Germany.

Chatelin, F. (1983). *Spectral approximation of linear operators*. Academic Press, New York.

Chicone, C. and Latushkin, Y. (1999). *Evolution semigroups in dynamical systems and differential equations*. Number 70 in SURV series. American Mathematical Society, USA.

Davis, P.J. (1975). *Interpolation & approximation*. Dover, New York, USA.

Delfour, M.C. (1977). State theory of linear hereditary differential systems. *J. Differ. Equations*, 60, 8–35.

Delfour, M.C. and Mitter, S.K. (1972). Hereditary differential systems with constant delays. I. General case. *J. Differ. Equations*, 12, 213–235.

Diekmann, O., van Gils, S.A., Verduyn Lunel, S.M., and Walther, H.O. (1995). *Delay Equations - Functional, Complex and Nonlinear Analysis*. Number 110 in AMS series. Springer Verlag, New York, USA.

Engel, K. and Nagel, R. (1999). *One-Parameter Semigroups for Linear Evolution Equations*. Number 194 in Graduate texts in mathematics. Springer-Verlag, New York, USA.

Engelborghs, K., Luzyanina, T., and Roose, D. (2002). Numerical bifurcation analysis of delay differential equations using DDE-BIFTOOL. *ACM T. Math. Software*, 28(1), 1–21.

Engelborghs, K. and Roose, D. (2002). On stability of LMS methods and characteristic roots of delay differential equations. *SIAM J. Numer. Anal.*, 40(2), 629–650.

Faber, G. (1914). Über die interpolatorische darstellung stetiger funktionen. *Jahresber. Deut. Math. Verein.*, 23, 192–210.

Farmer, D. (1982). Chaotic attractors of an infinite-dimensional dynamical system. *Physica D*, 4, 605–617.

Hadd, S., Rhandi, A., and Schnaubelt, R. (2008). Feedback theory for time-varying regular linear systems with input and state delays. *IMA J. Math. Control Inform.*, 25(1), 85–110.

Hale, J.K. (1977). *Introduction to functional differential equations*. Number 99 in AMS series. Springer Verlag, New York, USA, 1st edition.

Insperger, T. and Stépán, G. (2002). Semi-discretization method for delayed systems. *Int. J. Numer. Meth. Engng*, 55, 503–518.

Ito, K. and Kappel, F. (1991). A uniformly differentiable approximation scheme for delay systems using splines. *Appl. Math. Opt.*, 23, 217–262.

Jarlebring, E. (2008). *The spectrum of delay-differential equations: numerical methods, stability and perturbation*. Ph.D. thesis, Inst. Comp. Math, TU Braunschweig.

Kappel, F. (1986). *Semigroups and delay equations*. Number 152 (Trieste, 1984) in Pitman Res. Notes Math. Ser. Longman Sci. Tech., Harlow.

Krasovskii, N. (1959). *Stability of Motion*. Moscow. English transl. Stanford University Press, 1963.

Krylov, V.I. (1956). Convergence of algebraic interpolation with respect to roots of Chebyshev's polynomial for absolutely continuous functions of bounded variation. *Dokl. Akad. Nauk SSSR*, 107, 362–365.

Natanson, I.P. (1965). *Constructive function theory Vol. III*. Frederick Ungar Publ., New York, USA.

Peichl, G.H. (1982). A kind of "history space" for retarded functional differential equations and representation of solutions. *Funkc. Ekvacioj-SER I*, 25, 245–256.

Trefethen, L.N. (2000). *Spectral methods in MATLAB*. Software - Environment - Tools series. SIAM, Philadelphia, USA.

Verheyden, K., Luzyanina, T., and Roose, D. (2008). Efficient computation of characteristic roots of delay differential equations using lms methods. *J. Comput. Appl. Math.*, 214(1), 209–226.

Vinter, R.B. (1978). On the evolution of the state of linear differential delay equations in M^2: properties of the generator. *J. Inst. Maths. Applics.*, 21, 13–23.

Vyhlídal, T. and Zítek, P. (2009). Mapping based algorithm for large-scale computation of quasi-polynomial zeros. *IEEE T. Automat. Cont.*, 54(1), 171–177.

Wu, J. (1996). *Theory and applications of partial functional differential equations*. Number 119 in AMS series. Springer-Verlag, New York, USA.