Stability in Time-Delay Systems

Zur Erlangung des akademischen Grades einer

DOKTORIN DER INGENIEURWISSENSCHAFTEN
(Dr.-Ing.)

von der KIT-Fakultät für Maschinenbau des
Karlsruher Instituts für Technologie (KIT)

angenommene

DISSERTATION

von

M.Sc. Tessina H. Scholl

Tag der mündlichen Prüfung: 22. August 2024
Hauptreferent: apl. Prof. Dr.-Ing. Lutz Gröll
Korreferenten: Prof. Dr. Veit Hagenmeyer
Prof. Dr.-Ing. habil. Dipl.-Math. Klaus Röbenack

Abstract

The present thesis is dedicated to the stability analysis of systems with a constant state delay. A control engineer is confronted with this class of systems whenever a feedback law cannot react instantaneously to the system variables but only with a fixed time delay.

An essential prerequisite for the further stability analysis in a nonlinear or uncertain system is the ability to test for stability in a linear nominal system. Therefore, as a first contribution, the present thesis derives a necessary and sufficient criterion for delay-independent stability, and it shows how to determine the critical delay at which an initial stability is lost.

Concerning the further stability analysis, which equals the analysis of robustness of the nominal system, so-called complete-type Lyapunov–Krasovskii functionals can be used. The present thesis proposes a new numerical approach to this known concept. The approach is based on an approximation of the time-delay system by an ordinary differential equation. Only a matrix-valued Lyapunov equation has to be solved to obtain a finite-dimensional approximation of the Lyapunov–Krasovskii functional.

However, the numerical method also opens up the possibility to determine better adapted functionals that can rely on a more general algebraic Riccati equation instead of the Lyapunov equation. Motivated by this finding, the present thesis introduces a new class of Lyapunov–Krasovskii functionals—so-called Lyapunov–Krasovskii functionals of robust type. These draw on methods from the field of absolute stability. In particular, the achievable robustness statements are significantly less restrictive than those of Lyapunov–Krasovskii functionals of complete type.

Deutsche Kurzfassung

Die vorliegende Dissertation widmet sich der Stabilitätsanalyse von Systemen, die eine konstante Zustandstotzeit aufweisen. Der Regelungstechniker sieht sich mit dieser Systemklasse konfrontiert, sobald ein Reglergesetz nicht instantan, sondern nur mit einer festen Zeitverzögerung auf die Systemgrößen reagieren kann.

Für Stabilitätsbetrachtungen in nichtlinearen oder unsicherheitsbehafteten Systemen sollte zunächst die Stabilität eines linearen Nominalsystems bestimmt werden können. Diesbezüglich stellt die Arbeit ein notwendiges und hinreichendes Kriterium für totzeitunabhängige Stabilität auf und zeigt wie die kritische Totzeit, bei der eine anfängliche Stabilität verloren geht, bestimmt werden kann.

Zur weitergehenden Stabilitätsanalyse, die einer Analyse der Robustheit des Nominalsystems gleichkommt, lassen sich sogenannte vollständige Lyapunov-Krasovskii-Funktionale verwenden. Die Arbeit schlägt einen neuen numerischen Zugang zu diesem bekannten Konzept vor. Er basiert auf einer Approximation des Totzeitsystems durch eine gewöhnliche Differentialgleichung. Lediglich eine matrixwertige Lyapunov-Gleichung muss gelöst werden, um eine endlichdimensionale Approximation des Lyapunov-Krasovskii-Funktionals zu erhalten.

Die numerische Methode eröffnet aber auch die Möglichkeit, angepasstere Funktionale aufzustellen, die auf einer allgemeineren algebraischen Riccati-Gleichung anstelle der Lyapunov-Gleichung basieren können. Motiviert durch diese Erkenntnis führt die vorliegende Arbeit eine neue Klasse von Lyapunov-Krasovskii-Funktionalen ein – sogenannte robuste Lyapunov-Krasovskii-Funktionale. Diese beruhen auf Methoden aus dem Bereich der absoluten Stabilität. Insbesondere sind die erreichbaren Robustheitsaussagen signifikant weniger restriktiv als jene von vollständigen Funktionalen.

Preface / Danksagung

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftliche Mitarbeiterin am KIT. Ich bin dankbar, in der Arbeitsgruppe Systemtheorie und Regelungstechnik von Prof. Lutz Gröll am Institut für Automation und angewandte Informatik von Prof. Veit Hagenmeyer zu diesen Themen forschen zu dürfen.

Ausschlaggebend für die Entscheidung, bei Prof. Lutz Gröll zu promovieren, war seine mehrteilige Vorlesung zu modernen Regelungskonzepten. Auch während der Promotion haben mir die Diskussionen zu den sich stetig weiterentwickelnden Vorlesungsinhalten immer viel Freude und zahlreiche neue Einsichten bereitet. Ihm und Herrn Prof. Hagenmeyer möchte ich für die Förderung meiner Arbeit und die wissenschaftlichen Freiheiten außerordentlich danken. Herrn Prof. Röbenack danke ich vielmals für die Übernahme des Korreferats und für den Austausch bei den GMA Workshops in Salzburg. Ein besonderer Dank gilt auch Herrn Prof. Bretthauer für den Zuspruch. Herrn Prof. Stiller danke ich für die Übernahme des Prüfungsvorsitzes. Danken möchte ich außerdem Herrn Prof. Dimitri Breda für die einsichtsreichen Workshops in Udine und Leiden und dem Organisationsteam des kommenden IFAC Workshops on Time Delay Systems für die Aufnahme in das IPC der Konferenz. Nicht zuletzt gilt ein Dank meinen Wegbegleitern am Institut, vor allem Julian Hofmann sowie Jan Wachter, Dominique Jumar, Anton Ponomarev, Jan Schaßberger, Adam Kastner, Prof. Jörg Matthes, Anselm Erdmann, Kathrin Reibelt, Prof. Timm Faulwasser und Tillmann Mühlpfordt.

Ganz besonders will ich mich aber bei meinen Eltern und bei meinem Bruder bedanken, die mich immer so sehr unterstützt haben.

Karlsruhe, Mai 2024 *Tessina Scholl*

Contents

Notation

Notation Related to Functions

$C([a,b],\mathbb{R}^n)$	continuous $\mathbb{R}^n$-valued functions on $[a,b]$
C	short notation for $C([-h,0],\mathbb{R}^n)$
$L_2([a,b],\mathbb{R}^n)$	square integrable $\mathbb{R}^n$-valued functions on $[a,b]$
L_2	short notation for $L_2([-h,0],\mathbb{R}^n)$ or $L_2([-h,0],\mathbb{C}^n)$
M_2	$M_2 = L_2([-h,0],\mathbb{R}^n) \times \mathbb{R}^n$ or $L_2([-h,0],\mathbb{C}^n) \times \mathbb{C}^n$
$\|\phi\|_C$	norm in C, $\|\phi\|_C = \max_{\theta\in[-h,0]} \|\phi(\theta)\|$
$\langle \phi_1, \phi_2 \rangle_{L_2}$	L_2 inner product of $\phi_1, \phi_2 \in L_2([-h,0],\mathbb{C}^n)$, i.e., $\langle \phi_1, \phi_2 \rangle_{L_2} = \int_{-h}^{0} (\phi_2(\theta))^H \phi_1(\theta)\, \mathrm{d}\theta$
$\|\phi\|_{L_2}$	L_2-norm of $\phi \in L_2([-h,0],\mathbb{C}^n)$, i.e., $\|\phi\|_{L_2} = \sqrt{\langle \phi, \phi \rangle_{L_2}}$
$\left\langle \left[\begin{smallmatrix}\phi_1\\ r_1\end{smallmatrix}\right], \left[\begin{smallmatrix}\phi_2\\ r_2\end{smallmatrix}\right] \right\rangle_{M_2}$	M_2 inner product of $\left[\begin{smallmatrix}\phi_1\\ r_1\end{smallmatrix}\right], \left[\begin{smallmatrix}\phi_2\\ r_2\end{smallmatrix}\right] \in M_2$, i.e., $\left\langle \left[\begin{smallmatrix}\phi_1\\ r_1\end{smallmatrix}\right], \left[\begin{smallmatrix}\phi_2\\ r_2\end{smallmatrix}\right] \right\rangle_{M_2} = \int_{-h}^{0} (\phi_2(\theta))^H \phi_1(\theta)\, \mathrm{d}\theta + r_2^H r_1$
AC	absolutely continuous functions
$\mathrm{P}_N([a,b],\mathbb{R}^n)$	$\mathbb{R}^n$-valued polynomials of degree at most N defined on $[a,b]$
$0_{n_{[a,b]}}$	vector-valued zero function on $[a,b]$, i.e., $\phi = 0_{n_{[a,b]}} \iff \phi(\theta) \equiv 0_n \in \mathbb{C}^n, \theta \in [a,b]$

$\mathscr{A}^*$	adjoint of an operator $\mathscr{A}$
$\mathsf{a} \otimes \mathsf{b}$	dyadic product between two functions $\mathsf{a} \in M_2$ and $\mathsf{b} \in M_2$, see Def. A.2.3
$\mathcal{K}$	set of class-K functions, $\mathcal{K} = \{\kappa \in C([0,\infty), \mathbb{R}_{\geq 0}) : \kappa(0) = 0, \text{ strictly increasing}\}$
$D^+_{(\mathrm{eq})}V$	upper right-hand derivative of $V : X \to \mathbb{R}$ along solutions of (eq), i.e., if $x_t \in X$ denotes the state of (eq) at time $t \geq 0$, then $D^+_{(\mathrm{eq})}V(x_0) = \limsup_{t\to 0^+} \frac{V(x_t)-V(x_0)}{t}$
$D^+_f V$	$D^+_f V := D^+_{(1.4)}V$, derivative of V along solutions of the linear system (1.4)
$D^+_{(f+g)}V$	$D^+_{(f+g)}V := D^+_{(1.2)}V$, derivative of V along solutions of the perturbed system (1.2)
$p_k(\cdot)$	k-th Legendre polynomial
$T_k(\cdot)$	k-th Chebyshev polynomial of first kind
$\ell_k(\cdot)$	Lagrange interpolation polynomial w.r.t. the k-th interpolation node
$b_k(\cdot)$	k-th Bessel polynomial
$\mathrm{padé}_{[N/N]}(f)$	Padé approximant of order $[N/N]$ of the function f
$H_\infty(\mathbb{C}^+, \mathbb{C}^{p\times m})$	Hardy space of holomorphic $\mathbb{C}^{p\times m}$-valued functions on $\mathbb{C}^+$
H_∞	short notation for $H_\infty(\mathbb{C}^+, \mathbb{C}^{p\times m})$
$\|G\|_\infty$	H_∞-norm of $G \in H_\infty$, i.e., $\|G\|_\infty = \sup_{s\in\mathbb{C}^+} \|G(s)\|_2 = \sup_{\omega\in\mathbb{R}} \|G(\mathrm{i}\omega)\|_2$
$(\mathbb{R}(s))^{p\times m}$	$p \times m$ matrices with entries being real-rational functions (i.e., ratios of polynomials with real coefficients) of $s \in \mathbb{C}$

Notation Related to $\mathbb{R}^n$ or $\mathbb{C}^n$

$\lambda_k(M)$	k-th eigenvalue of a matrix $M \in \mathbb{C}^{n\times n}$, $k \in \{1,\dots,n\}$ (arbitrarily ordered)
$\sigma(M)$	spectrum of $M \in \mathbb{C}^{n\times n}$, i.e., $\sigma(M) = \bigcup_k \{\lambda_k\}$
$\lambda_k(M,E)$	k-th eigenvalue of a matrix pencil (M,E), $\lambda_k(M,E) \in \sigma(M,E) = \{\lambda \in \mathbb{C} : \det(\lambda E - M) = 0\}$
$\rho(M)$	spectral radius of $M \in \mathbb{C}^{n\times n}$, i.e., $\rho(M) = \max_k \lvert\lambda_k(M)\rvert$
$\alpha(M)$	spectral abscissa of $M \in \mathbb{C}^{n\times n}$, i.e., $\alpha(M) = \max_k \operatorname{Re}\lambda_k(M)$
$\operatorname{rk}(M)$	rank of a matrix $M \in \mathbb{C}^{m\times n}$
I_n	identity matrix in $\mathbb{C}^{n\times n}$
$0_{m\times n}$, 0_n	zero matrix of dimension $m \times n$, zero vector in $\mathbb{C}^n$
e_k	k-th canonical basis vector of $\mathbb{C}^n$
$M^H = \overline{M}^\top$	conjugate transpose of $M \in \mathbb{C}^{n\times n}$
$Q \succ (\succeq)\, 0_{n\times n}$	positive (semi)definiteness of $Q \in \mathbb{C}^{n\times n}$, implicitly requiring that $Q = Q^H$
$\lambda_{\min}(Q)$, $\lambda_{\max}(Q)$	smallest and largest eigenvalue of $Q = Q^H \in \mathbb{C}^{n\times n}$
$\operatorname{He}(M)$	Hermitian part of $M \in \mathbb{C}^{n\times n}$, i.e., $\operatorname{He}(M) = \frac{1}{2}(M + M^H)$
$\operatorname{sym}(M)$	symmetric part of $M \in \mathbb{C}^{n\times n}$, i.e., $\operatorname{sym}(M) = \frac{1}{2}(M + M^\top)$
$\operatorname{diag}(m_1,\dots,m_n)$	diagonal matrix with entries $m_1,\dots,m_n$
$\operatorname{blkdiag}(A_1,\dots,A_m)$	block diagonal matrix built from matrices $A_1,\dots,A_m$

$A \otimes B$	Kronecker product of matrices $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{r\times s}$, i.e., $\begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix} \otimes B = \begin{bmatrix} A_{11}B & \cdots & A_{1n}B \\ \vdots & & \vdots \\ A_{m1}B & \cdots & A_{mn}B \end{bmatrix} \in \mathbb{C}^{mr\times ns}$
$A \oplus B$	Kronecker sum of matrices $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{s\times s}$, i.e., $A \otimes I_r + I_n \otimes B$
A^-	generalized inverse of a matrix $A \in \mathbb{C}^{m\times n}$, i.e., $A^- \in \mathbb{C}^{n\times m}$ (not necessarily unique): $AA^-A = A$
$[\,-A_1-\ \ A_2\,]$	block matrix $[A_1\ \ A_2]$ emphasizing that the submatrices $A_1 \in \mathbb{C}^{n\times nN}$, $A_2 \in \mathbb{C}^{n\times n}$ are differently sized
$\mathrm{vec}(M)$	vectorization of $M \in \mathbb{C}^{m\times n}$
$\|x\|$	arbitrary norm of $x \in \mathbb{C}^n$
$\|x\|_2$	Euclidean norm of $x \in \mathbb{C}^n$, i.e., $\|x\|_2 = \sqrt{x^H x}$
$\|M\|_2$	spectral norm of $M \in \mathbb{C}^{m\times n}$, i.e., $\|M\|_2 = \sqrt{\lambda_{\max}(M^H M)}$
$\mu_2(M)$	logarithmic norm w.r.t. the spectral norm, i.e., $\mu_2(M) = \lambda_{\max}(\frac{1}{2}(A^H + A))$
$(w_k)_{k\in\mathcal{I}}$	a vector with entries w_k, e.g., $(w_k)_{k\in\{0,\ldots,N\}} = [w_0, \ldots, w_N]^\top$
$(w_k)_k$	$(w_k)_{k\in\mathcal{I}}$ with the index set $\mathcal{I}$ being clear from the context

Further Notation

$\mathbb{R}_{\geq 0}, \mathbb{R}_{>0}$	nonnegative and positive real numbers, i.e., $[0,\infty), (0,\infty)$
$\mathbb{C}^-, \mathbb{C}^+$	open left and open right complex half-plane, i.e., $\mathbb{C}^- = \{s \in \mathbb{C} : \mathrm{Re}(s) < 0\}$ and $\mathbb{C}^+ = \{s \in \mathbb{C} : \mathrm{Re}(s) > 0\}$
$\overline{S}$	closure of a set S

$S_1 \subseteq S_2$, $S_1 \subset S_2$	subset, proper subset
mod	modulo operator
$\overset{!}{=}$	the equality is emphasized to be a requirement
$\overset{(...)}{=}$	the relation is explained by $(\ldots)$
$\overset{\mathrm{def}}{=}$	definition (universally valid)
$:=$	assignment (definition of a variable, chosen in the special context)
δ_{jk}	Kronecker delta, i.e., $\delta_{jk} = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k \end{cases}$

Typeface

$\mathcal{A}, \mathcal{C}, \mathcal{K}$	operators with domain or codomain in $C = C([-h, 0], \mathbb{R}^n)$
ϕ, ψ	functions in C or L_2
$\mathscr{A}, \mathscr{B}, \mathscr{K}, \mathscr{P}, \mathscr{Q}$	operators with domain or codomain in M_2
$\mathsf{a}, \mathsf{b}, \mathsf{g}, \mathsf{h}, \mathsf{x}, \mathsf{y}, \mathsf{z}$	elements in M_2, e.g., $\mathsf{x} = \left[\begin{smallmatrix} \phi \\ r \end{smallmatrix}\right]$ with $\phi \in L_2$ and $r \in \mathbb{C}^n$

Abbreviations

a.e.	almost everywhere
s.t.	subject to
w.l.o.g.	without loss of generality
w.r.t.	with respect to
vs.	versus

Acronyms

ARE	algebraic Riccati equation
ARI	algebraic Riccati inequality
DAE	differential algebraic equation
ES	exponential stability
JP	Jacobi polynomial
KYP lemma	Kalman–Yakubovich–Popov lemma
LK functional	Lyapunov–Krasovskii functional
LMI	linear matrix inequality
LQR problem	linear-quadratic regulator problem
LTI system	linear time-invariant system
MIMO system	multiple-inputs multiple-outputs system
ODE	ordinary differential equation
PDE	partial differential equation
RFDE	retarded functional differential equation
SISO system	single-input single-output system

1 Introduction

Imagine a controller at an oscillating system always undertakes what would have been correct a short time ago. Clearly, such a well-meant controller can have the opposite effect to what was intended. Unfortunately, a delayed controller action is the rule rather than the exception: measurements, communication, data processing or the actuator response almost always cause a delay. Fortunately, in most standard control loops these time delays are very small and therefore negligible—but, how can we recognize how much delay is small and which one will be harmful? This is one of the first questions tackled in the present thesis. Already the otherwise effortless eigenvalue-based test for exponential stability of an equilibrium is no longer a trivial task once time delays are involved. However, regardless of whether delays are involved or not, such a confirmation of exponential stability for a nominal linear (respectively linearized) system is usually not enough. Neglected additional nonlinear terms or uncertainties call for a further analysis enabling regional or global stability results or robustness bounds. In a delay-free system, this motivates the construction of a Lyapunov function. How to proceed if a time delay is involved? Answering this second question, in fact, constitutes the major part of the present thesis.

Although this work is primarily theoretical in nature, it has a very clear practical motivation as time delays are ubiquitous in control systems (still noting that many applications call for generalizations to multiple or time-varying delays, which, for the sake of compactness, are not covered in the present thesis—see the outlook in Section 8.2). Consider, for instance, networked systems like energy systems that are to become smart. In this field, wide area measurement systems are a great innovation. However, the goal cannot be to just watch the energy system swing into a blackout, and, therefore, the wide-area measurement system must become

a wide-area control system. Because of the significant communication-induced delay, the closed-loop system becomes a time-delay system [141, 201, 193, 202, 145]. In fact, this delay is considered to be one of the most important obstacles for the implementation of wide area control systems [147]. The system need not even be such a large scale one. Already in the control of microgrids, the time delay problem is seen to be a very critical issue [195, 126, 164]. In the end, even small delays might be harmful since smallness is relative and a matter of the relevant time scales. For instance, in the control of power converters, the time delay is less than a half millisecond and still a cause of instability [58, 173, 92, 150, 205]. Contrary to the common intuition, in certain ranges, the delay in power converters can also have a stabilizing effect [206, 189, 188]. In any case, it definitely should not be neglected, although this is frequently done in the literature. Besides of the fact that any closed loop system becomes a time-delay system once the controller does not act instantaneously, many other examples could be mentioned from various disciplines, be it inherent time delays in systems from mechanical engineering, biology, economy, or other fields. See, e.g., [172, 95, 67, 112, 36] and references therein.

1.1 System Class

If a control law $u = k(x)$ is designed for a system $\dot{x} = p(x, u)$, it is usually assumed that the closed loop obeys the delay-free dynamics $\dot{x} = p(x, k(x)) =: F_0(x)$. If, however, the controller acts with a delay $h > 0$ (induced by the required measurements, by network communication, by calculation times or by the actuator response), then the resulting closed loop dynamics become $\dot{x}(t) = p(x(t), k(x(t-h))) =: F(x(t), x(t-h))$, which is a retarded functional differential equation (RFDE).

In contrast to the delay-free case, for time-delay systems the specification of an initial vector $x(0) \in \mathbb{R}^n$ does not bear enough information to calculate the solution $x(t), t \geq 0$. Already $\dot{x}(0)$ depends on $x(-h)$ and similar holds for $\dot{x}(t)$ on $t \in [0, h)$. That is why an initial function $x_0(t) = x(t)$ on $t = \theta \in [-h, 0]$

is required. The RFDE state $x_t \in C([-h,0],\mathbb{R}^n)$ at time $t \geq 0$ accordingly describes the solution segment on the preceding delay interval $[t-h,t]$, defined as

$$x_t(\theta) \stackrel{\text{def}}{=} x(t+\theta), \quad \theta \in [-h,0]. \tag{1.1}$$

In the present thesis, the thus considered RFDE

$$\begin{aligned} \dot{x}(t) &= F(x(t), x(t-h)) \\ &= \underbrace{A_0 x(t) + A_1 x(t-h)}_{f(x_t)} + \underbrace{\tilde{g}\big(x(t), x(t-h)\big)}_{g(x_t)} \end{aligned} \tag{1.2}$$

is decomposed into a linear part, $\dot{x}(t) = f(x_t)$ with $A_0, A_1 \in \mathbb{R}^{n\times n}$, and a possibly nonlinear term $g(x_t)$. For notational compactness, the perturbation $g\colon C([-h,0],\mathbb{R}^n) \to \mathbb{R}^n; \phi \mapsto g(\phi)$ is assumed to be time-invariant. Still, the results straightforwardly extend to a time-varying $(t,\phi) \mapsto g(t,\phi)$ (see Remark 5.3.3). Note that, throughout this thesis, the delay is assumed to be constant. Moreover, for simplicity, g is assumed to be locally Lipschitz continuous, ensuring well-posedness of (1.2), see [82]. Furthermore, $g(0_{n_{[-h,0]}}) = 0_n$. Various scenarios give rise to the decomposition in (1.2).

1. The nonlinearity $g(x_t)$ might consist of higher order terms representing the remainder from a linearization $\dot{x}(t) = A_0 x(t) + A_1 x(t-h)$.

2. The nonlinearity $g(x_t)$ might involve a saturation, which is frequently encountered if a delayed control law operates on a constrained input.

3. Uncertainties $\Delta_0, \Delta_1 \in \mathbb{R}^{n\times n}$ added to A_0, A_1 in (1.2) can be addressed by $g(x_t) = \Delta_0 x(t) + \Delta_1 x(t-h)$.

1.2 A Template from Delay-Free Systems

The present thesis is concerned with the time-delay system (1.2). However, as a template and for comparison, it is worth taking a look at the analogous delay-free system

$$\dot{x}(t) = F_0(x(t)) = Ax(t) + g(x(t)), \tag{1.3}$$

$A \in \mathbb{R}^{n\times n}$, $g \in C(\mathbb{R}^n, \mathbb{R}^n)$, $g(0_n) = 0_n$. In that delay-free system, the following simple approach is possible, see, e.g., [108].

i Clearly, exponential stability holds for the zero equilibrium of the nominal linear system $\dot{x} = Ax$ if and only if all **eigenvalues** of A have negative real parts. If $\dot{x} = Ax$ represents the linearization of $\dot{x} = Ax + g(x)$, the latter also implies local asymptotic stability of the zero equilibrium in the nonlinear system.

ii However, such a local statement is usually not enough, and a Lyapunov function is additionally desirable for the further analysis of the overall system. The probably simplest concept lies in the construction of a quadratic Lyapunov function $V(x) = x^\top Px$ that provides a desired negative definite derivative $D^+_{(\dot{x}=Ax)}V(x) = -x^\top Qx$ along the solutions of the nominal linear system. To this end, only the corresponding **Lyapunov equation** $PA + A^\top P = -Q$ must be solved for P, which requires just one line of Matlab code. With Q being chosen positive definite, the solution P of the Lyapunov equation is well known to be positive definite if and only if A is Hurwitz [108]. The largest possible coefficient $k_1 > 0$ in $k_1\|x\|_2^2 \leq V(x)$ is simply given by the smallest eigenvalue of P.

iii The Lyapunov function $V(x)$ constructed in this way also has a negative definite derivative along the solutions of $\dot{x} = Ax + g(x)$ provided the perturbation g is sufficiently small—to be more precise, if $\|g(x)\|_2 \leq \gamma\|x\|_2$ with $\gamma < \frac{\lambda_{\min}(Q)}{2\lambda_{\max}(P)}$ [108]. Consequently, this **linear norm bound** γ is a robustness bound, which characterizes admissible terms $g(x)$ that do not

compromise the global exponential stability known from the nominal linear system. Of course, a nonlinear term $g(x)$ might not remain globally below such a linear norm bound. The linear norm bound might instead be satisfied locally on some set $G = \{x \in \mathbb{R}^n : \|g(x)\|_2 \leq \gamma\|x\|_2\}$. Then any sublevel set of $V(x)$ within that set G is guaranteed to belong to the domain of attraction. For instance, if there is some radius $r > 0$ such that the linear norm bound is known to be satisfied on $B_r = \{x \in \mathbb{R}^n : \|x\|_2 < r\} \subseteq G$, then such a subset of the domain of attraction is given by the sublevel set $\{x \in \mathbb{R}^n : V(x) < k_1 r^2\} \subseteq B_r \subseteq G$ with k_1 from ii, cf. [108].

1.3 Methods for Proving Stability in Time-Delay Systems

As a counterpart to the simple delay-free **template i** from Section 1.2, the following considerations focus on the linear nominal time-delay system

$$\dot{x}(t) = A_0 x(t) + A_1 x(t-h) =: f(x_t) \tag{1.4}$$

with $A_0, A_1 \in \mathbb{R}^{n\times n}, h > 0$. Stability terms are defined similar to the ODE case and will formally be given in Definition 4.5.2. So far, however, it suffices to note that for the linear time-delay system (1.4), asymptotic stability of the zero equilibrium[1] coincides with (global) exponential stability of the zero equilibrium [82, Thm. 5.3 in Ch. 6], i.e., the solutions $x(t)$ from all initial functions $x_0 \in C([-h, 0], \mathbb{R}^n)$ decay exponentially according to

$$\exists b < 0, \exists c \geq 1, \forall t \geq 0 : \quad \|x(t)\| \leq c\mathrm{e}^{bt}\|x_0\|_C, \tag{1.5}$$

where $\|x_0\|_C \stackrel{\text{def}}{=} \max_{\theta\in[-h,0]} \|x_0(\theta)\|$.

[1] The zero equilibrium refers to the zero state $0_{n_{[-h,0]}}$ in the state space $C([-h, 0], \mathbb{R}^n)$, i.e., the vector-valued zero function on the delay interval $[-h, 0]$.

The effect of the additional perturbation $g(x_t)$ in (1.2) will be tackled in the subsequent Section 1.4. However, anyway, if $g(x_t)$ only represents the remainder from a linearization (scenario 1 in Section 1.1), asymptotic stability of the zero equilibrium in the linear system (1.4) already implies local asymptotic stability of the zero equilibrium in the overall system (1.2) (principle of linearized stability, [57]).

Remark 1.3.1. *This linearization-based local statement of stability, in fact, can also be drawn as a conclusion from the robustness considerations in the next section. Similar to template iii, these considerations provide some nonzero linear norm bound γ characterizing admissible perturbations $g(x_t)$ in (1.2). By definition, a little-o term*[2] *$g(x_t)$ where $g(\phi) \in o(\phi)$, respectively $\tilde{g}\big(\left[\begin{smallmatrix} x(t) \\ x(t-h) \end{smallmatrix}\right]\big)$ where*[3] *$\tilde{g}(\zeta) \in o(\zeta)$, locally satisfies any linear norm bound γ. Thus, however small the linear norm bound γ on admissible perturbations from the next section will become, higher order terms $g(x_t)$ that are neglected in a linearization always remain below that bound if the consideration is restricted to a local domain G of sufficiently small states x_t. Therefore, local asymptotic stability of the zero equilibrium in the overall system (1.2) can immediately be concluded from the asymptotic stability in the linearized system as the latter comes along with such a nonzero admissible linear norm bound.*

As described in Section 1.1, the considerations in the present thesis are not restricted to this special case of perturbations $g(x_t)$ that represent the remainder of a linearization. However, also in more general cases, the ability to prove stability of linear RFDEs, which is the subject of the present section, is immanent.

---

[2] In fact, the principle of linearized stability calls for the Fréchet differentiability of the RFDE right-hand side at $\phi = 0_{n_{[-h,0]}}$ in $C([-h,0],\mathbb{R}^n)$, which exactly amounts to the decomposability in a linear part $f(x_t)$ and a rest $g(x_t)$ that is a little-o term.

[3] Note that $\tilde{g}(\zeta) \in o(\zeta)$ is defined by $\lim_{\|\zeta\|\to 0} \frac{\|\tilde{g}(\zeta)\|}{\|\zeta\|} = 0$, which by the ε-δ-definition of the limit becomes $\forall \varepsilon > 0, \exists \delta_\varepsilon > 0 : \|\zeta\| < \delta_\varepsilon \Rightarrow \frac{\|\tilde{g}(\zeta)\|}{\|\zeta\|} < \varepsilon$, respectively, $\|\tilde{g}(\zeta)\| < \varepsilon\|\zeta\|$. Concerning (1.17), choose $\zeta = \left[\begin{smallmatrix} x(t) \\ x(t-h) \end{smallmatrix}\right]$ and $\varepsilon = \gamma$.

Similar to the delay-free template i, the **characteristic roots**, which are the solutions $s \in \mathbb{C}$ of the characteristic equation

$$\det(sI_n - A_0 - \mathrm{e}^{-sh}A_1) = 0 \tag{1.6}$$

(being derived from an exponential ansatz or from a Laplace transform), are still meaningful: The zero equilibrium of (1.4) is exponentially stable for a certain delay $h > 0$ if and only if all characteristic roots s have a negative real part [82]. However, in contrast to the n eigenvalues of A in the delay-free case i from Section 1.2, there is generically an **infinite number of roots** $s \in \mathbb{C}$ that solve (1.6), which makes the stability analysis more involved. In the last decades, various approaches for proving stability have been established, see, e.g., [76, 149, 113, 67, 31] and references therein.

There are approaches that aim at a **numerical calculation of the most important characteristic roots** for a given delay, see, e.g., [30, 194, 185, 64, 97]. Some of these rely on the eigenvalue computation of a matrix that describes an **ODE approximation** of the time-delay system. Such ODE approximations (see Chapter 3) will play a central role in the present thesis, but the field of application will be different.

Probably the majority of literature on stability in time-delay systems from the last two decades ends up with **linear matrix inequalities (LMIs)**. Therefore, a brief glance is in order. These LMIs might, for instance, result from **Lyapunov–Krasovskii (LK) functionals** (cf. Theorem 4.5.1). The following example demonstrates such an approach. Alternatively, LMIs might result from **Lyapunov–Razumikhin functions**, or from treating the time delay as an uncertainty in **robustness approaches**. See [82, 67, 31, 18].

Example 1.3.2. *Consider the simple LK functional*

$$V(x_t) = x^\top(t)Px(t) + \int_{t-h}^{t} x^\top(\eta)\, S\, x(\eta)\, \mathrm{d}\eta, \tag{1.7}$$

$P, S \succ 0_{n\times n}$. *Its derivative along solutions of (1.4)*

$$\begin{aligned} & D_f^+ V(x_t) \\ &= \left(A_0 x(t) + A_1 x(t-h)\right)^\top P x(t) + x^\top(t) P \left(A_0 x(t) + A_1 x(t-h)\right) \\ &\quad + x^\top(t) S x(t) - x^\top(t-h) S x(t-h) \\ &= \begin{bmatrix} x(t) \\ x(t-h) \end{bmatrix}^\top \left(\begin{bmatrix} PA_0 + A_0^\top P & PA_1 \\ A_1^\top P & 0_{n\times n} \end{bmatrix} + \begin{bmatrix} S & 0_{n\times n} \\ 0_{n\times n} & -S \end{bmatrix} \right) \begin{bmatrix} x(t) \\ x(t-h) \end{bmatrix} \end{aligned}$$

gives rise to the following LMI criterion:
If there exists a matrix $P = P^\top \in \mathbb{R}^{n\times n}$ *and a matrix* $S = S^\top \in \mathbb{R}^{n\times n}$ *such that*

$$\begin{bmatrix} PA_0 + A_0^\top P + S & PA_1 \\ PA_1 & -S \end{bmatrix} \prec 0_{2n\times 2n}, \quad P \succ 0_{n\times n}, \quad S \succ 0_{n\times n}, \tag{1.8}$$

then the zero equilibrium of (1.4) is asymptotically stable [31, Thm. 5.6.1], [67, Prop. 3.3], [76, Prop. 5.14].

The above criterion does not depend on the delay h in (1.4). Thus, it can either be concluded that the asymptotic, respectively exponential, stability holds for all delays $h > 0$, which is called **delay-independent stability**, or nothing is concluded. To obtain a **delay-dependent** result, double-integral terms are usually incorporated in V, and $D_f^+ V(x_t)$ is no longer equal to some quadratic form in $[x^\top(t), x^\top(t-h)]^\top$ but only upper bounded by some quadratic form in a possibly larger augmented vector [31]. A kind of competition has developed as to how nonconservative LMIs can be set up [167, Tab. 1]. What they have in common, however, is that non-conservative criteria usually come along with large LMIs.

When looking for non-conservative LK functionals, one will quickly stumble upon the fact that so-called **LK functionals of complete type** [110, 111] give rise to a necessary and sufficient condition for asymptotic stability. However, proving stability is originally not the raison d'être of such functionals. On the contrary, decisive results on these functionals (in particular, the formula for a lower bound

[110, 137]) only apply if stability has already been proven beforehand. Rather, complete-type and related LK functionals are the counterpart to template ii in the delay-free discussion from Section 1.2. Nevertheless, in the delay-free template, it would also be possible (but there is no reason to do so), instead of calculating the eigenvalues of A in i, first to solve the Lyapunov equation $PA + A^\top P = -Q$ from ii for some $Q \succ 0$, and then to test whether P (and thus also the Lyapunov function $x^\top Px$) is positive definite—which it is if and only if A was Hurwitz. However, how this approach translates to time-delay systems has long been an open question. It has only recently been resolved by Egorov, Gomez, and Mondié [60, 72, 142], describing how at all to conclude stability from the so-called delay Lyapunov matrix function, which is a building block of the corresponding complete-type and related LK functionals (see Remark 4.4.1). Still, in the present thesis, the reason why complete-type and related LK functionals are considered is not to prove stability in the linear system, but it is their role in the robustness analysis as a counterpart to the template iii from Section 1.2. The latter will be discussed in the next section. In accordance with the template i, stability will instead be proven a priori based on a consideration of the characteristic equation.

There are various methods for proving stability of the linear RFDE (1.4) that directly **rely on the characteristic equation** (1.6). If the stability only holds for sufficiently small delays $h \in [0, h_c)$, such approaches can exactly determine the (first) **critical delay** h_c at which exponential stability is lost, thus proving exponential (equivalently asymptotic) stability for the overall interval of delays $h \in [0, h_c)$ at once. Some of these methods resort to an evaluation of the determinant in (1.6) yielding the **quasipolynomial**

$$\det(sI_n - A_0 - \mathrm{e}^{-sh}A_1) = p_n(s)\mathrm{e}^{-nhs} + \ldots + p_1(s)\mathrm{e}^{-hs} + p_0(s) \tag{1.9}$$

(each $p_k(s)$, $k \in \{0, \ldots, n\}$, is a polynomial of degree at most $n-k$), respectively the corresponding **bivariate polynomial**

$$\det(sI_n - A_0 - zA_1) = p_n(s)z^n + \ldots + p_1(s)z + p_0(s), \tag{1.10}$$

[160, 175, 109, 87, 187, 44, 171, 169]. Others are based on eigenvalues of **matrix pencils** with block matrices containing Kronecker products of A_0 and A_1 [41, 132, 148]. Moreover, there are so-called **frequency sweeping tests**. In particular, delay-independent exponential stability holds if and only if A_0 and $A_0 + A_1$ are Hurwitz, and

$$\forall \omega > 0 : \quad \rho\left((\mathrm{i}\omega I_n - A_0)^{-1} A_1\right) < 1, \tag{1.11}$$

[50, Thm. 1.3], [42], [40], [96], where $\rho(M) = \max_k |\lambda_k(M)|$ is the spectral radius of $M \in \mathbb{C}^{n\times n}$.

The criteria listed up to this point, however, are not very insightful in terms of how A_0 and A_1 may look like without resorting to numerical computations. **Mori's criterion** for delay-independent exponential stability

$$\mu_2(A_0) + \|A_1\|_2 < 0, \tag{1.12}$$

[143], with $\mu_2(A_0) \stackrel{\text{def}}{=} \lambda_{\max}\left(\frac{1}{2}(A_0 + A_0^\top)\right)$, provides some more insights. However, Mori's criterion is rather conservative. Similarly insightful but less conservative is

$$\max_{\varphi\in[0,\pi]} \alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi} A_1) < 0, \tag{1.13}$$

relying on the spectral abscissa $\alpha(M) \stackrel{\text{def}}{=} \max_{k\in\{1,\ldots,n\}} \operatorname{Re}\lambda_k(M)$ of the parameter-dependent matrix $M(\varphi) = A_0 + \mathrm{e}^{\mathrm{i}\varphi} A_1 \in \mathbb{C}^{n\times n}$. Criterion (1.13) seems to be little known in this form, but it is a direct consequence of results by Datko [50, Thm. 1.2] and Kamen [106]. Unfortunately, (1.13) is only a sufficient not yet necessary criterion for delay-independent exponential stability. However, it will be further developed in Chapter 2 of the present thesis.

1.4 Lyapunov–Krasovskii Functionals of Complete Type

Complete-type [110, 111] and related [136] LK functionals are a recent field of research [204, 103, 117, 102, 142, 6]. As already mentioned above, these LK functionals are the counterpart to the Lyapunov-equation-based construction of Lyapunov functions in the **template ii** from Section 1.2.

Instead of the desired Lyapunov function derivative $D^+_{(\dot{x}=Ax)}V(x)$ encountered in ii, the **desired LK functional derivative** along solutions of the nominal linear RFDE (1.4) has to be prescribed. This derivative is commonly set as

$$\begin{aligned} D^+_f V(x_t) = &-x^\top(t)Q_0x(t) - x^\top(t-h)Q_1x(t-h) \\ &- \int_{-h}^{0} x^\top(t+\theta)Q_2x(t+\theta)\,\mathrm{d}\theta \end{aligned} \tag{1.14}$$

[110], with freely chosen $Q_0 \succ 0_{n\times n}$, $Q_1, Q_2 \succeq 0_{n\times n}$.

If $Q_{0,1,2} \succ 0_{n\times n}$, the LK functional $V(x_t)$ that accomplishes (1.14) is called an **LK functional of complete type** [110, Thm. 2.11]. The known formula for the solution of (1.14)

$$\begin{aligned} V(x_t) = &\, x^\top(t)\Psi(0;\tilde{Q})\,x(t) + 2\int_{-h}^{0} x^\top(t)\Psi(-h-\eta;\tilde{Q})A_1x(t+\eta)\,\mathrm{d}\eta \\ &+ \int_{-h}^{0}\int_{-h}^{0} x^\top(t+\xi)A_1^\top\Psi(\xi-\eta;\tilde{Q})A_1x(t+\eta)\,\mathrm{d}\eta\,\mathrm{d}\xi \\ &+ \int_{-h}^{0} x^\top(t+\eta)\big[Q_1+(h+\eta)Q_2\big]x(t+\eta)\,\mathrm{d}\eta \end{aligned} \tag{1.15}$$

relies on the so-called **delay Lyapunov matrix function**[4] $\Psi(\,\cdot\,;\tilde{Q})\colon [-h,h] \to \mathbb{R}^{n\times n}$ associated with $\tilde{Q} = Q_0 + Q_1 + hQ_2$. This matrix-valued function Ψ

[4] $\Psi(s;\tilde{Q})$ is commonly denoted by $U(s)$ in the literature

is defined via a matrix-valued time-delayed boundary-value problem (given in Appendix B.1.1) that first has to be solved semi-analytically (Proposition B.1.2) or numerically.

Regarding the numerical calculation of such functionals, the following should be noted. There are some seemingly related LMI [74, 75] or sum of squares approaches [151] that, however, do not seek for the actual complete-type LK functional. Rather they are inspired by the structure of the integral formula in (1.15), which promises to provide a necessary and sufficient stability criterion if appropriate kernel functions could be reproduced. Concerning the actual complete-type or related LK functional given in (1.15), it is a recent field of research how to make these functionals numerically traceable. Such **numerical results**, however, either rely on the knowledge of the delay Lyapunov matrix function[4] Ψ, [72, 60, 142, 136, 134, 48, 10, 62], or they aim to determine Ψ, [61, 140, 69, 94]. A different approach will be provided in Chapter 4.

As in the template ii, a lower bound on the functional plays a special role. If $Q_1 \succ 0_{n\times n}$, and if the nominal system is exponentially stable, then there exists a $k_1 > 0$ such that the functional has a **quadratic lower bound** of the form

$$k_1\|x(t)\|^2 \leq V(x_t) \tag{1.16}$$

[110, Lem. 2.10]. Similar to the template, the coefficient k_1 in (1.16) is, for instance, relevant in estimations of domains of attraction [137]. Still, in contrast to the ODE case ii, where the minimum eigenvalue of P gives the best possible coefficient k_1, nothing is reported about the conservatism of known formulas for k_1 in (1.16) [110, 137]—this issue, however, will also be tackled in Chapter 4.

It remains to discuss the counterpart of point iii in the template. Indeed, the robustness analysis of time-delay systems is an important application of complete-type [111, 137, 110] and related LK functionals [135, 6, 182]. The analogy to the **template iii** in terms of complete-type LK functionals is as follows. Not only the derivative $D_f^+V(x_t)$ along solutions of the nominal RFDE satisfies the requirements of the classical LK theorem (by construction), but also the derivative

$D^+_{(f+g)}V(x_t)$ along solutions of the perturbed RFDE (1.2) still satisfies these requirements, whenever $g(x_t) = \tilde{g}(x(t), x(t-h))$ in (1.2) remains below the **linear norm bound** [137]

$$\|\tilde{g}(x(t), x(t-h))\|_2 \leq \gamma \left\| \begin{bmatrix} x(t) \\ x(t-h) \end{bmatrix} \right\|_2 \tag{1.17}$$

$$\text{with } \gamma < \frac{\min\left\{ \frac{\lambda_{\min}(Q_0)}{2+h\|A_1\|_2}, \frac{\lambda_{\min}(Q_1)}{1+h\|A_1\|_2}, \frac{\lambda_{\min}(Q_2)}{\|A_1\|_2} \right\}}{\lambda_{\max}(\Psi(0))}. \tag{1.18}$$

See, e.g., [137, 182, 6, 3] on results regarding how to estimate the domain of attraction within the set on which that linear norm bound is valid.

1.5 Objectives and Outline

The objectives of the present thesis are in alignment with the template from delay-free systems given in Section 1.2.

I Concerning a counterpart to point i in the template, Section 1.3 above already makes clear that various techniques come into question for proving stability. However, a trade-off between the conservatism of the criteria and their complexity is noticeable. The increasing complexity also comes along with a vanishing insight in terms of how A_0 and A_1 actually might look like in order not to hamper stability. Therefore, the objective in **Chapter 2** is to extend the insightful formula (1.13) to a necessary and sufficient criterion for delay-independent exponential stability, and, if exponential stability only holds for sufficiently small delays $h \in [0, h_c)$, to provide the exact critical delay h_c at which this stability is lost. The focus is on the simplicity and interpretability of the gained results.

II The determination of complete-type and related LK functionals is far more elaborate than the simple Lyapunov equation from their delay-free template ii. With the objective to benefit from the enormous simplification that comes along with the treatment of ODEs in contrast to RFDEs, **Chapter 3**

discusses ODE-approximation schemes of time-delay systems. By employing the presented ODE approximations, **Chapter 4** intends to provide a new numerical approach to complete-type or related LK functionals which only requires to solve (a sequence of) Lyapunov equations. Another objective is the computation of a less conservative lower-bound coefficient k_1 in (1.16). Additionally to the computational advantages, the approach also aims at theoretical insights offered by the analogy to the stability theory of ODEs.

III Although robustness considerations have been the original purpose of complete-type LK functionals when introduced in [111], the linear norm bound (1.17) is very restrictive. Therefore, **Chapter 5** aims to introduce a class of quadratic LK functionals that provides less restrictive robustness bounds. This goal will be achieved by replacing the template of the Lyapunov equation from ii by the template of an algebraic Riccati equation and by using methods from the realm of absolute stability. **Chapter 6** tackles the thus introduced LK functionals of robust type by the numerical approach from Chapter 4, making the proposed functionals numerically traceable and establishing the link to known concepts in the simpler setting of delay-free systems. **Chapter 7** aims at proving the convergence of that numerical approach.

2 Stability Criteria from an Insightful Perspective on the Characteristic Equation

This chapter is devoted to the analysis of stability in the nominal linear system (1.4). As a first step, a framework of three possible perspectives on a two-variable formulation of the characteristic equation is proposed. One of these perspectives, perspective ($\mathrm{P_Z}$), gives rise to the known frequency-sweeping test (1.11). In contrast, the particularly insightful criterion mentioned in (1.13) is related to perspective ($\mathrm{P_S}$). Frequency-sweeping is not only associated with a necessary and sufficient delay-independent criterion but also with a formula for the critical delay h_c if stability holds for $h \in [0, h_c)$. Thus, the chapter aims to develop an analogous theory based on the simpler perspective ($\mathrm{P_S}$).

The chapter is organized as follows. After the introduction of the framework of possible perspectives in **Section 2.1**, the problem of determining the first critical delay h_c is tackled in **Section 2.2** by means of a constrained minimization problem. Then the new necessary and sufficient criterion for delay-independent exponential stability is derived in **Section 2.3**. It shows that the required extension of (1.13) only has to incorporate special cases of $\max_{\varphi\in[0,\pi]} \alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1) = 0$. To this end, the following non-trivial consequence of delay-independent exponential stability must be proven: Eigenvalues of $M(\varphi) := A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1$, as φ increases, cannot move between the left and right complex half-plane only by tunneling through the origin. The section closes with some corollaries, including Mori's criterion that has already been encountered in (1.12). Finally, the main points of the chapter are listed in **Section 2.4**.

The contributions of this chapter are prepublished in

[S1] Scholl, T. H.; Gröll, L.: Stability criteria for time-delay systems from an insightful perspective on the characteristic equation. IEEE Transactions on Automatic Control 68 (2023) 4, 2352–2359.

2.1 Considerations Based on a Two-Variable Formulation

Consider a reformulation [104] of the characteristic equation (1.6) with two variables s and $z := \mathrm{e}^{-sh}$

$$\det(sI_n - A_0 - zA_1) = 0. \tag{2.1}$$

Section 2.1.1 introduces various forbidden sets $\mathcal{S}_{\overline{\mathbb{C}^+}}, \mathcal{S}_{\mathrm{i}\mathbb{R}}, \mathcal{S}_0, \mathcal{S}_{\mathrm{str}}$ for s and z to describe some required preliminaries in a uniform manner. Relying on that foundation, possible perspectives on (2.1) are discussed in Section 2.1.2.

2.1.1 Forbidden Sets*

Delay-independent exponential stability holds if and only if for all delays $h \geq 0$ no characteristic root s from the characteristic equation (1.6) occurs in $\overline{\mathbb{C}^+}$, and thus if and only if no (s, z) from the forbidden set

$$\mathcal{S}_{\overline{\mathbb{C}^+}} := \left\{ (s,z) \in \mathbb{C}^2 \colon s \in \overline{\mathbb{C}^+}, z = \mathrm{e}^{-sh}, h \in \mathbb{R}_{\geq 0} \right\} \tag{2.2}$$

satisfies (2.1). Assume for a given delay $h \geq 0$ there is indeed no characteristic root s in $\overline{\mathbb{C}^+}$. If h is varied, the occurrence of a characteristic root in $\mathbb{C}^+$ must be preceded by a crossing of the imaginary axis, i.e., $s = \mathrm{i}\omega, \omega \in \mathbb{R}$, at some $h = h_c$

* The author has prepublished Section 2.1.1 in [S1], ©2023 IEEE.

(continuous dependence of the real part of the rightmost characteristic root on changes in h is proven in [49]). Hence, already a forbidden set

$$\mathcal{S}_{\mathrm{i}\mathbb{R}} := \left\{(s,z) \in \mathbb{C}^2 \colon s = \mathrm{i}\omega, z = \mathrm{e}^{-\mathrm{i}\omega h_c}, \omega \in \mathbb{R}, h_c \in \mathbb{R}_{\geq 0}\right\} \tag{2.3}$$

is decisive: The zero equilibrium of (1.4) is delay-independently exponentially stable if and only if no (s,z) from the forbidden set $\mathcal{S}_{\mathrm{i}\mathbb{R}}$ satisfies (2.1), provided exponential stability is proven for an arbitrary explicit delay, e.g., for zero delay. For zero delay, the system equation (1.4) simplifies to

$$\dot{x}(t) = (A_0 + A_1)x(t), \tag{2.4}$$

and therefore $A_0 + A_1$ is required to be Hurwitz. Furthermore, the range of ω considered in (2.3) can be restricted to $\omega \in \mathbb{R} \setminus \{0\}$ since

a) $s = \mathrm{i}\omega = 0$ with $h_c < \infty$ (and thus $z = 1$) would be a root for all $h \geq 0$ in (1.6) and contradicts $A_0 + A_1$ Hurwitz;

b) $s \to \mathrm{i}\omega = 0$ as $h \to \infty$ is beyond the definition of delay-independent exponential stability which considers only finite delays.

That is why a new parameter $\varphi \in \mathbb{R}$ can be introduced that substitutes the free parameter $h_c \geq 0$ by

$$h_c = -\frac{\varphi}{\omega}, \quad \omega \neq 0, \tag{2.5}$$

which will be decisive in Section 2.2. Hence, the variable z in (2.3),

$$z = \mathrm{e}^{-\mathrm{i}\omega h_c} = \mathrm{e}^{\mathrm{i}\varphi}, \quad \varphi \in \mathbb{R}, \tag{2.6}$$

becomes an arbitrary complex number with $|z| = 1$. This decoupling of s and z leads to the following two-variable criterion by Hale, Infante, and Tsen [81].

Theorem 2.1.1 (Two-variable criterion [81, Thm. 2.4]). *The zero equilibrium of (1.4) is delay-independently exponentially stable if and only if $A_0 + A_1$ is Hurwitz and no (s, z) from the forbidden set*

$$\mathcal{S}_0 := \{(s, z) \in \mathbb{C}^2 \colon s = \mathrm{i}\omega, |z| = 1, \omega \in \mathbb{R} \setminus \{0\}\} \tag{2.7}$$

satisfies (2.1).

Note that $s = \mathrm{i}\omega = 0$ is not part of the forbidden set $\mathcal{S}_0$. If $A_0 + A_1$ is Hurwitz and if even no (s, z) from the larger set

$$\mathcal{S}_{\mathrm{str}} := \{(s, z) \in \mathbb{C}^2 \colon s = \mathrm{i}\omega, |z| = 1, \omega \in \mathbb{R}\} \tag{2.8}$$

satisfies (2.1), the zero equilibrium is referred to as *strongly* delay-independently exponentially stable [21]. This stability term has no underlying definition in terms of the solutions $x(t)$, but it merely addresses what is covered by considering instead of (2.7) the coarser but simpler forbidden set (2.8). The latter amounts to the classical two-variable criterion by Kamen [106, 104].

The equilibrium might be delay-independently exponentially stable without being strongly delay-independently exponentially stable. In other words, $s = \mathrm{i}\omega = 0$ combined with $|z| = 1$, which does not hamper delay-independent exponential stability, must indeed be excluded in the forbidden set (2.7). This fact has initially been overseen in the literature, but already the following scalar system provides such an example [106].

Example 2.1.2 (Non-strong delay-independent exponential stability). *In*

$$\dot{x}(t) = -x(t) - x(t - h), \tag{2.9}$$

the zero equilibrium is, despite of being delay-independently exponentially stable, not strongly delay-independently exponentially stable. The decisive element turns out to be $(s_0, z_0) = (0, -1)$, which satisfies (2.1) with $s + 1 + z = 0$. It hampers strong delay-independent exponential stability by $(s_0, z_0) \in \mathcal{S}_{\mathrm{str}}$ from (2.8), while not belonging to the forbidden set $\mathcal{S}_0$ for delay-independent exponential stability

from (2.7). Indeed, although $s_0 = 0$ *is an element of the imaginary axis and* $|z_0| = 1$*, it holds* $(s_0, z_0) \notin \mathcal{S}_{\mathrm{i}\mathbb{R}}$ *in (2.3) since there is no finite* h_c *such that* $\mathrm{e}^{-\mathrm{i}\omega h_c} = z_0$ *when* $\omega = 0$*.*

2.1.2 A Framework of Three Perspectives*

In order to determine whether any forbidden $(s, z) \in \mathcal{S}_0$ satisfies (2.1), i.e., satisfies $\det(sI_n - A_0 - zA_1) = 0$, several perspectives on this two-variable formulation in terms of classical eigenvalue problems are appropriate.

($\mathrm{P_S}$) For any given z, the variable s in (2.1) can be seen as an eigenvalue of the matrix $A_0 + zA_1$, cf. [105]. In $\mathcal{S}_0$, or whenever (2.6) is used, values $z = \mathrm{e}^{\mathrm{i}\varphi}$ are of interest. This motivates to define

$$\begin{aligned} s_k(\varphi) &:= \lambda_k(A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1), \\ &\varphi \in \mathbb{R}, \quad k \in \{1, \ldots, n\}. \end{aligned} \tag{2.10}$$

Theorem 2.1.1 requires that no $(s_k(\varphi), \mathrm{e}^{\mathrm{i}\varphi})$ belongs to $\mathcal{S}_0$, i.e., $s_k(\varphi)$ must satisfy

$$\nexists(\varphi, k) : \mathrm{Re}(s_k(\varphi)) = 0 \text{ with } \mathrm{Im}(s_k(\varphi)) \neq 0. \tag{2.11}$$

($\mathrm{P_Z}$) For any given s, the variable z in (2.1) can be seen as an eigenvalue of the matrix pencil $(sI_n - A_0,\ A_1)$, cf. [40]. In $\mathcal{S}_0$, values $s = \mathrm{i}\omega \neq 0$ are of interest. This motivates to define[1]

$$\begin{aligned} z_k(\omega) &:= \lambda_k(\mathrm{i}\omega I_n - A_0,\ A_1), \\ &\omega \in \mathbb{R} \setminus \{0\}, \quad k \in \{1, \ldots, n\}. \end{aligned} \tag{2.12}$$

Theorem 2.1.1 requires $(\mathrm{i}\omega, z_k(\omega)) \notin \mathcal{S}_0$, i.e., $z_k(\omega)$ must satisfy

$$\nexists(\omega, k) : |z_k(\omega)| = 1. \tag{2.13}$$

Alternatively, $\frac{1}{z}$ is seen as an eigenvalue of the dual pencil $(A_1,\ sI_n - A_0)$. If A_0 is Hurwitz, $sI_n - A_0$ in the latter is invertible for $s = \mathrm{i}\omega$, and thus (2.12) also results from

$$\frac{1}{z_k(\omega)} = \lambda_k\Big((\mathrm{i}\omega I_n - A_0)^{-1} A_1\Big) \tag{2.14}$$

(provided k in (2.12) and (2.14) is chosen correspondingly).

($\mathrm{P_{SZ}}$) Both s and z can be seen as eigenvalues, cf. [100]. To this end,

$$(sI_n - A_0 - zA_1)\, v = 0_{n\times 1}, \quad s, z \in \mathbb{C}, v \in \mathbb{C}^n \tag{2.15}$$

must be complemented by a second equation, incorporating that $\overline{s} = \overline{\mathrm{i}\omega} = -s$ and $\overline{z} = \overline{\mathrm{e}^{-\mathrm{i}\varphi}} = \frac{1}{z}$ are characteristic properties in $\mathcal{S}_0$, cf. [87, 44]. The conjugate complex of (2.15) with $w := \overline{v}$ serves this purpose. A quadratic two-parameter eigenvalue problem in s and z emerges [100]

$$(A_0 - sI_n + zA_1\qquad\quad)\, v = 0_{n\times 1} \tag{2.16a}$$
$$(A_1\qquad\quad + zA_0 + szI_n)\, w = 0_{n\times 1} \tag{2.16b}$$

with solution tuples (s_k, z_k), $k \in \{1, \ldots, 2n^2\}$. Theorem 2.1.1 requires that $(s_k, z_k) \notin \mathcal{S}_0$, i.e., s_k and z_k must satisfy

$$\nexists k : \mathrm{Re}(s_k) = 0, \mathrm{Im}(s_k) \neq 0, \text{ and } |z_k| = 1. \tag{2.17}$$

If elements (s, z) in the forbidden set $\mathcal{S}_0$ have been identified, no matter by which perspective, stability is only delay-dependent. Then, based on the corresponding values of φ in $z = \mathrm{e}^{\mathrm{i}\varphi}$ and ω in $s = \mathrm{i}\omega$, critical delays at which roots on the imaginary axis occur can be concluded from (2.5), cf. [40, 44, 41]. Provided

---

* The author has prepublished Section 2.1.2 in [S1], ©2023 IEEE.

1 Some of the n eigenvalues can be infinite. Since their number n_∞ (algebraic multiplicity) equals the dimension of the nilpotent matrix N in Weierstrass' canonical form with $\mathrm{rk}(N) \geq 0$, it holds $n_\infty \geq n - \mathrm{rk}(A_1)$.

exponential stability holds for the system with zero delay (2.4), the initial exponential stability gets lost at the smallest of these critical delays, which thus bounds the exponentially stable initial delay interval $[0, h_c)$.

2.2 Delay-Dependent Stability

The following section is divided into two parts. First, the approach to determine h_c is explained in Section 2.2.1. Then, some examples are given in Section 2.2.2.

2.2.1 Determination of the First Critical Delay*

The main result of this section, Theorem 2.2.2, gives a constrained minimization problem based on perspective ($\mathrm{P_S}$) for the first critical delay h_c. The result is complementary to the ($\mathrm{P_Z}$)-based frequency sweeping approach [76, Thm. 2.2], which is reformulated into a constrained optimization form in the following theorem. Note that (2.18) describes the minimum[2] positive value of (2.5) for (ω, k)-pairs in ($\mathrm{P_Z}$) that hamper (2.13).

Theorem 2.2.1 (Delay interval of exponential stability by ($\mathrm{P_Z}$), cf. [76, Thm. 2.2]). *Based on (2.12) define*

$$h_c := \inf_{(\omega,k)\in(0,\infty)\times\{1,\dots,n\}} \left(-\frac{\arg^- z_k(\omega)}{\omega} \right) \quad \text{subject to} \quad |z_k(\omega)| = 1 \tag{2.18}$$

with $\arg^- z := \varphi \in (-2\pi, 0]$ *such that* $z = |z|\mathrm{e}^{\mathrm{i}\varphi}$. *If* $A_0 + A_1$ *is Hurwitz and* $h_c < \infty$, *then the equilibrium of (1.4) is exponentially stable for* $h \in [0, h_c)$ *and not exponentially stable at* $h = h_c$.

* The author has prepublished Section 2.2.1 in [S1], ©2023 IEEE.

2 In (2.18) and (2.20), the infimum is only required to cope with an empty constraint set and can otherwise be replaced by a minimum operator.

Similarly, the proposed approach will describe the minimum positive value of (2.5) for (φ, k)-pairs in $(\mathrm{P_S})$ that hamper (2.11).

A disadvantage of (2.18) is that $z_k(\omega)$ is considered on an unbounded set $\omega \in (0, \infty)$. In contrast, for perspective $(\mathrm{P_S})$, the evaluation of eigenvalues $s_k(\varphi)$, (2.10), can be restricted to the bounded set $\varphi \in [0, \pi]$. Values of $s_k(\varphi)$ on the whole domain $\varphi \in (-\infty, \infty)$ are still needed to gain all critical delays from (2.5), but they can be reconstructed from those on $[0, \pi]$ due to symmetry.

Note that the objective function in (2.20) below is nothing more than

$$h(\varphi, k) := \begin{cases} -\frac{\varphi}{\operatorname{Im} s_k(\varphi)} & \text{if } \operatorname{Im} s_k(\varphi) < 0, \\ \infty & \text{if } \operatorname{Im} s_k(\varphi) = 0, \\ \frac{2\pi - \varphi}{\operatorname{Im} s_k(\varphi)} & \text{if } \operatorname{Im} s_k(\varphi) > 0, \end{cases} \tag{2.19}$$

which, for the sake of compactness, is written with the modulo operation, i.e., $(\pm\varphi \mod 2\pi) \in [0, 2\pi)$.

Theorem 2.2.2 (Delay interval of exponential stability). *Based on the eigenvalues $s_k(\varphi) := \lambda_k(A_0 + \mathrm{e}^{\mathrm{i}\varphi} A_1)$ define*

$$h_c := \inf_{(\varphi,k) \in (0,\pi] \times \{1,\dots,n\}} \frac{\operatorname{sgn}\big(-\operatorname{Im} s_k(\varphi)\big)\varphi \mod 2\pi}{|\operatorname{Im} s_k(\varphi)|} \quad \text{subject to} \quad \operatorname{Re} s_k(\varphi) = 0 \tag{2.20}$$

with $\inf \emptyset = \infty$ and $\frac{\operatorname{sgn}(0)\varphi \mod 2\pi}{0} := \infty$. If $A_0 + A_1$ is Hurwitz and $h_c < \infty$, then the zero equilibrium of (1.4) is exponentially stable for $h \in [0, h_c)$ and not exponentially stable at $h = h_c$. The zero equilibrium of (1.4) is delay-independently exponentially stable if and only if $A_0 + A_1$ is Hurwitz and $h_c = \infty$.

Proof. Due to (2.4), $A_0 + A_1$ being Hurwitz is necessary. As defined in (2.5), critical delays can be expressed by $h_c = -\frac{\varphi}{\omega}$, where $\omega = \operatorname{Im} s_k(\varphi)$ in perspective $(\mathrm{P_S})$. According to Section 2.1, critical delays occur for any $(\varphi, k) \in \mathbb{R} \times$

$\{1,\dots,n\}$ with $\operatorname{Re} s_k(\varphi) = 0$ and $\operatorname{Im} s_k(\varphi) \neq 0$. Hence, the minimum positive critical delay is

$$\begin{aligned} h_c = \inf h_0(\varphi,k) \quad & \text{with } h_0(\varphi,k) := -\frac{\varphi}{\operatorname{Im} s_k(\varphi)} \\ \text{subject to} \quad & \operatorname{Re} s_k(\varphi) = 0 \text{ and } \operatorname{Im} s_k(\varphi) \neq 0 \\ & h_0(\varphi,k) > 0 \\ & \varphi \in (-\infty,\infty) \text{ and } k \in \{1,\dots,n\}. \end{aligned}$$

The requirement $\operatorname{Im} s_k(\varphi) \neq 0$ can be dropped since the corresponding objective function value, which is ∞ by definition, can only be optimal if there is no other element in the constraint set, while an empty constraint set yields the same result $h_c = \inf \emptyset = \infty$. It has to be shown that only $\varphi \in (0,\pi]$ instead of $\varphi \in (-\infty,\infty)$ is relevant. Only $\varphi \geq 0$ must be considered because

$$s_k(-\varphi) = \lambda_k\Big(\overline{A_0 + \mathrm{e}^{\mathrm{i}\varphi} A_1}\Big) = \overline{\lambda_{\tilde{k}}(A_0 + \mathrm{e}^{\mathrm{i}\varphi} A_1)} = \overline{s_{\tilde{k}}(\varphi)} \tag{2.21}$$

holds for some $\tilde{k} \in \{1,\dots,n\}$, and thus $h_0(-\varphi,k) = h_0(\varphi,\tilde{k})$. Furthermore, only $\varphi \in [0,2\pi)$ can lead to an optimum since $s_k(\varphi + 2l\pi) = s_k(\varphi)$, $l \in \mathbb{N}$, implies $h_0(\varphi + 2l\pi,k) > h_0(\varphi,k)$. Additionally, $\varphi \neq 0$ since $A_0 + A_1$ Hurwitz implies $\operatorname{Re} s_k(0) \neq 0$. Hence, in a first step, only $\varphi \in (0,2\pi)$ is relevant. Positivity $h_0(\varphi,k) > 0$ is achieved if and only if $\operatorname{Im} s_k(\varphi) < 0$, which gives the first case in (2.19) and allows to drop $h_0(\varphi,k) > 0$ from the constraints. In a second step, the domain can be restricted to $(0,\pi]$ by considering for any $\varphi \in [\pi,2\pi)$ the corresponding $\tilde{\varphi} \in (0,\pi]$ with $\varphi = 2\pi - \tilde{\varphi}$. Since $s_k(\varphi) = \overline{s_{\tilde{k}}(\tilde{\varphi})}$, it holds $h_0(\varphi,k) = -\frac{\varphi}{\operatorname{Im} s_k(\varphi)} = -\frac{2\pi-\tilde{\varphi}}{(-\operatorname{Im} s_{\tilde{k}}(\tilde{\varphi}))}$, which gives the third case in (2.19) for $\operatorname{Im} s_{\tilde{k}}(\tilde{\varphi}) > 0$. □

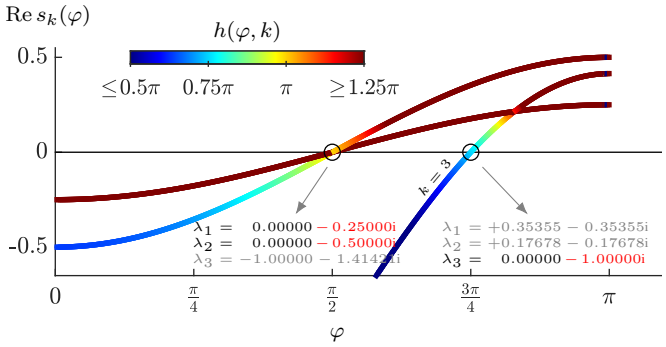


Figure 2.1[*]**:** Example 2.2.3. Zeros of $\operatorname{Re} s_k(\varphi) = \operatorname{Re} \lambda_k(A_0 + \mathrm{e}^{\mathrm{i}\varphi} A_1)$, $k \in \{1, 2, 3\}$, form the constraint set in Theorem 2.2.2 (circles). Among these (φ, k) pairs, the smallest value of $h(\varphi, k) = \big(\operatorname{sgn}\big(-\operatorname{Im} s_k(\varphi)\big)\, \varphi \mod 2\pi\big) \,/\, (|\operatorname{Im} s_k(\varphi)|)$, see (2.19), is attained by $h_c = h(\frac{3\pi}{4}, 3) = \frac{3}{4}\pi$, which is indicated by the color bar. Thus, stability holds for delays $h \in [0, \frac{3}{4}\pi)$.

2.2.2 Examples on the First Critical Delay*

The constraint set in (2.20)

$$C := \{(\varphi, k) : \operatorname{Re} s_k(\varphi) = 0\} \tag{2.22}$$

contains zeros of the real parts from all n eigenvalue functions $\varphi \mapsto \operatorname{Re} s_k(\varphi)$, $k \in \{1 \ldots, n\}$. The next example demonstrates that a restriction to zeros of the spectral abscissa function $\varphi \mapsto \max_k \operatorname{Re} s_k(\varphi)$ is indeed not possible. Moreover, Figure 2.1 provides a simple graphical evaluation of Theorem 2.2.2.

Example 2.2.3 (Relevance of all eigenvalues in Theorem 2.2.2). *Consider*

$$\dot{x}(t) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & -1 \end{bmatrix} x(t) + \begin{bmatrix} -\frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -\sqrt{2} \end{bmatrix} x(t-h).$$

[*] The author has prepublished Section 2.2.2 in [S1], ©2023 IEEE.
[*] The author has prepublished Figure 2.1 in [S1], ©2023 IEEE.

The eigenvalues required in Theorem 2.2.2 are $s_1(\varphi) = -\frac{1}{4}\mathrm{e}^{\mathrm{i}\varphi}$, $s_2(\varphi) = -\frac{1}{2}\mathrm{e}^{\mathrm{i}\varphi}$, $s_3(\varphi) = -1 - \sqrt{2}\mathrm{e}^{\mathrm{i}\varphi}$. *Figure 2.1 visualizes that the constraint set consists of three elements* $C = \{(\frac{\pi}{2}, 1), (\frac{\pi}{2}, 2), (\frac{3\pi}{4}, 3)\}$. *At these* (φ, k)*-pairs, the objective function of (2.20), see the color bar, takes the values* $\{2\pi, \pi, \frac{3}{4}\pi\} =: H$. *Hence,* $h_c = \min H = \frac{3}{4}\pi$.

Remark 2.2.4 (Numerical determination of φ). *The graphical evaluation of the optimization problem already provides a rough result for the critical delay with little effort. To evaluate the optimum (2.20) precisely, the values of* φ *in (2.22) can, e.g., be determined as minima of* $\varphi \mapsto \min_{k\in\{1,\dots,n\}} |\operatorname{Re} s_k(\varphi)|$. *To this end, local minima from a pointwise evaluation can be refined by* `fminbnd` *in Matlab with tightened tolerances. Alternatively, the values of* φ *or* $2\pi - \varphi$ *are derived by perspective (*$\mathrm{P_{SZ}}$*) as arguments of* $z = \mathrm{e}^{\mathrm{i}\varphi}$ *if a routine like* `quad_twopareig` *[155, 146] is available to compute* (s_k, z_k). *The proof of [41, Thm. 3.1] reveals that another way is to search for valid values of* z *in the spectrum of a* $2n^2$*-dimensional matrix pencil* $\sigma\left(\begin{bmatrix} 0_{n^2\times n^2} & I_{n^2} \\ -I_n\otimes A_1^\top & -A_0\oplus A_0^\top \end{bmatrix}, \begin{bmatrix} I_{n^2} & 0_{n^2\times n^2} \\ 0_{n^2\times n^2} & A_1\otimes I_n \end{bmatrix}\right)$, *where* $\otimes, \oplus$ *denote Kronecker product and sum.*

The system in the following example is a frequently used benchmark in the context of LMI-based criteria [166, 79].

Example 2.2.5 (Analytically determined h_c). *The zero equilibrium of*

$$\dot{x}(t) = \begin{bmatrix} -2 & 0 \\ 0 & -0.9 \end{bmatrix} x(t) + \begin{bmatrix} -1 & 0 \\ -1 & -1 \end{bmatrix} x(t-h)$$

is exponentially stable for $h < h_c := \frac{\arccos(-0.9)}{\sqrt{1-0.9^2}} \approx 6.17258$ *and not exponentially stable at* $h = h_c$.

Reasoning: $A_0 + A_1$ *is Hurwitz as required in Theorem 2.2.2. The eigenvalues* $s_k(\varphi) := \lambda_k(M(\varphi))$ *of the triangular matrix* $M(\varphi) := A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1$ *are obvious. Since* $\operatorname{Re} s_2(\varphi) = -2 - \cos\varphi \neq 0$ *and* $\operatorname{Re} s_1(\varphi) = -0.9 - \cos\varphi$, *the constraint set (2.22) consists of the single point* $(\varphi_c, k_c) := (\arccos(-0.9), 1)$ *with*

$s_{k_c}(\varphi_c) = -\mathrm{i}\sin\varphi_c = -\mathrm{i}\sqrt{1-0.9^2}$. *Since* $\mathrm{Im}(s_{k_c}(\varphi_c)) < 0$, *Theorem 2.2.2 yields* $h_c = \frac{\varphi_c}{|\,\mathrm{Im}\, s_{k_c}(\varphi_c)|} = \frac{\arccos(-0.9)}{\sqrt{1-0.9^2}}$.

In fact, the triangular structure of Example 2.2.5 allows the stability analysis to be reduced to an analysis of scalar systems (so-called Hayes equations [85]). The following example shows that perspective ($\mathrm{P_S}$) is particularly insightful in this case. The results are well known [82, p. 135 / Fig. 5.1], but the usual derivations are quite laborious.

Example 2.2.6 (The scalar case). *Consider (1.4) with* $n = 1$, *i.e.,* $\dot{x}(t) = a_0x(t) + a_1x(t-h)$, $a_0, a_1 \in \mathbb{R}$. *The Hurwitz condition*[3]

$$a_0 + a_1 < 0 \tag{2.23}$$

forms the non-red open region in Figure 2.2. Theorem 2.2.2 only depends on $s_1(\varphi) = a_0 + \mathrm{e}^{\mathrm{i}\varphi}a_1$ *for* $\varphi \in [0, \pi]$. *In the complex plane, the latter can be visualized as a* $180°$ *rotation around* a_0 *from* $a_0 + a_1$ *to* $a_0 - a_1$, *see the gray semicircle in Figure 2.2 with* $(a_0, a_1) = (-3, 0.5)$. *Depending on the end point* $a_0 - a_1$, *three cases have to be distinguished:*

a) For (2.23) combined with $a_0 - a_1 < 0$, *i.e.,* $a_0 + |a_1| < 0$, *the constraint set in (2.20) is empty, and thus* $h_c = \infty$ *(dark-green triangle in the stability chart, strong delay-independent exponential stability).*

b) For (2.23) combined with $a_0 - a_1 = 0$, *i.e.,* $a_0 = a_1 < 0$, *the denominator of the objective function in (2.20) is zero, and thus* $h_c = \infty$ *(white dashed line in the stability chart, non-strong delay-independent exponential stability).*

c) For (2.23) combined with $a_0 - a_1 > 0$, *consider in Figure 2.2 the blue and ochre arcs. Theorem 2.2.2 yields* $h_c = \frac{\varphi_c}{|\omega(\varphi_c)|}$ *as the quotient of the*

[3] According to [46], exponential stability cannot be gained by increasing the delay in the scalar system. Hence, $a_0 + a_1 < 0$, which ensures that exponential stability holds for $h = 0$, is not only necessary for exponential stability $\forall h \geq 0$ or $\forall h \in [0, h_c)$, but even necessary for exponential stability at some $h \in \mathbb{R}_{\geq 0}$.

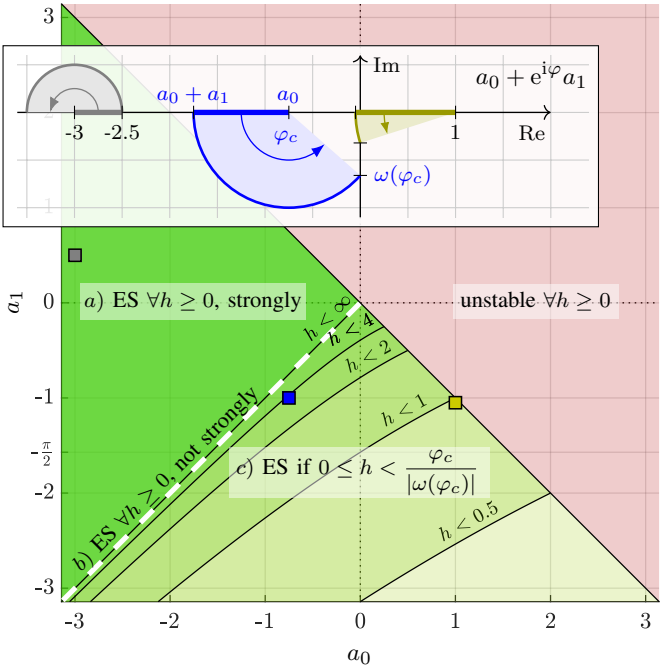


Figure 2.2[*]: Example 2.2.6. Stability chart for scalar systems $\dot{x}(t) = a_0 x(t) + a_1 x(t-h)$ indicating for which parameter combinations exponential stability (ES) holds. The white box above gives for highlighted points an evaluation of $\varphi \mapsto a_0 + e^{i\varphi} a_1$ in the complex plane. The gray point $(a_0, a_1) = (-3, 0.5)$ provides delay-independent exponential stability, whereas $(-0.75, -1)$ in blue gives exponential stability for $h < 3.66$ and $(1, -1.05)$ in ochre gives exponential stability for $h < 0.97$.

[*] The author has prepublished Figure 2.2 in [S1], ©2023 IEEE.

rotation angle $\varphi_c \in (0, \pi)$ at which the imaginary axis is met and the corresponding section of the imaginary axis $|\omega(\varphi_c)|$. That is

$$h_c = \frac{\varphi_c}{|a_1| \sin \varphi_c} \quad \text{with } \varphi_c = \arccos(-\tfrac{a_0}{a_1}) \qquad (2.24)$$
$$= \frac{\arccos(-\frac{a_0}{a_1})}{\sqrt{a_1^2 - a_0^2}}.$$

Moreover, the following should be noted.

**) If (2.23) is satisfied, exponential stability holds at least for sufficiently small delays (differently shaded green areas in Figure 2.2). Given a fixed delay $h > 0$, the exponentially stable region in the (a_0, a_1)-parameter plane is derived by solving $h < h_c$ in (2.24) for $|a_1| = -a_1 < \frac{\varphi_c}{h \sin \varphi_c}$ and $a_0 = -a_1 \cos(\varphi_c) < \frac{\varphi_c}{h} \cot(\varphi_c)$. Its boundary $\{[a_0, a_1]^\top = \frac{\varphi_c}{h \sin \varphi_c}[\cos \varphi_c, -1]^\top : \varphi_c \in (0, \pi)\}$ for $h = h_c \in \{0.5, 1, 2, 4\}$ is shown as boundary of the $(h < h_c)$ regions in Figure 2.2.*

***) Provided $a_0 < 0$, the critical delay is always larger than the critical delay with $a_0 = 0$ since the blue arc reveals that a larger angle $\varphi_c > \frac{\pi}{2}$ is combined with a smaller $|\omega(\varphi_c)| < |a_1|$. Hence, $h_c > \frac{\pi}{2|a_1|}$ if $a_0 < 0$ and $a_1 < 0$.*

2.3 Delay-Independent Stability

The section is structured as follows: Section 2.3.1 introduces the criterion for delay-independent exponential stability, which is then applied to some examples in Section 2.3.2. Moreover, Section 2.3.3 derives some corollaries.

2.3.1 A Necessary and Sufficient Criterion*

The main result in this section, stated in Theorem 2.3.4, provides a necessary and sufficient delay-independent stability criterion based on perspective ($\mathrm{P_S}$). It aims to be complementary to the following well-known ($\mathrm{P_Z}$)-based criterion[4], where (2.25) ensures that $z_k(\omega)$ in (2.14) satisfies (2.13).

Theorem 2.3.1 (Frequency sweeping [76, Thm. 2.1]). *The zero equilibrium of (1.4) is delay-independently exponentially stable if and only if A_0 and $A_0 + A_1$ are Hurwitz and*

$$\forall \omega > 0: \quad \rho\big((\mathrm{i}\omega I_n - A_0)^{-1} A_1\big) < 1. \tag{2.25}$$

A graphical evaluation of the spectral radius (2.25) over $\omega \in (0, \infty)$ is proposed in [76, 42]. Similarly, the present section is concerned with an evaluation of the spectral abscissa

$$\alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi} A_1) = \max_{k \in \{1,\dots,n\}} \operatorname{Re} s_k(\varphi) \tag{2.26}$$

over $\varphi \in [0, \pi]$ to ensure that $s_k(\varphi)$ in (2.10) satisfies (2.11).

According to Theorem 2.2.2, there are two scenarios that lead to delay-independent exponential stability:

(i) $A_0 + A_1$ is Hurwitz and the constraint set of (2.20) is empty, i.e., for all $\varphi \in [0, \pi]$ no eigenvalue $s_k(\varphi)$ occurs on the imaginary axis;

* The author has prepublished Section 2.3.1 in [S1], ©2023 IEEE.

4 The spectral radius $\rho\big((\mathrm{i}\omega I_n - A_0)^{-1} A_1\big)$ can also be recognized as the structured singular value of $G_{ux}(s) = (sI - A_0)^{-1} A_1$, addressing $\dot{x}(t) = A_0 x(t) + A_1 u(t)$, to which the diagonally structured $G_{xu}(s) = \mathrm{e}^{-sh} I_n$ is applied in closed loop [42].

(ii) $A_0 + A_1$ is Hurwitz and there is no other denominator in (2.20) than zero, i.e., eigenvalues on the imaginary axis occur for some (φ, k), but these are exclusively located at the origin.

Case (i) describes strong delay-independent exponential stability (2.8). It will be addressed by $\alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1) < 0$ for all $\varphi \in [0, \pi]$, cf. (1.13). Case (ii) is the special case of non-strong delay-independent exponential stability. Obviously, it can occur with $\max_{\varphi\in[0,\pi]} \alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1) = 0$, which becomes visible in Example 2.2.6 b). Whether case (ii) can also be accompanied by $\max_{\varphi\in[0,\pi]} \alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1) > 0$ is not that obvious. Starting in $\mathbb{C}^-$ for $\varphi = 0$, the n eigenvalues $s_k(\varphi)$, $k \in \{1, \ldots, n\}$, move continuously in the complex plane as φ increases. Case (i) bans eigenvalues from the imaginary axis, and thus they cannot reach the right half-plane. However, case (ii) describes a gap in the imaginary axis: the occurrence of eigenvalues at the origin does not hamper delay-independent exponential stability. Thus, the question arises whether eigenvalues can move from the left half-plane to the right half-plane only by tunneling through the origin. This question is motivated further in the following remark.

Remark 2.3.2 (Crossing of the origin). *In contrast to the roots*[5] *of (1.6), the eigenvalues $s_k(\varphi)$ in (2.10) can move from $\mathbb{C}^-$ through the origin to $\mathbb{C}^+$ as φ increases. Such an example is provided in Figure 2.3a, which shows for $A_0 = \left[\begin{smallmatrix} -1 & 1 \\ -1 & -1 \end{smallmatrix}\right]$, $A_1 = \sqrt{2}\left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$ the union of eigenvalue paths*[6] *$\{s_k(\varphi) : \varphi \in \mathbb{R}, k \in \{1, 2\}\}$. However, since the movement back to $\mathbb{C}^-$ in this example is not through the origin, $h_c = \infty$ does not result as a minimum in Theorem 2.2.2. In Figure 2.3b there is indeed no other crossing point of the imaginary axis than the origin. Thus, Theorem 2.2.2 would lead to the conclusion of delay-independent exponential stability, provided $A_0 + A_1$ was Hurwitz. Figure 2.3b can, e.g.,*

[5] Because of case a) in Section 2.1, roots s of the characteristic quasipolynomial (1.6) cannot move through the origin in the complex plane as h increases. Note that the only relation between $M(\varphi) = A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1$ and the delay equation (1.4) is that non-zero purely imaginary eigenvalues $s_k(\varphi) = \mathrm{i}\omega \neq 0$ of $M(\varphi)$ at some φ coincide with non-zero purely imaginary roots s of (1.6) at some h.

[6] An evaluation of the union of eigenvalue paths is proposed in [139], where, however, zero crossings in the manner of Figure 2.3b are not taken into account.

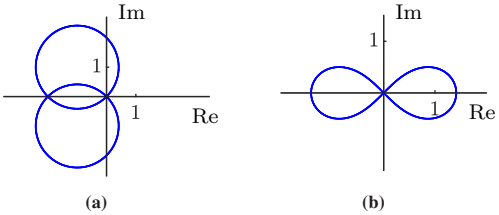


Figure 2.3[*]**:** As discussed in Remark 2.3.2, eigenvalues $s_k(\varphi)$ can cross the origin.

be realized by $A_0 = \left[\begin{smallmatrix} -1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$, $A_1 = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]$ *or by* $A_0 = \left[\begin{smallmatrix} -1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$, $A_1 = \left[\begin{smallmatrix} 0 & 0 \\ -1 & 0 \end{smallmatrix}\right]$, *but both are not fulfilling* $A_0 + A_1$ *Hurwitz, although the latter example with* $\alpha(A_0 + A_1) = 0$ *is very close.*

The proof of Theorem 2.3.4 will show that case (ii) indeed cannot be accompanied by zero crossings in the manner of Figure 2.3b. Thus, if there is a $\varphi \in [0, \pi]$ with $\alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi} A_1) > 0$, non-delay-independent exponential stability is proven immediately. The proof makes use of the following statement.

Lemma 2.3.3 ([81, Cor. 2.7]). *If Theorem 2.1.1 holds, then* A_0 *is Hurwitz.*

The main result of this section is as follows.

Theorem 2.3.4 (Delay-independent exponential stability). *The zero equilibrium of (1.4) is delay-independently exponentially stable if and only if either*

$$\begin{aligned} &(i) \quad \max_{\varphi\in[0,\pi]} \alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi} A_1) < 0 \\ \textit{or} \quad &(ii) \quad \alpha(A_0 + A_1) < 0, \quad \max_{\varphi\in(0,\pi]} \alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi} A_1) = 0, \\ &\textit{and } \nexists (\varphi_c, k) : \operatorname{Re} s_k(\varphi_c) = 0,\ \operatorname{Im} s_k(\varphi_c) \neq 0, \end{aligned} \tag{2.27}$$

[*] The author has prepublished Figure 2.3 in [S1], ©2023 IEEE.

where $\varphi_c \in \operatorname*{argmax}_{\varphi\in[0,\pi]} \alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1)$ *and* $s_k(\varphi) := \lambda_k(A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1)$, $k \in \{1,\dots,n\}$. *It is strongly delay-independently exponentially stable, see (2.8), if and only if* (i) *holds.*

Proof. According to Theorem 2.2.2, $h_c = \infty$ occurs if and only if either

$$\begin{aligned}
&(i')\ A_0 + A_1 \text{ is Hurwitz and } \nexists(\varphi,k) : \operatorname{Re} s_k(\varphi) = 0 \quad \text{or}\\
&(ii')\ A_0 + A_1 \text{ is Hurwitz, } \exists(\varphi,k) : s_k(\varphi) = 0,\\
&\qquad \text{and } \nexists(\varphi,k) : s_k(\varphi) = \mathrm{i}\omega, \omega \in \mathbb{R}\setminus\{0\}.
\end{aligned}$$

In the following, it will be inferred that (i) and (ii) in Theorem 2.3.4 under consideration are necessary and sufficient for (i') and (ii'). Unless otherwise stated, consider $\varphi \in [0,\pi], k \in \{1,\dots,n\}, s \in \mathbb{C}, \omega \in \mathbb{R}, M(\varphi) := A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1$.

$(i) \Rightarrow (i')$: For $\varphi = 0$, (i) yields $\alpha(M(0)) = \alpha(A_0 + A_1) < 0$, i.e., $A_0 + A_1$ must be Hurwitz. Furthermore, since (i) requires $\max_{\varphi\in[0,\pi]} \alpha(M(\varphi)) = \max_{\varphi\in[0,\pi]} \max_{k\in\{1,\dots,n\}} \operatorname{Re} s_k(\varphi) < 0$, it holds that $\nexists(\varphi,k) : \operatorname{Re} s_k(\varphi) = 0$.

$(i') \Rightarrow (i)$: Continuity of $\varphi \mapsto \alpha(M(\varphi))$ and $\alpha(A_0 + A_1) = \alpha(M(0)) < 0$, while $\nexists\varphi : \alpha(M(\varphi)) = 0$, implies $\alpha(M(\varphi)) < 0, \forall\varphi$.

$(ii) \Leftrightarrow (ii')$: Both, (ii) and (ii') consist of three requirements. For the first one ($A_0 + A_1$ Hurwitz) as well as the third one (nonexistence of non-zero purely imaginary roots) equivalence between (ii) and (ii') is obvious. The second combined with the third one in (ii) implies $\exists(\varphi,k) : s_k(\varphi) = 0$ in (ii'). It remains to show that (ii') implies $\max_{\varphi\in(0,\pi]} \alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1) = 0$ by proving that (ii') implies

$$\nexists(\varphi,k) : s_k(\varphi) \in \mathbb{C}^+, \tag{2.28}$$

see the discussion above and Figure 2.3b. To this end, define $s \mapsto \tilde{z}_k(s)$ with $k \in \{1, \ldots, n\}$ as a mapping of $s \in \mathbb{C}$ to a corresponding z in (2.1), i.e.,

$$\det(sI_n - A_0 - \tilde{z}_k(s)\, A_1) = 0. \tag{2.29}$$

As in (2.14), and because of Lemma 2.3.3, $\tilde{z}_k(s)$ can be determined by

$$\frac{1}{\tilde{z}_k(s)} = \lambda_k\big((sI_n - A_0)^{-1} A_1\big) \text{ for } s \in \overline{\mathbb{C}^+}. \tag{2.30}$$

Consider at each s the largest absolute value of (2.30)

$$\frac{1}{\min\limits_{k\in\{1,\ldots,n\}} |\tilde{z}_k(s)|} = \rho\big((sI_n - A_0)^{-1} A_1\big). \tag{2.31}$$

Due to Lemma 2.3.3, $s \mapsto \tilde{N}(s) := (sI_n - A_0)^{-1} A_1$ is holomorphic on $\mathbb{C}^+$. For holomorphic $\tilde{N}(s)$ the spectral radius $s \mapsto \rho(\tilde{N}(s)) = \max_{k\in\{1,\ldots,n\}} |\lambda_k(\tilde{N}(s))|$ is a subharmonic function [24]. Since the domain $\mathbb{C}^+$ is unbounded, the maximum principle for subharmonic functions only applies if the function is bounded above [89, Thm. A.2.28], which, however, is true because $\lim_{s\to\infty} \rho(\tilde{N}(s)) = 0$. Thus, for $s \in \overline{\mathbb{C}^+}$ the maximum of (2.31) is attained on the imaginary axis $\partial\overline{\mathbb{C}^+} = \mathrm{i}\mathbb{R}$, and, consequently, the minimum of $s \mapsto \min_{k\in\{1,\ldots,n\}} |\tilde{z}_k(s)|$ as well. Hence, unless $s \mapsto \min_{k\in\{1,\ldots,n\}} |\tilde{z}_k(s)|$ is constant on $\mathbb{C}^+$, the strict inequality

$$\forall (s,k) \text{ with } s \in \mathbb{C}^+ : |\tilde{z}_k(s)| > \min_{\omega\in\mathbb{R}} \min_{k\in\{1,\ldots,n\}} |\tilde{z}_k(\mathrm{i}\omega)| \tag{2.32}$$

holds. Consider the right hand side of (2.32). On the one hand, because of (2.31) with $\rho(\tilde{N}(\mathrm{i}\omega)) \to 0$ as $\omega \to \pm\infty$, it holds

$$\min_{k\in\{1,\ldots,n\}} |\tilde{z}_k(\mathrm{i}\omega)| \to \infty \text{ as } \omega \to \pm\infty. \tag{2.33}$$

On the other hand, the second requirement in (ii') yields

$$\begin{aligned} &\exists(\varphi, k) : s_k(\varphi) = 0 \\ \overset{(2.10)}{\Leftrightarrow}\quad &\exists\varphi : \det(-A_0 - \mathrm{e}^{\mathrm{i}\varphi} A_1) = 0 \\ \overset{(2.29)}{\Leftrightarrow}\quad &\exists k : |\tilde{z}_k(0)| = 1 \end{aligned} \tag{2.34}$$

for the point $\omega = 0$. Consequently, by continuity, $s \mapsto \min_{k\in\{1,\ldots,n\}} |\tilde{z}_k(s)|$ is indeed non-constant and (2.32) applies. Furthermore, the third requirement in (ii') gives

$$\begin{aligned} &\nexists(\varphi, k) : s_k(\varphi) = \mathrm{i}\omega, \quad \omega \neq 0 \\ \overset{(2.10)}{\Leftrightarrow}\quad &\nexists\varphi : \det(\mathrm{i}\omega I_n - A_0 - \mathrm{e}^{\mathrm{i}\varphi} A_1) = 0, \quad \omega \neq 0 \\ \overset{(2.29)}{\Leftrightarrow}\quad &\nexists k : |\tilde{z}_k(\mathrm{i}\omega)| = 1 \quad \text{for all } \omega \neq 0. \end{aligned} \tag{2.35}$$

Continuity of $\omega \mapsto \tilde{z}_k(\mathrm{i}\omega)$ combined with the results for $\omega = 0$, $\omega \in \mathbb{R} \setminus \{0\}$, and $\omega \to \pm\infty$, which are obtained in (2.34), (2.35), and (2.33), leads (similar to [81, Lemma 2.5]) to

$$\min_{\omega\in\mathbb{R}} \min_{k\in\{1,\ldots,n\}} |\tilde{z}_k(\mathrm{i}\omega)| = 1$$

for the right hand side in (2.32). Thus, (2.32) implies

$$\begin{aligned} &\nexists(s, k) : |\tilde{z}_k(s)| = 1 \text{ with } s \in \mathbb{C}^+ \\ \overset{(2.29)}{\Leftrightarrow}\quad &\nexists s \in \mathbb{C}^+ : \det(sI_n - A_0 - \mathrm{e}^{\mathrm{i}\varphi} A_1) = 0,\ \varphi \in (-\pi, \pi] \\ \overset{(2.10),(2.21)}{\Leftrightarrow}\quad &\nexists(\varphi, k) : s_k(\varphi) \in \mathbb{C}^+, \end{aligned}$$

which completes the proof of (2.28). □

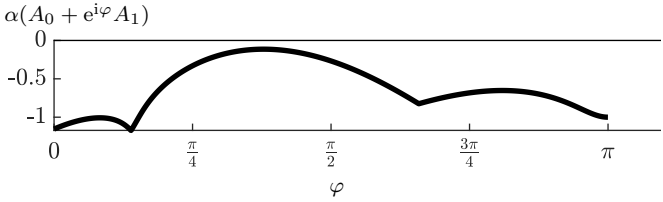


Figure 2.4*: In Example 2.3.5, $\alpha(A_0 + e^{i\varphi}A_1)$ takes only negative values on $\varphi \in [0, \pi]$. Consequently, strong delay-independent exponential stability holds.

2.3.2 Examples on Delay-Independent Stability*

A main advantage of Theorem 2.3.4 (i) is its simple implementation.

Example 2.3.5 (Numerical evaluation of Theorem 2.3.4 (i)). *Theorem 2.3.4* (i) *only requires an evaluation of the function* $\varphi \mapsto \alpha(A_0 + e^{i\varphi}A_1)$ *over* $\varphi \in [0, \pi]$. *In Matlab, the following lines serve this purpose for an exemplary step size of* φ, *provided the system matrices have been assigned to* `A0` *and* `A1`.

```
P=0:1e-3:pi;
ALPHA=arrayfun(@(x) max(real(eig(A0+exp(1i*x)*A1))), P);
plot(P,ALPHA);
```

Exclusively negative values indicate strong delay-independent exponential stability. For $A_0 = \begin{bmatrix} -1 & 2 & 1 \\ -1 & -2 & 0 \\ -1 & 0 & -2 \end{bmatrix}$ *and* $A_1 = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 0 & -2 \\ 1 & 2 & 2 \end{bmatrix}$ *the result is shown in Figure 2.4. Thus, the zero equilibrium of* $\dot{x}(t) = A_0x(t) + A_1x(t-h)$ *is exponentially stable for any delay* $h \geq 0$.

In the following Examples 2.3.6a and 2.3.6b, the maximum of the spectral abscissa function $\varphi \mapsto \alpha(A_0 + e^{i\varphi}A_1)$ is zero. Hence, part (ii) of Theorem 2.3.4 must be considered and the eigenvalues $s_k(\varphi_c), k \in \{1, \ldots, n\}$, at the maximizers φ_c become decisive. Examples 2.3.6a and 2.3.6b lead to identical spectral abscissa

* The author has prepublished Section 2.3.2 in [S1], ©2023 IEEE.
* The author has prepublished Figure 2.4 in [S1], ©2023 IEEE.

functions $\varphi \mapsto \alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1)$, shown as a black line in Figure 2.5. Nevertheless, in Example 2.3.6a, the zero equilibrium is delay-independently exponentially stable, whereas in Example 2.3.6b, with values of $s_k(\varphi_c)$ explicitly listed in Figure 2.5, it is only delay-dependently exponentially stable, i.e., the exponential stability is lost at some critical delay h_c.

Example 2.3.6 (Zero as maximum of $\varphi \mapsto \alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1)$). *Consider (1.4) with $x(t) \in \mathbb{R}^{2p}$, $p \in \mathbb{N}_{>0}$, where the coefficients are given by the block diagonal matrices*

$$A_0 = \operatorname{blkdiag}\Big(\left[\begin{smallmatrix} -1 & -\beta \\ \beta & -1 \end{smallmatrix}\right], -I_2, \dots, -I_2 \Big)$$
$$A_1 = \operatorname{blkdiag}\big(Q(\vartheta_1), \dots, Q(\vartheta_p)\big),$$

with $Q(\vartheta) := \left[\begin{smallmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{smallmatrix}\right]$ and $\vartheta_j \in (-\pi, \pi] \setminus \{0\}$, $j \in \{1, \dots, p\}$.

a) With $\beta = 0$, i.e., $A_0 = -I_{2p}$, delay-independent exponential stability holds.

b) With $\beta = 1$, the zero equilibrium is exponentially stable for $h < h_c$ and not exponentially stable for $h = h_c$,

$$h_c := \begin{cases} -\vartheta_1 & \text{if } \vartheta_1 < 0, \\ 2\pi - \vartheta_1 & \text{if } \vartheta_1 > 0. \end{cases}$$

Reasoning: The $2p$ eigenvalues of $M(\varphi) = A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1$ are

$$\begin{aligned} s_{1,2}(\varphi) &= -1 \pm \mathrm{i}(-\beta) + \mathrm{e}^{\mathrm{i}(\varphi \pm \vartheta_1)} \\ &= -1 + \cos(\varphi_1 \pm \vartheta_1) + \mathrm{i}\big(\mp \beta + \sin(\varphi \pm \vartheta_1)\big) \\ s_k(\varphi) &= -1 + \mathrm{e}^{\mathrm{i}(\varphi \pm \vartheta_j)}, \end{aligned}$$

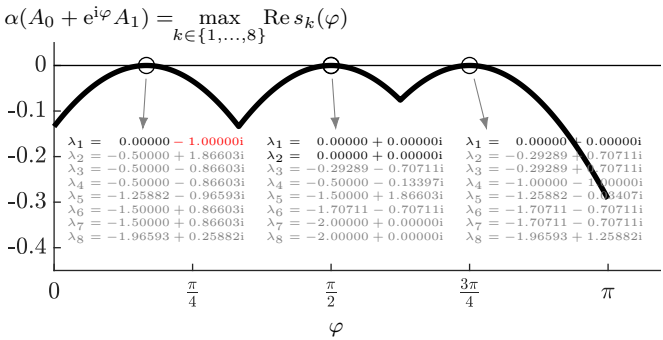


Figure 2.5[*]: Example 2.3.6b with $x(t) \in \mathbb{R}^8$ and $\vartheta_1 = -\pi/6, \vartheta_2 = \vartheta_3 = \pi/2, \vartheta_4 = 3/4\,\pi$. Since a non-zero purely imaginary eigenvalue $s_k(\varphi) = \mathrm{i}\omega \neq 0$ occurs (red), Theorem 2.3.4 (ii) does not apply. Instead, Theorem 2.2.2 yields $h_c = \pi/6$.

$k := 2j - \frac{1}{2}(1 \pm 1)$, $j \in \{2, \ldots, p\}$, *and their largest real part is*

$$\alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1) = -1 + \max_{j\in\{1,\ldots,p\}} \cos(\varphi \pm \vartheta_j)$$

with $\max_{\varphi\in[0,\pi]} \alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1) = 0$. *This maximum on* $[0, \pi]$ *is attained at*

$$\varphi_c \in \Phi := \{|\vartheta_1|, \ldots, |\vartheta_p|\}.$$

a) If $\beta = 0$, *then at any* $\varphi_c \in \Phi$ *no non-zero purely imaginary eigenvalues exist since* $\mathrm{Im}(s_k(\varphi_c)) = \sin(\varphi_c \pm \vartheta_j) = 0$ *when* $\cos(\varphi_c \pm \vartheta_j) = 1$. *Hence, case* (ii) *in Theorem 2.3.4 applies.*

b) In contrast, if $\beta = 1$, *there is a non-zero purely imaginary eigenvalue at* $\varphi_c = |\vartheta_1|$, *namely* $s_1(-\vartheta_1) = -\mathrm{i}$ *if* $\vartheta_1 < 0$, *see Figure 2.5, or* $s_2(\vartheta_1) = \mathrm{i}$ *if* $\vartheta_1 > 0$. *Hence, Theorem 2.3.4* (ii) *does not hold. Instead, Theorem 2.2.2 provides the delay interval of exponential stability. Since* $\mathrm{Im}\, s_1(|\vartheta_1|) = -1 < 0$ *for* $\vartheta_1 < 0$, *the modulo operation in (2.20) is without effect for* $\vartheta_1 < 0$, *while for* $\vartheta_1 > 0$ *the numerator of* h_c *becomes* $2\pi - \varphi = 2\pi - |\vartheta_1|$.

[*] The author has prepublished Figure 2.5 in [S1], ©2023 IEEE.

2.3.3 Corollaries*

Considerations so far are based on the spectral abscissa. The logarithmic norm is related to the spectral abscissa, but it exhibits advantageous properties allowing further simplifications. Based on these, some known stability criteria can be inferred directly from Theorem 2.3.4 without elaborate proofs.

The spectral abscissa $\alpha(M)$ of a matrix $M \in \mathbb{C}^{n\times n}$ can be approached as close as desired by a logarithmic norm of M

$$\mu_\nu(M) \stackrel{\text{def}}{=} \lim_{h\to 0^+} \frac{\|I_n + hM\|_\nu - 1}{h}, \tag{2.36}$$

provided the involved matrix norm $\|\cdot\|_\nu$ is chosen in an optimal way depending on M. To be more precise, $\alpha(M) = \inf_\nu \mu_\nu(M)$ [54]. Thus, Theorem 2.3.4(i) can equivalently be expressed in this manner. Usually, however, a matrix norm is chosen a priori. For common norms, (2.36) simplifies to well-known formulas [53, p. 33], e.g., the logarithmic norm w.r.t. the spectral norm $\|\cdot\|_2$ equals the maximum eigenvalue of the Hermitian part of M

$$\mu_2(M) = \lambda_{\max}\big(\tfrac{1}{2}(M + M^H)\big). \tag{2.37}$$

In any case, inequality (2.38a) holds.

Lemma 2.3.7 (Properties of $\mu_\nu(\cdot)$ [53]). *Let $M, N \in \mathbb{C}^{n\times n}$. Then*

$$\alpha(M) \le \mu_\nu(M), \tag{2.38a}$$
$$\mu_\nu(M+N) \le \mu_\nu(M) + \mu_\nu(N), \tag{2.38b}$$
$$\mu_\nu(M) \le \|M\|_\nu. \tag{2.38c}$$
$$\mu_\nu(\overline{M}) = \mu_\nu(M), \tag{2.38d}$$
$$\mu_\nu(M) = \sup_{\|x\|_\nu=1} \mathrm{Re}[Mx, x]_\nu \quad \textit{[133, Lem. 12]}, \tag{2.38e}$$

* The author has prepublished Section 2.3.3 in [S1], ©2023 IEEE.

where $[\cdot,\cdot]_\nu$ *is a semi-inner product with* $[x,x]_\nu = \|x\|_\nu^2$.

Consider the following expressions and relations

$$\max_{\varphi\in[0,\pi]} \alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1) \overset{(2.38a)}{\leq} \max_{\varphi\in[0,\pi]} \mu_\nu(A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1) =: \mathcal{B}_\mu \tag{2.39}$$

$$\overset{(2.38b)}{\leq} \mu_\nu(A_0) + \max_{\varphi\in[0,\pi]} \mu_\nu(\mathrm{e}^{\mathrm{i}\varphi}A_1) =: \tilde{\mathcal{B}}_r$$

$$\overset{(2.38c)}{\leq} \mu_\nu(A_0) + \|A_1\|_\nu =: \mathcal{B}_M, \tag{2.40}$$

as well as

$$\mu_\nu(A_0) + r_{\mathrm{num},\nu}(A_1) =: \mathcal{B}_r, \tag{2.41}$$

with the numerical radius $r_{\mathrm{num},\nu}(A) = \sup_{\|x\|_\nu=1} |[Ax,x]_\nu|$.

Lemma 2.3.8. *The equality* $\tilde{\mathcal{B}}_r = \mathcal{B}_r$ *holds.*

Proof.

$$\max_{\varphi\in[0,\pi]} \mu_\nu(\mathrm{e}^{\mathrm{i}\varphi}A_1) \overset{(2.38d,e)}{=} \max_{\varphi\in(-\pi,\pi]} \sup_{\|x\|_\nu=1} \mathrm{Re}(\mathrm{e}^{\mathrm{i}\varphi}[A_1x,x]_\nu)$$

$$= \sup_{\|x\|_\nu=1} \left|[A_1x,x]_\nu\right| \overset{\mathrm{def}}{=} r_{\mathrm{num},\nu}(A_1).$$

□

Based on the relations of $\mathcal{B}_M, \mathcal{B}_\mu$, and $\tilde{\mathcal{B}}_r = \mathcal{B}_r$ from the inequality chain (2.40), three immediate corollaries of Theorem 2.3.4 can be summarized.

Corollary 2.3.9. *Consider the definitions in (2.39), (2.40), and (2.41). If for some norm $\|\cdot\|_\nu$ one of the following inequalities holds, which are ordered by decreasing conservatism,*

$$\begin{aligned} &\text{(I)} && \mathcal{B}_M < 0, && \textit{(Mori's criterion [143, Thm. 1])},\\ &\text{(II)} && \mathcal{B}_r < 0, && \textit{(cf. [43, Thm. 2.2])},\\ &\text{(III)} && \mathcal{B}_\mu < 0, && \textit{(cf. [90, Thm. 1])}, \end{aligned}$$

then the zero equilibrium of (1.4) is strongly delay-independently exponentially stable.

2.4 Revisiting the Main Points of the Chapter

- A framework of three possible perspectives, $(\mathrm{P_S})$, $(\mathrm{P_Z})$, and $(\mathrm{P_{SZ}})$, on the two-variable formulation of the characteristic equation is introduced (Section 2.1.2).
- Based on the latter, consequent analogues to the delay-dependent (Theorem 2.2.1) and delay-independent (Theorem 2.3.1) frequency-sweeping criteria are developed.
- The derived criteria focus on eigenvalues of the matrix $M(\varphi) = A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1$. Contrary to the frequency sweeping approach, no generalized eigenvalues or matrix inverses are needed and eigenvalues of $M(\varphi)$ must only be evaluated on the bounded domain $\varphi \in [0, \pi]$.
- Concerning delay-independent stability (Theorem 2.3.4), the spectral abscissa function $\varphi \mapsto \alpha(A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1)$ on $\varphi \in [0, \pi]$ is considered.
  - A plot like Figure 2.4 does not require more than three lines of code.
  - Exclusively negative values of the spectral abscissa function indicate delay-independent stability.

  - The ambiguous case of a zero maximum is discussed in Example 2.3.6.
  - If positive values of the function occur, delay-independent exponential stability can be excluded immediately.
  - The latter statement is central for the necessity of the given criterion and relies on the proof that, if delay-independent stability holds, eigenvalues of $A_0 + \mathrm{e}^{\mathrm{i}\varphi} A_1$ cannot move between the left and right complex half-plane only by tunneling through the origin (see Figure 2.3).
  - Mori's famous logarithmic-norm-based criterion (1.12), which is more conservative, follows from simple inequality estimations in (2.40).

- Concerning delay-dependent stability (Theorem 2.2.2, discussed before delay-independent stability in this chapter), the angles $\varphi \in [0, \pi]$ for which real parts of eigenvalues of $A_0 + \mathrm{e}^{\mathrm{i}\varphi} A_1$ become zero must be determined.
  - These zeros include the zeros of the spectral abscissa function from the delay-independent consideration. Example 2.2.3 shows that the zeros of the remaining eigenvalue real parts indeed cannot be ignored.
  - The zeros constitute the constraint set in a proposed minimization problem for the first critical delay (Figure 2.1).
  - The delay-dependent stability chart for scalar systems can easily be derived (Figure 2.2).

3 Method: ODE-Approximation Schemes

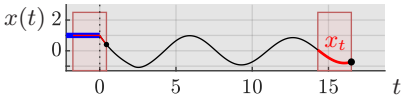
The present chapter is devoted to preliminaries concerning the discretization of time-delay systems in terms of an ODE approximation. The latter is needed for the numerical approaches proposed in Chapter 4 and Chapter 6 of this thesis.

The chapter is structured as follows. **Section 3.1** outlines what the desired ODE approximations are about. **Section 3.2** shows how to describe the exact time evolution of the overall state x_t for the linear time-delay system (1.4). In Section 3.3 et seq., these dynamics are discretized via two spectral methods: Chebyshev collocation (**Section 3.4**) and Legendre tau (**Section 3.5**). As both methods rely on different coordinate representations of polynomials, **Section 3.6** considers how to change the basis. Finally, stability of the equilibrium in the obtained ODE approximations is discussed in **Section 3.7**, before **Section 3.8** revisits the most relevant aspects of the chapter.

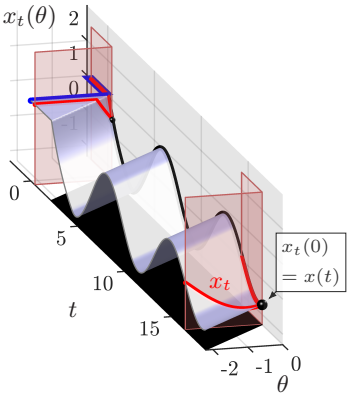
In a very shortened form, the author has prepublished important points of this chapter in [S2] (which is the prepublication of the contributions of Chapter 4).

3.1 Problem Statement

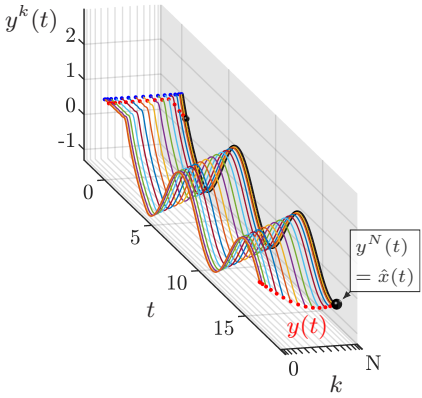
The state $x_t(\theta) = x(t+\theta), \theta \in [-h, 0]$, at time $t \geq 0$ of the RFDE (1.4) describes the solution segment on the preceding delay interval $[t-h, t]$. See Figure 3.1a



(a) RFDE solution, state x_t as solution segment



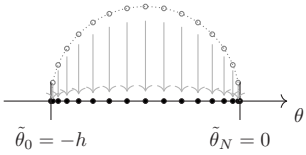
(b) Evolution of the RFDE state x_t



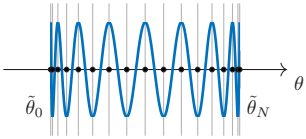
(c) Components $y^k(t)$ of the ODE solution ($N = 16$, A_y from (3.16))

Figure 3.1[*]**:** Solution of the scalar (i.e., $n = 1$) RFDE $\dot{x}(t) = -0.5x(t) - x(t - 2.2)$ for the initial function $x_0(\theta) \equiv 1$.

[*] The author has prepublished Figure 3.1 in [S2], ©2024 IEEE.



(a) Geometric intuition of the cosine in (3.2): the non-equidistant (one-dimensional) grid as a projection of equidistant circle points.



(b) Relation to Chebyshev polynomials: The $N{+}1$ (Gauss–Lobatto) Chebyshev nodes coincide with the extrema of the N-th Chebyshev polynomial (shifted and scaled to the domain $\theta \in [\tilde{\theta}_0, \tilde{\theta}_N]$)

Figure 3.2: Chebyshev nodes form a nonequidistant grid $\{\tilde{\theta}_k\}_{k\in\{0,\dots,N\}}$ (black points). The exemplary shown resolution N in the above plots is $N = 16$, i.e., the grid has $N{+}1 = 17$ nodes.

and Figure 3.1b. In contrast, the state of an ODE approximation must be a finite-dimensional vector. For instance, this state vector $y(t)$ at time t can approximate the values of the segment x_t in $N{+}1$ distinct points $\tilde{\theta}_0 = -h, \dots, \tilde{\theta}_N = 0$,

$$\begin{bmatrix} x(t-h) \\ x(t+\tilde{\theta}_1) \\ \vdots \\ x(t+\tilde{\theta}_{N-1}) \\ x(t) \end{bmatrix} = \begin{bmatrix} x_t(-h) \\ x_t(\tilde{\theta}_1) \\ \vdots \\ x_t(\tilde{\theta}_{N-1}) \\ x_t(0) \end{bmatrix} \approx \underbrace{\begin{bmatrix} y^0(t) \\ y^1(t) \\ \vdots \\ y^{N-1}(t) \\ y^N(t) \end{bmatrix}}_{y(t)\in\mathbb{R}^{n(N+1)}} =: \left[\begin{array}{c} z^0(t) \\ z^1(t) \\ \vdots \\ z^{N-1}(t) \\ \hline \hat{x}(t) \end{array}\right], \tag{3.1}$$

where upper indices $k \in \{0, \dots, N\}$ address vector-valued components $y^k(t) \in \mathbb{R}^n$. The decomposition $y = [z^\top, \hat{x}^\top]^\top$, which is indicated in the right-hand side of (3.1), is henceforth employed whenever the special interest in y^N shall be emphasized.

For $\tilde{\theta}_k$ in (3.1), a non-equidistant (one-dimensional) grid $\{\tilde{\theta}_k\}_{k\in\{0,\dots,N\}}$,

$$\tilde{\theta}_k = \tfrac{h}{2}(\tilde{\vartheta}_k - 1), \qquad \text{with} \quad \tilde{\vartheta}_k = -\cos(\tfrac{k}{N}\pi), \tag{3.2}$$

constructed from shifting and scaling classical Chebyshev nodes[1] $\tilde{\vartheta}_k \in [-1, 1]$ to $\tilde{\theta}_k \in [-h, 0]$, is known to be advantageous [177]. See Figure 3.2. Note that the latter is also at the core of the open-source Matlab toolbox Chebfun by Trefethen and co-workers [59], which proves to be very helpful in some implementations.

The objective of this chapter is to find an ODE

$$\dot{y}(t) = A_y\, y(t), \tag{3.3}$$

$A_y \in \mathbb{R}^{n(N+1)\times n(N+1)}$, that describes the dynamics of y in (3.1). Figure 3.1c depicts the solution of such an ODE (3.3), given the initial vector $y(0)$, indicated by the blue points, is a discretization of the initial function $x_0 \in C([-h, 0], \mathbb{R}^n)$ (see (3.23) with $\phi = x_0$). To obtain the desired ODE, this thesis employs two alternative methods: the Chebyshev collocation method (Section 3.4) and the Legendre tau method (Section 3.5) combined with a change of basis. The resulting system matrices A_y are given in (3.16) and (3.63) below.

3.2 The Infinite-Dimensional Dynamics

In the following, it is clarified how the exact time evolution of $t \mapsto x_t$ in Figure 3.1b looks like. These considerations build a foundation for the subsequent sections that derive a finite-dimensional approximation thereof to describe the time evolution of $t \mapsto y(t)$ in Figure 3.1c.

Consider the linear RFDE $\dot{x}(t) = A_0x(t)+A_1x(t-h)$ from (1.4) with the solution segment $x_t(\theta) = x(t + \theta)$, $\theta \in [-h, 0]$, at time $t \geq 0$. The time evolution of $t \mapsto x_t \in C([-h, 0], \mathbb{R}^n) = C$ obeys an abstract differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}x_t = \mathcal{A}x_t \tag{3.4}$$

[1] also called Gauss–Lobatto Chebyshev nodes (see Table 3.1) or Chebyshev points of the second kind (despite of referring to extrema of the "Chebyshev polynomials of the first kind", see Figure 3.2b) or endpoints-and-extrema Chebyshev nodes.

on the infinite-dimensional state space C of continuous ($\mathbb{R}^n$-valued) functions on $[-h, 0]$. Alternatively, by appending the important boundary value $x_t(0) = x(t)$, emphasized by the black point in Figure 3.1b, any $x_t \in C$ gives rise to a tuple $\left[\begin{smallmatrix} x_t(\cdot) \\ x_t(0) \end{smallmatrix}\right] \in C \times \mathbb{R}^n \subset L_2 \times \mathbb{R}^n = M_2$ that obeys an abstract differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\begin{smallmatrix} x_t(\cdot) \\ x_t(0) \end{smallmatrix}\right] = \mathscr{A}\left[\begin{smallmatrix} x_t(\cdot) \\ x_t(0) \end{smallmatrix}\right] \tag{3.5}$$

on the space M_2. Both will be made precise below.

To make sense of these abstract ODEs and, actually, even to apply the numerical methods in the subsequent sections, it suffices to understand the dynamics of the map $(t, \theta) \mapsto x_t(\theta) \in \mathbb{R}^n$. If both t and θ are considered as independent variables, then $(t, \theta) \mapsto x_t(\theta)$ can be recognized as the surface in Figure 3.1b. The latter is the solution of a PDE in (t, θ). First, note that $x_t(\theta) \stackrel{\text{def}}{=} x(t + \theta)$ inevitably has equal derivatives w.r.t. both t and θ. Second, note that the actual RFDE $\dot{x}(t) = A_0 x(t) + A_1 x(t - h)$ from (1.4) relates $\dot{x}(t) = \frac{\partial}{\partial t} x_t(\theta)|_{\theta=0}$ with $x(t) = x_t(0)$ and $x(t - h) = x_t(-h)$. Hence, the map $(t, \theta) \mapsto x_t(\theta)$ must obey the dynamics

$$\frac{\partial}{\partial t} x_t(\theta) = \frac{\partial}{\partial \theta} x_t(\theta), \qquad \theta \in [-h, 0), \qquad t > 0, \tag{3.6a}$$

$$\frac{\partial}{\partial t} x_t(0) = A_0 x_t(0) + A_1 x_t(-h), \qquad t > 0. \tag{3.6b}$$

The latter, in fact, is the foundation of the abstract ODE for $(x_t(\cdot), x_t(0))$ in M_2 that will be discussed in Section 3.2.2. To establish the connection to a more standard PDE description, (3.6b) can be rewritten as a boundary condition in terms of the spatial derivative by recognizing that $\frac{\partial}{\partial t} x_t(\theta)|_{\theta=0} = \frac{\partial}{\partial \theta} x_t(\theta)|_{\theta=0}$,

$$\frac{\partial}{\partial \theta} x_t(0) = A_0 x_t(0) + A_1 x_t(-h), \qquad t > 0. \tag{3.6b$'$}$$

In fact, (3.6a) with the boundary condition (3.6b′) is the foundation of the abstract ODE for $x_t(\cdot)$ in C that will be discussed in Section 3.2.1. Clearly, in any case, for a Cauchy problem of the RFDE with $x_0 = \varphi \in C$, the initial condition

$$x_0(\theta) = \varphi(\theta), \qquad \theta \in [-h, 0] \tag{3.7}$$

must also be appended.

Remark 3.2.1 (PDE description). *Written in terms of $u(t,\theta) := x_t(\theta)$ the overall PDE system from (3.6a), (3.6b′), and (3.7) becomes*

$$\frac{\partial u}{\partial t}(t,\theta) = \frac{\partial u}{\partial \theta}(t,\theta), \qquad t > 0,\ \theta \in [-h, 0),$$

boundary condition at θ=0 (involving $\theta = -h$)

$$\frac{\partial u}{\partial \theta}(t,0) = A_0 u(t,0) + A_1 u(t,-h), \quad t > 0,$$

initial condition at t=0

$$u(0,\theta) = \varphi(\theta), \qquad \theta \in [-h, 0],$$

where the boundary condition is a nonlocal (Robin) boundary condition that involves both the boundary $\theta = 0$ and $\theta = -h$.

The finite-dimensional ODE approximations that will be discussed in this chapter are semi-discretizations of the PDE w.r.t. the spatial variable θ.

3.2.1 Abstract ODE on the Banach Space C

The abstract ODE for the state $x_t \in C([-h,0],\mathbb{R}^n) = C$ from the RFDE $\dot{x}(t) = A_0 x(t) + A_1 x(t-h)$ in (1.4) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t} x_t = \mathcal{A} x_t, \qquad t > 0, \tag{3.9a}$$

$$x_0 = \varphi \in C \tag{3.9b}$$

with the unbounded operator $\mathcal{A}\colon C \supset D(\mathcal{A}) \to C$,

$$\mathcal{A}\phi = \phi', \tag{3.9c}$$

$$D(\mathcal{A}) = \{\phi \in C : \phi' \in C, \phi'(0) = A_0\phi(0) + A_1\phi(-h)\} \tag{3.9d}$$

(where $\phi'(\theta) = \frac{\mathrm{d}}{\mathrm{d}\theta}\phi(\theta)$). In particular, the operator $\mathcal{A}$ is the infinitesimal generator of a C_0-semigroup $\{\mathcal{T}(t)\}_{t\geq 0}$ built from the solution operators in $x_t = \mathcal{T}(t)x_0$ for $t \geq 0$. See, e.g., [82, 30] for details.

Note that the actual RFDE (1.4), respectively the boundary condition (3.6b′), only affects the domain of $\mathcal{A}$, whereas $\mathcal{A}$ itself—independently from A_0 and A_1—is always only a differentiation operator that takes account of (3.6a).

3.2.2 Embedding in the Hilbert Space M_2

Throughout this thesis, the considered state space is the Banach space of continuous functions, i.e., $x_t \in C$. However, the fact that any continuous function on $[-h, 0]$ is L_2-integrable will be used. The Hilbert space L_2 itself is no appropriate state space for time-delay systems since, from a given initial function x_0, the pointwise initial value $x(0) = x_0(0)$ is particularly decisive for the solution. However, instead of x_t, the pair $(x_t, x(t)) \in L_2 \times \mathbb{R}^n = M_2$ can be used, which treats the decisive boundary value $x(t) = x_t(0)$ separately (cf. Figure 3.1). If only linear time-delay systems were of interest, the overall space M_2 would indeed also be an appropriate state space. Nevertheless, the usage of M_2 in the present thesis only relies on the fact that, for any $\phi \in C$,

$$\begin{bmatrix} \phi \\ \phi(0) \end{bmatrix} \in C([-h, 0], \mathbb{R}^n) \times \mathbb{R}^n \subset L_2([-h, 0], \mathbb{R}^n) \times \mathbb{R}^n \tag{3.10}$$

is an element of M_2. See [156, Prop. 6.12 et seq.] for details on the underlying embedding $C \to M_2; \phi \mapsto \begin{bmatrix} \phi \\ \phi(0) \end{bmatrix}$.

The evolution of $\left[\begin{smallmatrix} x_t \\ x_t(0) \end{smallmatrix}\right] \in M_2$ from the linear RFDE (1.4) obeys the abstract ODE on the Hilbert space M_2

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\begin{smallmatrix} x_t \\ x_t(0) \end{smallmatrix}\right] = \mathscr{A}\left[\begin{smallmatrix} x_t \\ x_t(0) \end{smallmatrix}\right], \qquad t > 0, \tag{3.11a}$$

$$\left[\begin{smallmatrix} x_0 \\ x_0(0) \end{smallmatrix}\right] = \left[\begin{smallmatrix} \varphi(\cdot) \\ \varphi(0) \end{smallmatrix}\right] \tag{3.11b}$$

with the unbounded operator $\mathscr{A}\colon M_2 \supset D(\mathscr{A}) \to M_2$,

$$\mathscr{A}\begin{bmatrix} \phi \\ r \end{bmatrix} = \begin{bmatrix} \phi' \\ A_0 r + A_1 \phi(-h) \end{bmatrix}, \tag{3.12}$$

$$D(\mathscr{A}) = \{\left[\begin{smallmatrix} \phi \\ r \end{smallmatrix}\right] \in M_2 : r = \phi(0), \phi' \in L_2, \phi \in AC\}$$

(denoting $\phi'(\theta) = \frac{\mathrm{d}}{\mathrm{d}\theta}\phi(\theta)$). This operator $\mathscr{A}$ is again the infinitesimal generator of a C_0-semigroup $\{\mathscr{T}(t)\}_{t\geq 0}$, see, e.g., [47, 15, 52].

In contrast to (3.9), the RFDE is no longer hidden in the domain[2]. Instead, due to the incorporation of the boundary value $x_t(0) = x(t)$ in the definition of the state $(x_t, x(t))$, the RFDE $\dot{x}(t) = A_0 x(t) + A_1 x(t-h)$, respectively (3.6b), can explicitly be recognized in the lower part of the operator $\mathscr{A}$ in (3.12).

[2] In this respect, a similar effect is achieved as in the so-called $\odot *$ (sun-star) calculus, where $\mathcal{X} = C([-h,0],\mathbb{R}^n)$ is instead embedded in the space $\mathcal{X}^{\odot *} = L_\infty([-h,0],\mathbb{R}^n) \times \mathbb{R}^n$. The latter offers a mature theoretical framework [57]. In particular, it resolves that the variation-of-constants formula relies on a fundamental-matrix solution with the initial function not being continuous at $t = 0$ (see Table B.1) and thus not belonging to $C([-h,0],\mathbb{R}^{n\times n})$ although C is the state space. The notation is as follows: Consider the shift semigroup $\{T_0(t)\}_{t\geq 0}$ which is generated by (3.9c)-(3.9d) with the trivial choice $A_0 = A_1 = 0$. Consequently, $T_0(t)\colon \mathcal{X} \to \mathcal{X}$; $\phi \mapsto T_0(t)\phi := x_t$ on $\mathcal{X} = C([-h,0],\mathbb{R}^n)$. On the dual space $\mathcal{X}^* \cong NBV([-h,0],\mathbb{R}^n)$, the adjoint semigroup $\{T_0^*(t)\}_{t\geq 0}$, $T_0^*(t) : \mathcal{X}^* \to \mathcal{X}^*$, is defined. However, the adjoint semigroup (see [57]) is not strongly continuous, unless considering it on the restricted dual space $\mathcal{X}^\odot \cong L^1 \times \mathbb{R}^n \subset X^*$. The dual of that restricted dual space is the space of interest $\mathcal{X}^{\odot *} \cong L_\infty([-h,0],\mathbb{R}^n) \times \mathbb{R}^n$ as it offers a perturbation theory to tackle the non-trivial RFDE. Again $\{T_0^{\odot *}(t)\}_{t\geq 0}$ is not strongly continuous unless restricting the space to the smaller space $\mathcal{X}^{\odot\odot}$, which, however, is isomorphic to the original state space $\mathcal{X} = C([-h,0],\mathbb{R}^n) \cong \mathcal{X}^{\odot\odot} \subset \mathcal{X}^{\odot *}$. See [57] for details.

3.3 Spectral Methods

Both numerical methods that are used in the sequel, Chebyshev collocation and Legendre tau, belong to the class of spectral methods. An introduction to these methods is, e.g., given in [88], see also [38, 68, 177, 65, 154, 73]. Like finite difference methods, finite element methods, or finite volume methods, spectral methods are used for the (spatial) discretization of differential equations. Such methods might be more prominently known for the numerical solution of PDEs. However, they are more generally applicable to general evolution equations on infinite-dimensional Banach or Hilbert spaces like the ones given above (which, in fact, also bear an underlying PDE description, see Remark 3.2.1).

Spectral methods are based on polynomial approximations. For instance, the finite-dimensional vector y introduced in (3.1) is understood as a representation of the uniquely defined interpolating polynomial through these points. The thus described polynomial serves as an approximation of the unknown function x_t. Consequently, the resulting approximation for x_t is a globally (in contrast to piecewise) defined high-order polynomial, and as such it is a smooth function on the whole domain $\theta \in [-h, 0]$. The latter, in fact, is the difference to finite element methods, where the ansatz for the numerical solution is a linear combination of basis functions with compact support (linear hat functions or, more generally, piecewise defined low-order polynomials), and which consequently only generate piecewise defined linear or low-order polynomial functions as numerical solutions.

Table 3.1 classifies all polynomial methods employed throughout this thesis. Alongside the spectral methods used for the discretization of the abstract ODE (second column), the table also lists the underlying polynomial approximation method (first column). In practice, this polynomial approximation becomes, e.g., relevant when discretizing a given initial function. The term discretization refers to the projection on a finite-dimensional space, with the result being representable by a finite number of coordinates. For instance, this might be the vector $y(0)$

* The author has prepublished a shortened version of Table 3.1 in [S2], ©2024 IEEE.

derived from an initial function x_0 as already indicated by the blue points in Figure 3.1. However, these coordinates are no loose points, but in the present context they always represent an underlying polynomial that is uniquely defined by these coordinates—in the case of the interpolation coordinates y, it is the unique interpolating polynomial through these points. Besides of the use in differential equations giving rise to spectral methods, the polynomial approximation of a continuous function also provides a simple but extremely effective means for the numerical integration of that function. These quadrature rules are listed in the right column of Table 3.1 and will become relevant in Section 4.2.

3.4 Chebyshev Collocation Method

To apply the Chebyshev collocation method (which is also called pseudospectral discretization) in order to discretize the infinitesimal generator $\mathcal{A}$ of the time-delay problem (3.9), goes back to Breda and co-workers [29, 30, 26]. The original purpose has been the numerical approximation of characteristic roots, which, in fact, are the eigenvalues of $\mathcal{A}$, via the eigenvalues of the resulting finite-dimensional system matrix A_y in (3.3). Due to the fast convergence of the most important characteristic roots and the capability to tackle RFDEs with large matrices A_0, A_1, this eigenvalue analysis is even encountered in stability considerations of large practical systems with time delays like power systems [141]. Besides of the eigenvalue calculation, applying the Chebyshev collocation method to time-delay systems is also successfully applied in fields like bifurcation analysis [28], Lyapunov exponents [27], or H_2-norm computations [101].

The core principle of a collocation method is interpolation. As mentioned above, the vector $y(t)$ at time t in (3.1), cf. Figure 3.1, determines the interpolating N-th degree approximating polynomial for x_t. This polynomial is described by

$$x_t(\theta) \approx \sum_{k=0}^{N} y^k(t)\ \ell_k\big(\vartheta(\theta)\big), \tag{3.13}$$

<table>
<tr><th colspan="2">Polynomial approximation
of functions</th><th colspan="3">Spectral methods
for differential equations</th><th colspan="3">Numerical integration
for integral expressions</th></tr>
<tr><td colspan="2">(JP stands for Jacobi polynomials like Chebyshev [Fig. 3.2b] or Legendre [Fig. 3.5])</td><td colspan="2">JP:</td><td>Chebyshev Legendre</td><td colspan="2">JP:</td><td>Chebyshev Legendre</td></tr>
<tr><td rowspan="2">Interpolation / coincidence in the chosen nodes [Fig. 3.4] (natural basis: Lagrange polynomials w.r.t. the chosen nodes [Fig. 3.3]), equivalently, “discrete expansion”[3] with the 0-th to N-th JP as basis, related[4] to an approximation of the series truncation below via quadrature</td><td>through Gauss–Lobatto nodes (extrema of the N-th JP), [Fig. 3.2b]</td><td rowspan="2">Collocation / pseudospectral method / method of selected points / at the nodes vanishing residual or—alternatively—G-NI / “Galerkin with numerical integration” methods</td><td colspan="2">Chebyshev col.</td><td rowspan="2">Interpolatory quadrature / integration of an interpolating polynomial instead of the original function / integral approximated by a weighted sum of interpolation coordinates</td><td colspan="2">Clenshaw–Curtis</td></tr>
<tr><td>through Gauss nodes (roots of the $(N+1)$-th JP)</td><td colspan="2">$\asymp$ (no boundary nodes for boundary conditions in collocation methods)</td><td colspan="2">Gauss quad.</td></tr>
<tr><td colspan="2">Series truncation / orthogonal[5] projection to the 0-th to N-th JP / “continuous expansion” with the 0-th to N-th JP as basis / generalized Fourier truncation / least squares best approximation</td><td colspan="2">Tau method / Lanczos’ tau m. / Galerkin with boundary bordering or—alternatively—Galerkin method</td><td>Legendre tau</td><td colspan="2">$\asymp$ (The projection requires itself integral evaluations)</td><td>$\asymp$ $\asymp$</td></tr>
</table>

Table 3.1*: Classification of the used methods (“/” marks synonymous terms or explanations). See, e.g., [88, 178, 177, 38, 68, 65, 23, 154, 73].

[3] orthogonal projection w.r.t. the so-called discrete inner product associated to the chosen nodes [88, 38].

[4] See [88, Thm. 5.3]. For Gauss–Lobatto nodes the result is not exactly the application of the quadrature rule to the integrals in the continuous expansion formula as the N-th order coefficient must be corrected.

[5] orthogonal w.r.t. the (weighted) inner product in which the chosen basis polynomials are orthogonal. In (3.47), the modified c^N makes the projection non-orthogonal, unless the discretization is interpreted in terms of (3.50).

see Figure 3.3j, where the notation is as follows. The functions $\ell_k \colon [-1,1] \to \mathbb{R}$ are interpolating Lagrange basis polynomials w.r.t. (Gauss–Lobatto) Chebyshev nodes $\{\tilde{\vartheta}_k\}_{k\in\{0,\dots,N\}}$ on $[\tilde{\vartheta}_0, \tilde{\vartheta}_N] = [-1,1]$, and

$$\vartheta \colon [-h,0] \to [-1,1]; \qquad \theta \mapsto \vartheta(\theta) := \tfrac{2}{h}\theta + 1 \tag{3.14}$$

maps the argument $\theta \in [\tilde{\theta}_0, \tilde{\theta}_N] = [-h,0]$ to this interval. Figure 3.3 exemplarily depicts the resulting scaled and shifted Lagrange basis polynomials $\ell_k\big(\vartheta(\theta)\big)$ for a grid of $N+1 = 9$ Chebyshev nodes.

3.4.1 Chebyshev Collocation: System Matrix (in Interpolation Coordinates)

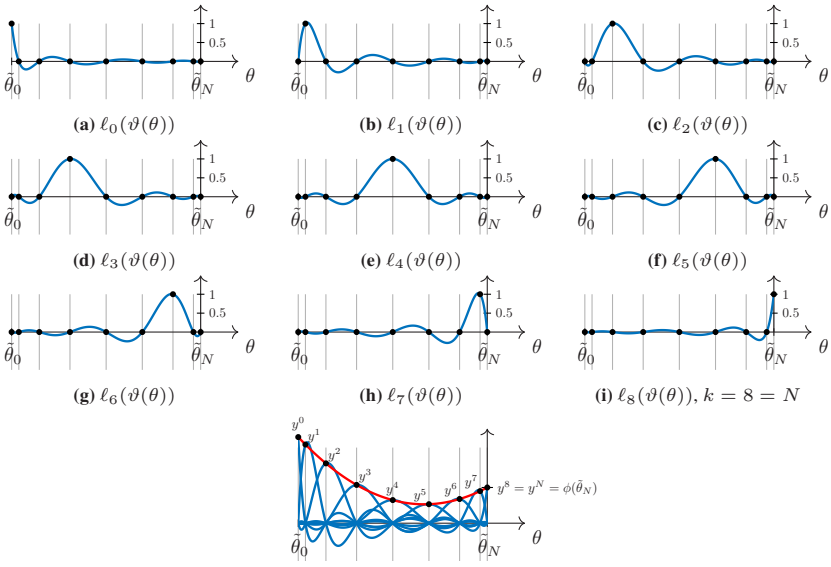
The Chebyshev collocation method yields for the dynamics of the unknown coefficients $y^k(t)$ in (3.13) the ODE

$$\dot{y}(t) = A_y^C\, y(t) \tag{3.15}$$

$$\text{with } A_y^C := \begin{bmatrix} \frac{2}{h}\ell_0'(\tilde{\vartheta}_0)I_n & \cdots & \cdots & \cdots & \frac{2}{h}\ell_N'(\tilde{\vartheta}_0)I_n \\ \vdots & & & & \vdots \\ \frac{2}{h}\ell_0'(\tilde{\vartheta}_{N-1})I_n & \cdots & \cdots & \cdots & \frac{2}{h}\ell_N'(\tilde{\vartheta}_{N-1})I_n \\ A_1 & 0_{n\times n} & \cdots & 0_{n\times n} & A_0 \end{bmatrix}, \tag{3.16}$$

cf. [28]. For the interested reader, it will be discussed below how this known result is derived. In view of the meaning of y from (3.1), the actual RFDE $\dot{x}(t) = A_0 x(t) + A_1 x(t-h)$ from (1.4) is easily recognized in the last (block-) row $\dot{y}^N(t) = A_0 y^N(t) + A_1 y^0(t)$ of (3.15). The remaining rows, i.e., the upper part $\frac{2}{h}(\ell_k'(\tilde{\vartheta}_j))_{j\in\{0,\dots,N-1\},k\in\{0,\dots,N\}} \otimes I_n$ of the system matrix (3.16), rely on the first N rows of the $(N+1)\times(N+1)$ differentiation matrix

$$D_{y,[-1,1]} = (\ell_k'(\tilde{\vartheta}_j))_{j,k\in\{0,\dots,N\}} \tag{3.17}$$



(a) $\ell_0(\vartheta(\theta))$ **(b)** $\ell_1(\vartheta(\theta))$ **(c)** $\ell_2(\vartheta(\theta))$

(d) $\ell_3(\vartheta(\theta))$ **(e)** $\ell_4(\vartheta(\theta))$ **(f)** $\ell_5(\vartheta(\theta))$

(g) $\ell_6(\vartheta(\theta))$ **(h)** $\ell_7(\vartheta(\theta))$ **(i)** $\ell_8(\vartheta(\theta))$, $k = 8 = N$

(j) $\phi(\theta) = \sum_{k=0}^{N} y^k \, \ell_k(\vartheta(\theta))$

Figure 3.3: Lagrange basis polynomials $\{\ell_k(\vartheta(\theta))\}_{k\in\{0,\dots,N\}}$ w.r.t. a grid of $N+1$ Chebyshev nodes. In the above plots, $N = 8$ is chosen. The k-th Lagrange basis polynomial $\ell_k(\vartheta(\theta))$ is the unique polynomial of degree N that is equal to one at the k-th Chebyshev node $\tilde{\theta}_k$ and zero at the remaining Chebyshev nodes. As a consequence, if $N+1$ given values $(y^k)_{k\in\{0,\dots,N\}}$ shall be attained on the Chebyshev grid $\{\tilde{\theta}_k\}_{k\in\{0,\dots,N\}}$, the unique interpolating polynomial of degree N through these points (red curve) is described by the superposition $y(\theta) = \sum_{k=0}^{N} y^k \, \ell_k(\vartheta(\theta))$. Moreover, given an arbitrary polynomial of degree at most N, it can be written as $\phi(\theta) = \sum_{k=0}^{N} y^k \, \ell_k(\vartheta(\theta))$ since the $N+1$ Lagrange polynomials $\{\ell_k(\vartheta(\theta))\}_{k\in\{0,\dots,N\}}$ form a basis for the finite-dimensional space of polynomials of degree at most N. In other words, any such polynomial ϕ is uniquely determined by its interpolation coordinates $(y^k)_{k\in\{0,\dots,N\}}$, which are simply given by the $N+1$ pointwise evaluations $y^k = \phi(\tilde{\theta}_k)$.

taking account of the differentiation operator from the abstract ODEs in Section 3.2. In fact, the scaled differentiation matrix

$$\begin{aligned} D_y &= \left(\tfrac{\partial}{\partial\theta}\ell_k(\vartheta(\theta))|_{\theta=\tilde\theta_j}\right)_{j,k\in\{0,\dots,N\}} = (\vartheta'\ell_k'(\tilde\vartheta_j))_{j,k\in\{0,\dots,N\}} \\ &= \tfrac{2}{h}D_{y,[-1,1]} \end{aligned} \tag{3.18}$$

occurs in (3.16) due to the definition of x_t on the scaled domain $[-h,0]$, cf. (3.2) and (3.14) with $\vartheta'(\theta) = \frac{2}{h}$. A further discussion on the differentiation matrix will be given in Remark 3.6.2. Its numerical implementation is addressed in the following remark.

Remark 3.4.1 (Implementation of A_y^C)*. *For an implementation of the skew-centrosymmetric differentiation matrix, see, e.g., [177, p. 54] (in terms of $x_k = -\tilde\vartheta_k$). A ready to use Matlab routine is given by* `diffmat` *in the Chebfun toolbox [59]. Based on the latter, $A_y^C =$* `Ay` *in (3.16) is obtained via*

```
Dy=diffmat(N+1,[-delay,0]); Ay=kron(Dy,eye(n));
Ay(end-n+1:end,:)=[A1,zeros(n,n*(N-1)),A0]
```

(if A_0, A_1, h, n, N is assigned to `A0,A1,delay,n,N`*).*

Interlude: How the Chebyshev-Collocation System Matrix is Derived

A collocation method derives the searched approximating finite-dimensional ODE from the infinite-dimensional problem by requiring that the residual shall vanish pointwise on a grid of selected values of θ. If this grid is a grid of Chebyshev nodes, the method is called Chebyshev collocation.

First, consider the interior nodes $\{\tilde\theta_k\}_{k\in\{0,\dots,N-1\}}$ without the boundary node $\tilde\theta_N = 0$, which, due to the boundary condition, plays a special role. The ansatz

* The implementation hint is part of [S2].

for the solution is the polynomial already stated in (3.13) above. With that ansatz, the residual from (3.6a), respectively (3.9a), becomes

$$\frac{\partial}{\partial t}\left(\sum_{k=0}^{N} y^k(t)\,\ell_k(\vartheta(\theta))\right) - \frac{\partial}{\partial \theta}\left(\sum_{k=0}^{N} y^k(t)\,\ell_k(\vartheta(\theta))\right). \tag{3.19}$$

Thus, it vanishes on the interior nodes if

$$\forall \theta \in \{\tilde{\theta}_0, \dots, \tilde{\theta}_{N-1}\}: \qquad \sum_{k=0}^{N} \dot{y}^k(t)\,\ell_k(\vartheta(\theta)) - \sum_{k=0}^{N} y^k(t)\,\vartheta'(\theta)\ell_k'(\vartheta(\theta)) = 0. \tag{3.20}$$

It is the gist of Lagrange basis polynomials that $\ell_k(\vartheta(\tilde{\theta}_j)) = 0$ if $k \neq j$ and $\ell_k(\vartheta(\tilde{\theta}_j)) = 1$ if $k \neq j$, see Figure 3.3—which, in fact, is the reason why the polynomial ansatz is represented in that basis when using the Chebyshev collocation method. As a consequence, the evaluation of (3.20) at one node after the other yields

$$\begin{aligned} \theta = \tilde{\theta}_0: &\qquad \dot{y}^0(t) - \sum_{k=0}^{N} y^k(t)\,\vartheta'\ell_k'(\vartheta(\tilde{\theta}_0)) = 0, \\ &\qquad\qquad \vdots \\ \theta = \tilde{\theta}_{N-1}: &\qquad \dot{y}^{N-1}(t) - \sum_{k=0}^{N} y^k(t)\,\vartheta'\ell_k'(\vartheta(\tilde{\theta}_{N-1})) = 0, \end{aligned} \tag{3.21}$$

where, by (3.14), $\vartheta' = \frac{2}{h}$. This already explains all but the last row in the ODE (3.15).

The last row realizes the boundary condition from (3.6b), respectively (3.9d), which is actually the RFDE (1.4) and, in view of (3.1), obviously becomes

$$\theta = \tilde{\theta}_N: \qquad \dot{y}^N(t) = A_0 y^N(t) + A_1 y^0(t) \tag{3.22}$$

(clearly, the above formula is also obtained from (3.6b) using the ansatz (3.13) and noting that the thus derived $\sum_{k=0}^{N} \dot{y}^k(t)\ell_k(\vartheta(0)) = A_0 \sum_{k=0}^{N} y^k(t)\ell_k(\vartheta(0)) + A_1 \sum_{k=0}^{N} y^k(t)\ell_k(\vartheta(-h))$ simplifies to (3.22) by the construction of the Lagrange basis polynomials, where $0 = \tilde{\theta}_N$ and $-h = \tilde{\theta}_0$).

3.4.2 Discretization: Interpolation through Chebyshev Nodes

When considering a Cauchy problem with an initial function x_0, a discretization must be employed to obtain the initial vector $y(0)$ for the finite-dimensional ODE. Moreover, for the approach that will be proposed in the next chapter, a discretization of a given function ϕ, being the argument of the Lyapunov–Krasovskii functional, will also be needed.

In view of (3.1), such a discretization y of $\phi \in C$ is simply obtained by evaluating the vector-valued function ϕ at the grid points (3.2) and stacking these $(N+1)$ vectors in

$$y = \begin{bmatrix} | \\ z \\ | \\ \hat{x} \end{bmatrix} = \begin{bmatrix} \phi(-h) \\ \phi(\tilde{\theta}_1) \\ \vdots \\ \phi(\tilde{\theta}_{N-1}) \\ \phi(0) \end{bmatrix}. \tag{3.23}$$

If ϕ is already a polynomial of degree at most N, the above vector y is the unique Lagrange interpolation coordinate representation of that polynomial, see Figure 3.3j. Otherwise, if ϕ is not yet such a polynomial, let us denote by $\phi^{[N]}$ the

resulting polynomial of degree N that is represented by the Lagrange interpolation coordinates y from (3.23), i.e.,

$$\phi^{[N]}(\theta) = \sum_{k=0}^{N} y^k \ell_k(\vartheta(\theta)). \tag{3.24}$$

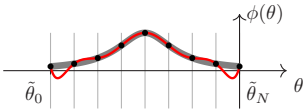
The latter is nothing else than the projection of ϕ from the space of continuous functions to an element $\phi^{[N]}$ in the finite-dimensional space of polynomials via interpolation, see the left column and upper row in Table 3.1.

Remark 3.4.2 (Near-best approximation). *According to Weierstrass' approximation theorem, see, e.g., [51, Thm. 6.1.1], continuous functions on a bounded closed interval like $[-h, 0]$ can arbitrarily well in the uniform norm be approximated by polynomials as the polynomial degree N increases. However, this convergence need not be achieved by a prescribed polynomial approximation scheme. In particular, interpolation through equidistant nodes turns out to be a poor choice, see Figure 3.4a and 3.4c. In contrast, interpolation through Chebyshev nodes (Figure 3.4b and 3.4d) is a near-best approximation [179], meaning that the error $\|\phi - \phi^{[N]}\|_C$ is close to the minimum achievable error $\|\phi - \phi^{*[N]}\|_C$.*

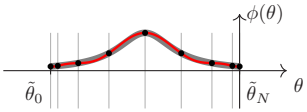
Still, without further regularity assumptions on ϕ, it theoretically might not mirror the uniform convergence $\|\phi - \phi^{*[N]}\|_C \to 0$ as $N \to \infty$, the Weierstrass approximation theorem attests for the best approximation $\phi^{*[N]}$ [51]. Nevertheless, the interpolation through Chebyshev nodes turns out to be, from a practical point of view, usually not very inferior to the best approximation [179].

- *First, although the comparison with the best approximation error, given by $\|\phi - \phi^{[N]}\|_C \leq (1 + \Lambda_N)\|\phi - \phi^{*[N]}\|_C$ [179, Thm. 15.1], involves an unbounded coefficient[6] $(1 + \Lambda_N)$, even for absurd large degrees $N \leq 10^5$ it holds $(1 + \Lambda_N) < 10$. Thus, for practical discretization resolutions, no*

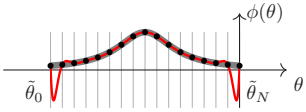
[6] Λ_N is the Lebesque constant of the linear projection. For the interpolation through Chebyshev nodes, $\Lambda_N \leq 1 + \frac{2}{\pi}\log(N + 1)$, whereas for the interpolation through equidistant nodes the Lebesque constant grows rapidly to huge numbers with $\Lambda_N > (2^{N-2})/(N^2)$, [179, Thm. 15.2].



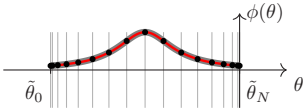
(a) Unique polynomial of degree 8 through $N + 1 = 9$ graph points of $\phi(\theta)$ from an equidistant grid



(b) Unique polynomial of degree 8 through $N + 1 = 9$ graph points of $\phi(\theta)$ from Chebyshev nodes



(c) Unique polynomial of degree 16 through $N + 1 = 17$ graph points of $\phi(\theta)$ from an equidistant grid



(d) Unique polynomial of degree 16 through $N + 1 = 17$ graph points of $\phi(\theta)$ from Chebyshev nodes

Figure 3.4: Reason for choosing a non-equidistant grid: The thick gray curve in the background is a continuous but non-polynomial function ϕ. Because of the so-called Runge phenomenon [161], interpolating polynomials of degree N (red curves in the left column) through evaluations of ϕ from equidistant nodes (black points in the left column) might give large local deviations. In contrast, the sequence $\{\phi^{[N]}\}_N$ of interpolating polynomials of degree N (red curves) through evaluations $\phi(\tilde{\theta}_k)$ from Chebyshev nodes (black points in the right column) is a so-called near-best approximation, cf. Remark 3.4.2. The plotted example function is $\phi(\theta) = \frac{1}{1+8(\vartheta(\theta))^2}$ with the affine map $\vartheta : [\tilde{\theta}_0, \tilde{\theta}_N] \to [-1, 1]$ from (3.14).

more than one digit accuracy can be lost compared to the best approximation.

- *Second, already if $\phi \in AC$ and the first derivative ϕ' is of bounded variation, uniform convergence holds. Moreover, the smoother[7] the function ϕ the faster $\|\phi - \phi^{[N]}\|_C$ converges to zero, see [179, Thm. 7.2].*
- *Third, already without further regularity assumptions, for any $\phi \in C$ the integral of the interpolation through Chebyshev nodes converges to the original value, i.e., $\int_{-h}^{0} \phi^{[N]}(\theta)\,\mathrm{d}\theta \to \int_{-h}^{0} \phi(\theta)\,\mathrm{d}\theta$, as $N \to \infty$. The latter, in fact, is the numerical integration via Clenshaw–Curtis quadrature [179, Thm. 19.4], see the right column in Table 3.1.*

Altogether, the above section can be summarized as follows. The Chebyshev collocation method leads to the ODE (3.3) with the dense system matrix (3.16), and its initial value $y(0)$ is simply obtained from the given initial function $x_0 = \phi$ by stacking pointwise evaluations of that function as described in (3.23).

3.5 Legendre Tau Method

The second spectral method used in this thesis—which, for the applications in the present thesis, will even turn out to be the recommended method—is the Legendre tau method. Applying this method to time-delay systems goes back to Ito and Teglas [98, 99, 97]. Already in their initial three-part work, the method is considered as a numerical solution approach, as an early lumping approach to optimal control problems, and as an approach for the approximation of eigenvalues. Still, the method is clearly less widespread than the Chebyshev collocation method for time-delay systems.

[7] With $\phi \in C^k$ it converges algebraically with $O(N^{-k})$, and for analytic functions ϕ, like the one in Figure 3.4, it converges even geometrically with $O(c^{-N})$, $c > 1$.

Again, a polynomial of degree N is taken as the ansatz for the numerical solution of x_t. Instead of Lagrange interpolation coordinates, Legendre coordinates are considered—at least in the derivations. Once the polynomial is determined, it can of course be represented arbitrarily, and the same holds for the underlying ODE. The associated change of basis will be discussed in Section 3.6.

Let $p_k\colon [-1,1] \to \mathbb{R}$ denote the k-th Legendre polynomial. See, e.g., [88] for formulas and further details. Since p_k is defined on $[-1,1]$, the affine map $\vartheta\colon [\tilde{\theta}_0, \tilde{\theta}_N] = [-h, 0] \to [-1,1]$ from (3.14) again comes into play. Using $\{p_k(\vartheta(\cdot))\}_{k=0}^N$ as basis—plotted in Figure 3.5—an approximating polynomial of degree N for x_t becomes

$$x_t(\theta) \approx \sum_{k=0}^{N} c^k(t)\; p_k(\vartheta(\theta)). \tag{3.25}$$

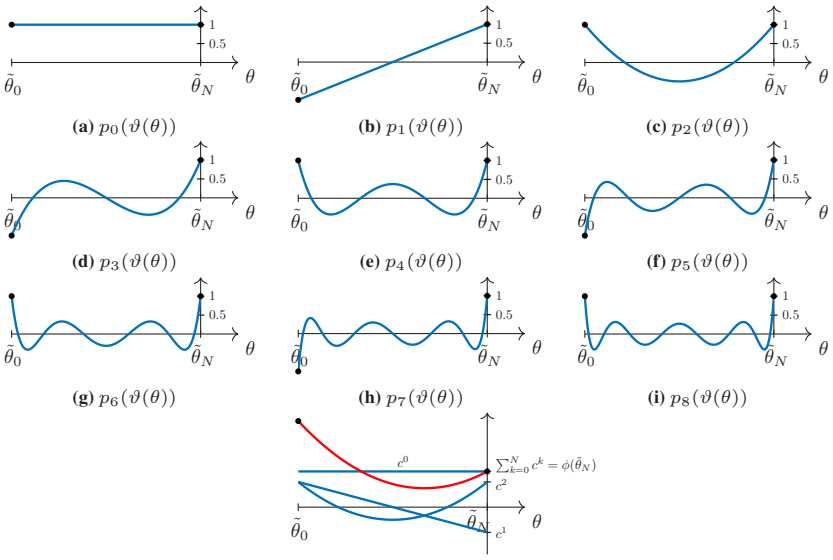
The evolution of the coefficients $c^k(t) \in \mathbb{R}^n$, stacked as $c := [(c^0)^\top, \dots, (c^N)^\top]^\top$, shall be described by the searched ODE

$$\dot{c}(t) = A_c\, c(t). \tag{3.26}$$

3.5.1 Legendre Tau: System Matrix (in Legendre Coordinates)

The Legendre tau method leads to the ODE (3.26) with $A_c = A_c^L$ having the block entries

$$A_c^{L,jk} = \begin{cases} \frac{2}{h}(2j+1) I_n & \text{if } j \in \{0,\dots,N-1\},\\ & \quad k > j, \text{ and } j+k \text{ odd},\\ A_0 + (-1)^k A_1 - \frac{2}{h}\frac{k(k+1)}{2} I_n & \text{if } j = N,\\ 0_{n\times n} & \text{otherwise}, \end{cases} \tag{3.27}$$



(a) $p_0(\vartheta(\theta))$ **(b)** $p_1(\vartheta(\theta))$ **(c)** $p_2(\vartheta(\theta))$

(d) $p_3(\vartheta(\theta))$ **(e)** $p_4(\vartheta(\theta))$ **(f)** $p_5(\vartheta(\theta))$

(g) $p_6(\vartheta(\theta))$ **(h)** $p_7(\vartheta(\theta))$ **(i)** $p_8(\vartheta(\theta))$

(j) $\phi(\theta) = \sum_{k=0}^{N} c^k \, p_k(\vartheta(\theta))$

Figure 3.5: The 0-th to 8-th shifted and scaled Legendre polynomials $\{p_k(\vartheta(\theta))\}_{k\in\{0,\dots,8\}}$. In contrast to Lagrange basis polynomials (Figure 3.3) w.r.t. a given grid, the k-th Legendre polynomial p_k itself is only a polynomial of degree k and remains unaltered if the overall dimension $N+1$ of the considered space of polynomials is increased. Again, any polynomial of degree at most N can be rewritten as $\phi(\theta) = \sum_{k=0}^{N} c^k \, p_k(\vartheta(\theta))$ since the $N+1$ Legendre polynomials $\{p_k(\vartheta(\theta))\}_{k\in\{0,\dots,N\}}$ form a basis for this $(N+1)$-dimensional space. Thus, any such polynomial ϕ is uniquely determined by its Legendre coordinates $(c^k)_{k\in\{0,\dots,N\}}$. Since the red curve represents a polynomial of degree two, $c^k = 0$ for $k > 2$.

cf. [98]. Thus, A_c^L exhibits the structure (exemplarily for N even)

$$A_c^L = \begin{bmatrix} 0_{n\times n} & 0_{n\times n} & 0_{n\times n} & \cdots & 0_{n\times n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0_{n\times n} & 0_{n\times n} & 0_{n\times n} & \cdots & 0_{n\times n} \\ A_0{+}A_1 & A_0{-}A_1 & A_0{+}A_1 & \cdots & A_0{-}A_1 \end{bmatrix}$$

$$+ \frac{2}{h} \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 3 & 0 & 3 & 0 & \cdots & 3 \\ 0 & 0 & 0 & 5 & 0 & 5 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 7 & 0 & \cdots & 7 \\ 0 & 0 & 0 & 0 & 0 & 9 & \cdots & 0 \\ \vdots & & & & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & (2N{-}1) \\ 0 & -1 & -3 & -6 & -10 & -15 & \cdots & -\frac{N(N+1)}{2} \end{bmatrix} \otimes I_n. \quad (3.28)$$

A detailed derivation of this known result will be given below for the sake of completeness. Similarly to the result from the Chebyshev collocation in Section 3.4.1, the upper part contains the first N of the $(N+1)$ rows of the triangular Legendre differentiation matrix D_c (see Remark 3.6.2 for the relation to D_y), whereas the lower (block-)row addresses the boundary condition (3.6b) which entails the actual RFDE (1.4).

Remark 3.5.1 (Implementation of A_c^L)*. *In standard Matlab code, the triangular unscaled Legendre differentiation matrix (see Figure 3.6 with $\vartheta' = 1$) can be obtained via*

```
DC=zeros(N+1,N+1); for j=0:N-1; DC(j+1,j+2:2:end)=2*j+1; end
```

based on which the scaled differentiation matrix D_c becomes `Dc=2/delay*DC`. *The overall system matrix $A_c^L =$* `Ac` *shown in (3.28) is*

```
DC(end,:)=-(0:N).*(1:N+1)/2;   Ac=2/delay*kron(DC,eye(n));
Ac(end-n+1:end,:)= Ac(end-n+1:end,:)+...
     kron(ones(1,N+1),A0)+kron((-1).^(0:N),A1)
```

(with `A0,A1,delay,n,N` *as in Remark 3.4.1).*

Interlude: How the Legendre-Tau System Matrix is Derived

A tau method derives the searched approximating finite-dimensional ODE (3.26) from the infinite-dimensional problem by requiring that the residual shall be orthogonal to the $(N+1-\beta)$-dimensional space of polynomials of degree at most $N-\beta$, where $\beta = 1$ is the number of boundary conditions. In other words, the inner products of the residual with basis polynomials of that space must vanish. If the orthogonality is understood in terms of the unweighted L_2 inner product, which is the inner product with respect to which Legendre polynomials are orthogonal, the method is called Legendre tau method. See [88] for further background on the method.

Using the polynomial ansatz (3.25) for the approximation of x_t, the residual from (3.6a) becomes

$$\frac{\partial}{\partial t}\left(\sum_{k=0}^{N} c^k(t)\, p_k(\vartheta(\theta))\right) - \frac{\partial}{\partial \theta}\left(\sum_{k=0}^{N} c^k(t)\, p_k(\vartheta(\theta))\right). \tag{3.29}$$

Thus, to establish that this residual is L_2-orthogonal to the 0-th to $(N-1)$-th Legendre polynomial (which span the space of polynomials of degree at most $N-1$) the inner products

$$\forall j \in \{0,\dots,N-1\}: \quad \left\langle \sum_{k=0}^{N} \dot{c}^k(t) p_k(\vartheta(\cdot)) - \sum_{k=0}^{N} c^k(t)\, \vartheta' p_k'(\vartheta(\cdot)),\quad p_j(\vartheta(\cdot)) \right\rangle_{L_2([-h,0],\mathbb{R}^n)} = 0 \tag{3.30}$$

must vanish. The occurring factor $\vartheta' = \frac{2}{h}$ denotes the derivative of (3.14). In the above expression, compositions with the affine mapping $\vartheta\colon [-h,0] \to [-1,1]$

* The implementation hint is part of [S2].

from (3.14) are encountered that accomplish the scaling of the domain. However, such a scaling causes in the inner product of two functions $p, q \in L_2([-1,1],\mathbb{R}^n)$

$$\begin{aligned}\langle q(\vartheta(\cdot)), p(\vartheta(\cdot))\rangle_{L_2([-h,0],\mathbb{R}^n)} &= \int_{-h}^{0} q^\top(\vartheta(\theta))p(\vartheta(\theta))\,\mathrm{d}\theta \\ &= \int_{-1}^{1} q^\top(\vartheta)p(\vartheta)\tfrac{1}{\vartheta'}\,\mathrm{d}\vartheta = \tfrac{1}{\vartheta'}\langle q,p\rangle_{L_2([-1,1],\mathbb{R}^n)}\end{aligned}$$

only a constant factor. Consequently, in (3.30), the functions on the unscaled and unshifted domain can be used equally well in

$$\forall j \in \{0,\dots,N-1\}: \quad \left\langle \sum_{k=0}^{N} \dot{c}^k(t)p_k(\cdot) - \vartheta' \sum_{k=0}^{N} c^k(t)\,p_k'(\cdot), \quad p_j(\cdot) \right\rangle_{L_2([-1,1],\mathbb{R}^n)} = 0. \quad (3.31)$$

The second term incorporates the Legendre polynomial derivative p_k', which, for the sake of clarity, is explicitly stated in the following lemma (alternatively, [88, eq. (5.8)] could directly be used in (3.31)).

Lemma 3.5.2. *The derivative of the (unscaled) k-th Legendre polynomial is*

$$\frac{\mathrm{d}}{\mathrm{d}\vartheta}p_k(\vartheta) = \sum_{\substack{m=0 \\ m+k\ odd}}^{k-1} (2m+1)p_m(\vartheta). \quad (3.32)$$

(See, in Remark 3.6.2, the third item and, in Figure 3.6b, the Legendre differentiation matrix D_c, where $\vartheta' = 1$ corresponds to this unscaled result.)

Proof. According to [88, eq. (5.8)], $\phi(\vartheta) = \sum_{j=0}^{\infty} c^j p_j(\vartheta)$ and its derivative $\phi'(\vartheta) = \sum_{m=0}^{\infty} \gamma^m p_m(\vartheta)$ are related via $\gamma^m = (2m+1)\sum_{j=m+1,m+j\text{ odd}}^{\infty} c^j$. Using $c^j = \delta_{kj}$ with $\delta_{kj} = 1$ if $k = j$ and $\delta_{kj} = 0$ if $k \neq j$ yields (3.32). □

Hence, (3.31) becomes

$$\forall j \in \{0,\dots,N-1\}:$$
$$\left\langle \sum_{k=0}^{N} \dot{c}^k(t)p_k(\cdot) - \vartheta' \sum_{k=0}^{N} c^k(t) \sum_{m=0}^{N} (2m+1)p_m(\cdot) \begin{cases} 1 & \text{if } m \le k-1, \\ & m+k \text{ odd}, \\ 0 & \text{otherwise} \end{cases}, \; p_j(\cdot) \right\rangle_{L_2} = 0. \tag{3.33}$$

It is the gist of Legendre polynomials that $\langle p_k, p_j \rangle_{L_2} = 0$ if $j \neq k$, which, in fact, is the reason why the polynomial ansatz is represented in that basis when using the Legendre tau method.

Lemma 3.5.3 ([88, Sec. B.1]). *Legendre polynomials are orthogonal in L_2 with*

$$\langle p_i, p_j \rangle_{L_2} = \begin{cases} 0 & \text{if } i \neq j, \\ \frac{2}{2j+1} & \text{if } i = j, \end{cases} \tag{3.34}$$

where $\langle p_i, p_j \rangle_{L_2} = \int_{-1}^{1} p_i(\vartheta)p_j(\vartheta)\,\mathrm{d}\vartheta$.

As a consequence, only $k = j$ must be considered in the first sum of (3.33) and $m = j$ in the second sum. The induced factor $\langle p_j, p_j \rangle_{L_2} = \frac{2}{2j+1}$ is irrelevant for the equation and thus

$$\forall j \in \{0,\dots,N-1\}:$$
$$\dot{c}^j(t) - (2j+1)\vartheta' \sum_{k=0}^{N} c^k(t) \begin{cases} 1 & \text{if } j \le k-1 \text{ and } j+k \text{ odd}, \\ 0 & \text{otherwise} \end{cases} = 0 \tag{3.35}$$

explains all but the last row in (3.28), respectively (3.27).

Similar to the collocation method, the last row shall tackle the boundary condition. That is why tau methods are also called Galerkin methods with boundary bordering. In fact, the treatment of boundary conditions is the only difference between tau methods and the more complicated polynomial Galerkin methods,

where boundary conditions instead must be satisfied by construction of the basis polynomials. See [88, 65, 38, 23] for examples on both approaches.

Accordingly, the last row has to ensure that the ansatz $\sum_{k=0}^{N} c^k(t)p_k(\vartheta(\theta))$ from (3.25) satisfies the boundary condition (3.6b), which is actually the RFDE (1.4), and which becomes

$$\sum_{k=0}^{N} \dot{c}^k(t)p_k(\vartheta(0)) = A_0 \sum_{k=0}^{N} c^k(t)p_k(\vartheta(0)) + A_1 \sum_{k=0}^{N} c^k(t)p_k(\vartheta(-h)). \quad (3.36)$$

Legendre polynomials attain at the right boundary the value $p_k(\vartheta(0)) = p_k(1) = 1$ and at the left boundary the value $p_k(\vartheta(-h)) = p_k(-1) = (-1)^k$, see Figure 3.5. Thus, the above equation reads

$$\sum_{k=0}^{N} \dot{c}^k(t) = A_0 \sum_{k=0}^{N} c^k(t) + (-1)^k A_1 \sum_{k=0}^{N} c^k(t). \quad (3.37)$$

The latter combined with (3.35) already yields a suitable description of the dynamics (as an index-0 DAE). To obtain an ODE, note that the derivatives $\dot{c}^j(t)$ for $j \in \{0, \dots, N-1\}$ are already determined by the first N rows, derived in (3.35) above, whereas the last row shall give an expression for the remaining unknown $\dot{c}^N(t)$, which consequently is

$$\dot{c}^N(t) = -\sum_{j=0}^{N-1} \dot{c}^j(t) + A_0 \sum_{k=0}^{N} c^k(t) + (-1)^k A_1 \sum_{k=0}^{N} c^k(t). \quad (3.38)$$

Thus, only the first term on the right-hand side must be made explicit in terms of $c(t)$. The following result can easily be confirmed by calculating column sums in (3.28).

Lemma 3.5.4. *If $\dot{c}^j(t)$, for $j \in \{0, \dots, N-1\}$, obeys (3.35), then*

$$\sum_{j=0}^{N-1} \dot{c}^j(t) = \vartheta' \sum_{k=0}^{N} \frac{k(k+1)}{2} c^k(t). \quad (3.39)$$

Proof.

$$\sum_{j=0}^{N-1} \dot{c}^j(t) \overset{(3.35)}{=} \vartheta' \sum_{j=0}^{N-1} (2j+1) \sum_{k=0}^{N} c^k(t) \begin{cases} 1 & \text{if } j \leq k-1 \text{ and } j+k \text{ odd}, \\ 0 & \text{otherwise} \end{cases} \tag{3.40}$$

$$= \vartheta' \sum_{k=0}^{N} c^k(t) \sum_{j=0}^{k-1} (2j+1) \begin{cases} 1 & \text{if } j+k \text{ odd}, \\ 0 & \text{otherwise} \end{cases} \tag{3.41}$$

$$= \vartheta' \sum_{k=0}^{N} c^k(t) \sum_{\substack{j=0 \\ j+k \text{ odd}}}^{k-1} \big(j + (j+1)\big). \tag{3.42}$$

Since for $j = k-1$, the condition $j+k$ being odd is satisfied independently from k, the last sum simplifies to

$$\sum_{\substack{j=0 \\ j+k \text{ odd}}}^{k-1} \big(j + (j+1)\big) = \sum_{\substack{j=0 \\ j+k \text{ odd}}}^{k-1} j + \sum_{\substack{m=1 \\ m+k \text{ even}}}^{k} m \quad = \sum_{j=0}^{k} j, \tag{3.43}$$

which is the triangular number $\sum_{j=1}^{k} j = \frac{k(k+1)}{2}$. □

Using (3.39) in (3.38) explains the last row in (3.28), respectively the case $j = N$ in (3.27).

3.5.2 Discretization

It remains to clarify how to choose, in the context of the Legendre tau method, for a given function ϕ the discretization $c = (c^k)_{k \in \{0,\dots,N\}}$.

3.5.2.1 Legendre Series Truncation with Corrected Boundary Value

The searched vector c intends to be the Legendre coordinate representation of an N-th degree polynomial $\phi^{[N]}$ that approximates the function ϕ. Still, unless ϕ is already a polynomial of degree at most N, the searched $\phi^{[N]}$ will not straightforwardly be taken from the Legendre series[8] truncation

$$\phi(\theta) = \sum_{k=0}^{\infty} \tilde{c}^k p_k(\vartheta(\theta)) \approx \sum_{k=0}^{N} \tilde{c}^k p_k(\vartheta(\theta)) \tag{3.44}$$

$$\begin{aligned} \tilde{c}^k &:= \frac{1}{\langle p_k, p_k \rangle_{L_2}} \langle \phi(\vartheta^{-1}(\cdot)), p_k \rangle_{L_2} \\ &= \frac{2k+1}{2} \int_{-1}^{1} \phi(\vartheta^{-1}(\tilde{\vartheta})) p_k(\tilde{\vartheta}) \,\mathrm{d}\tilde{\vartheta} = \underbrace{\vartheta'}_{\frac{2}{h}} \frac{2k+1}{2} \int_{-h}^{0} \phi(\theta) p_k(\vartheta(\theta)) \,\mathrm{d}\theta. \end{aligned} \tag{3.45}$$

The latter would be the L_2-orthogonal projection to the space of polynomials of degree at most N (establishing, similar to (3.31), that the error is orthogonal to that space, i.e., $\forall k \in \{0, \dots, N\} : \langle \phi(\vartheta^{-1}(\cdot)) - \sum_{j=0}^{N} \tilde{c}^j p_j, p_k \rangle_{L_2} = 0$). However, an accordingly obtained polynomial would not necessarily coincide at $\theta = 0$ with the value $\phi(0)$—which is decisive if $\phi = x_0$ is an initial function, and thus $\phi(0) = x_0(0) = x(0)$ the initial solution value for the time-delay system, cf. Figure 3.1.

Instead, the projection to be chosen is the same as the one applied to the residual in the Legendre tau method. This projection, in fact, only involves the L_2-orthogonal projection to the space of polynomials of degree at most $N-1$, being uniquely described by one less coefficient $\tilde{c}^0, \dots, \tilde{c}^{N-1}$ from (3.45). The searched N-th degree polynomial $\phi^{[N]}$, which has an additional coefficient c^N,

[8] For any $\phi \in L_2$, the series is convergent in L_2 and convergent a.e. [157]. The coefficients $\tilde{c}^k$ defined in (3.45) can numerically be obtained in Matlab via `legcoeffs` from the Chebfun toolbox [59].

uses this remaining degree of freedom for a correction in the end point such that $\phi^{[N]}(0) = \phi(0)$, cf. [98]. Consequently,

$$\phi^{[N]}(\theta) = \sum_{k=0}^{N-1} \tilde{c}^k p_k(\vartheta(\theta)) + c^N p_N(\vartheta(\theta)) = \sum_{k=0}^{N} c^k p_k(\vartheta(\theta)), \tag{3.46}$$

$$c^k := \begin{cases} \tilde{c}^k & \text{if } k < N, \\ \hat{x} - \sum_{k=0}^{N-1} \tilde{c}^k & \text{if } k = N, \end{cases} \qquad \hat{x} := \phi(0), \tag{3.47}$$

where $\phi^{[N]}(0) = \phi(0)$ since $\phi^{[N]}(0) = \sum_{k=0}^{N} c^k p_k(\vartheta(0)) = \sum_{k=0}^{N} c^k = \hat{x}$ (based on the fact that all Legendre polynomials have the boundary value $p_k(\vartheta(0)) = p_k(1) = 1$, see Figure 3.5).

If ϕ is already a polynomial of degree N or less, $(\tilde{c}^k)_{k\in\{0,\dots,N\}} = (c^k)_{k\in\{0,\dots,N\}}$ is simply the unique Legendre coordinate representation of that polynomial $\phi = \phi^{[N]}$, cf. Figure 3.5j. As such it can also be derived by a change of basis from any other coordinate representation (see Section 3.6).

3.5.2.2 Discontinuous Interpretation

Despite of considering only continuous state functions ϕ, there are situations where a function with a jump discontinuity at $\theta = 0$ is of interest, e.g., as the limit of a sequence of continuous functions (in particular when considering $\sup_{\phi\in C}(\dots)$ or $\inf_{\phi\in C}(\dots)$ of some expression incorporating the functional $V(\phi)$ in the next chapter). In these cases, it is convenient to consider, instead of $\phi^{[N]}$ from (3.46), rather a piecewise defined $(N-1)$-th degree polynomial

$$\phi_{\mathrm{d}}^{[N]}(\theta) := \begin{cases} \bar{\phi}_{\mathrm{d}}^{[N]}(\theta) := \sum_{k=0}^{N-1} c^k p_k(\vartheta(\theta)) & \text{if } \theta \in [-h, 0), \\ \hat{x} & \text{if } \theta = 0, \end{cases} \tag{3.48}$$

with a discontinuous end point

$$\hat{x} = \sum_{k=0}^{N} c^k. \tag{3.49}$$

This approximating function $\phi_{\mathrm{d}}^{[N]}(\theta)$ achieves in (3.47) the same discretization c.

3.5.2.3 Associated Projection Operators

As a consequence, if $c = [(c^0)^\top, \ldots, (c^N)^\top]^\top \in \mathbb{R}^{n(N+1)}$ is built from discretizing a given function ϕ according to (3.47), then ϕ can be approximated by the discontinuous function $\phi_{\mathrm{d}}^{[N]}$ from (3.48) or by the continuous function $\phi^{[N]}$ from (3.46). Of course, there is an infinite number of alternative realizations or "reconstructions" in terms of functions that also provide the same discretization c in (3.47).

For the original function ϕ, consider the corresponding $\mathsf{x} := \begin{bmatrix} \phi \\ \phi(0) \end{bmatrix}$ in $M_2 = L_2 \times \mathbb{R}^n$, cf. Section 3.2.2 (the sans-serif typeface, e.g., $\mathsf{x}, \mathsf{y}, \mathsf{z}$, will henceforth be used for elements in M_2). What characterizes the discontinuous interpretation $\phi_{\mathrm{d}}^{[N]}$ in (3.48) is that the associated mapping[9]

$$\begin{aligned} \mathrm{Proj}_{\mathrm{d}}^{[N]}\colon \qquad & M_2 \to \mathrm{P}_{N-1}([-h,0],\mathbb{R}^n) \times \mathbb{R}^n \subset M_2; \\ \mathsf{x} = \begin{bmatrix} \phi(\cdot) \\ \phi(0) \end{bmatrix} & \mapsto \mathrm{Proj}_{\mathrm{d}}^{[N]}\mathsf{x} = \begin{bmatrix} \bar{\phi}_{\mathrm{d}}^{[N]}(\cdot) \\ \phi_{\mathrm{d}}^{[N]}(0) \end{bmatrix} \end{aligned} \tag{3.50}$$

is an M_2-orthogonal projection to the subspace of M_2 which in the first component involves a polynomial of degree at most $N-1$.

In contrast, what characterizes $\phi^{[N]}$ in (3.46) is only that this is the function represented by c, when the entries of c are understood as the Legendre coordinates of an N-th degree polynomial, i.e., straightforwardly as the coefficients of the 0-th

[9] See in [98] the projector Π^N or in [99] the projector Q^N.

to N-th Legendre polynomial. However, because of the correction $c^N \neq \tilde{c}^N$ in (3.47), the continuous $\phi^{[N]}$ from (3.46) stems from a projection

$$\begin{aligned} \mathrm{Proj}_{\mathrm{cont}}^{[N]}\colon \qquad & M_2 \to \mathrm{P}_N([-h,0],\mathbb{R}^n) \times \mathbb{R}^n \subset M_2; \\ \mathsf{x} = \begin{bmatrix} \phi(\cdot) \\ \phi(0) \end{bmatrix} & \mapsto \mathrm{Proj}_{\mathrm{cont}}^{[N]}\mathsf{x} = \begin{bmatrix} \phi^{[N]}(\cdot) \\ \phi^{[N]}(0) \end{bmatrix} \end{aligned} \tag{3.51}$$

that is no M_2-orthogonal projection to the stated subspace which in the first component involves a polynomial of degree at most N.

All that is irrelevant if ϕ is already a polynomial of degree at most $N-1$ since then $c^N = 0$, and the functions $\phi = \phi^{[N]} = \phi_{\mathrm{d}}^{[N]}$ coincide. In fact, (3.46) and (3.48) only associate a different basis function with the last coordinate c^N, see (A.3) vs. (A.8) in Appendix A. Therefore, a significant difference between $\phi^{[N]}$ and $\phi_{\mathrm{d}}^{[N]}$, which manifests in a jump discontinuity at the end point in (3.48), only occurs if a non-small last coordinate c^N is involved.

3.5.3 The Approximated Operators

The above discussed projection operators can be used to describe the operator $\mathscr{A}^{[N]}$ that is actually represented by the matrix A_c^L in (3.28) from the Legendre tau method and that intends to approximate the operator $\mathscr{A}\colon M_2 \supset D(\mathscr{A}) \to M_2$ from (3.12). There are two projection steps involved.

1. The projection of the argument $\mathsf{x} \in D(\mathscr{A})$ of $\mathscr{A}\mathsf{x}$: The result of that projection must correspond to the ansatz function $\phi^{[N]} \approx \phi$ that is used in the numerical approach and that is represented by the Legendre coordinates c taken in $A_c^L c$. The ansatz (3.25) is unambiguous. Incorporating how the involved coordinate vector c is derived from a given initial function in (3.47), it amounts to the polynomial (3.46) of degree N. The latter results from the projection $\mathrm{Proj}_{\mathrm{cont}}^{[N]}\mathsf{x}$ defined in (3.51). Thus, once the ansatz (3.25) is incorporated, no longer $\mathscr{A}$ but $\mathscr{A}\mathrm{Proj}_{\mathrm{cont}}^{[N]}$ is applied to $\mathsf{x} \in D(\mathscr{A})$.

2. The projection of the resulting residual, and thus the projection of the involved $\mathscr{A}\mathrm{Proj}_{\mathrm{cont}}^{[N]}\mathsf{x}$: The result of the latter is represented by $A_c^L c \in \mathbb{R}^{n(N+1)}$ forming the right-hand side of the ODE (where $c \in \mathbb{R}^{n(N+1)}$ are the coordinates of $\mathrm{Proj}_{\mathrm{cont}}^{[N]}\mathsf{x}$). Therefore, the projection operator depends on how to interpret that coordinate vector $A_c^L c$. The standard interpretation, cf. [73, eq. 2.16] or [99, eq. 3.13], is to view the resulting vector $A_c^L c$ again as representing the Legendre coordinates of an N-th degree polynomial. Then $\mathscr{A}^{[N]} = \mathscr{A}_{\mathrm{cc}}^{[N]}$,

$$\mathscr{A}_{\mathrm{cc}}^{[N]} = \mathrm{Proj}_{\mathrm{cont}}^{[N]}\, \mathscr{A}\, \mathrm{Proj}_{\mathrm{cont}}^{[N]}. \tag{3.52}$$

An alternative interpretation, however, is to view the resulting vector $A_c^L c$ as representing a discontinuous function (3.48). Then $\mathscr{A}^{[N]} = \mathscr{A}_{\mathrm{dc}}^{[N]}$,

$$\mathscr{A}_{\mathrm{dc}}^{[N]} = \mathrm{Proj}_{\mathrm{d}}^{[N]}\, \mathscr{A}\, \mathrm{Proj}_{\mathrm{cont}}^{[N]}, \tag{3.53}$$

cf. [99, eq. 3.15], which amounts to the M_2-orthogonal projection of the residual in M_2. In particular, convergence results from [99] refer to $\mathscr{A}_{\mathrm{dc}}^{[N]}$.

In terms of the coordinate representations, both interpretations $\mathscr{A}_{\mathrm{cc}}^{[N]}$ and $\mathscr{A}_{\mathrm{dc}}^{[N]}$ are indistinguishable. That is why, the chosen interpretation is not relevant for the results in the next chapters, which only rely on the coordinates.

Where, however, the operator $\mathscr{A}^{[N]}$ occurs in the derivations, the interpretation (3.53) will be meant (although even in the arising operator-valued equations of the next chapters, the choice does not make a difference, see Appendix A.3.3). The interpretation (3.53) has the great advantage that the outer projection $\mathrm{Proj}_{\mathrm{d}}^{[N]}$ is actually without effect: The inner projection $\left[\begin{smallmatrix}\tilde{\phi}\\ \tilde{\phi}(0)\end{smallmatrix}\right] := \mathrm{Proj}_{\mathrm{cont}}^{[N]}\left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right]$ in (3.53) results in a function $\tilde{\phi}$ that is a polynomial of degree at most N. Therefore, the first component $\tilde{\phi}'$ in $\mathscr{A}\left[\begin{smallmatrix}\tilde{\phi}\\ \tilde{\phi}(0)\end{smallmatrix}\right] = \left[\begin{smallmatrix}\tilde{\phi}'\\ A_0\tilde{\phi}(0)+A_1\tilde{\phi}(-h)\end{smallmatrix}\right]$ is already a polynomial of degree at most $N-1$. Therefore,

$$\begin{aligned}\mathscr{A}_{\mathrm{dc}}^{[N]} &= \mathrm{Proj}_{\mathrm{d}}^{[N]}\mathscr{A}\mathrm{Proj}_{\mathrm{cont}}^{[N]} && (3.54)\\ &= \mathscr{A}\mathrm{Proj}_{\mathrm{cont}}^{[N]} \qquad =: \mathscr{A}^{[N]}. && (3.55)\end{aligned}$$

The precise meaning of the statement that A_c^L is a coordinate representation of $\mathscr{A}^{[N]}$ is discussed in Appendix A.

If ϕ is already a polynomial of degree at most N, the remaining inner projection $\mathrm{Proj}_{\mathrm{cont}}^{[N]}$ in (3.55) is also without effect, and $\mathscr{A}^{[N]}\left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right]$ even gives the exact result $\mathscr{A}\left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right] = \left[\begin{smallmatrix}\phi'\\ A_0\phi(0)+A_1\phi(-h)\end{smallmatrix}\right]$. In terms of the coordinates, note that the rectangular upper part of the differentiation matrix D_c used in A_c^L from the Legendre tau method indeed suffices to obtain the exact coordinates of ϕ' since the omitted last row of of the differentiation matrix D_c is zero anyway, see Remark 3.6.2. Consequently, $\mathscr{A}^{[N]}$ and $\mathscr{A}$ coincide on the restriction to the polynomial subspace $\mathrm{range}(\mathrm{Proj}_{\mathrm{cont}}^{[N]}) \subset D(\mathscr{A})$ (which, however, does not yet imply a statement on the solutions from polynomial initial functions since that polynomial subspace is no invariant subspace—still, at least it becomes a dense subspace of the state space as N becomes arbitrarily large).

The following convergence statement is known for the operator sequence $\{\mathscr{A}^{[N]}\}_N$ with increasing discretization resolution N.

Lemma 3.5.5 (Convergence of $\mathscr{A}^{[N]}$, [99, Lem. 3.4]). *Let $D(\mathscr{A}^2)$ be the domain of the second power of $\mathscr{A}$, i.e., $D(\mathscr{A}^2) := \{\mathsf{x} \in D(\mathscr{A}) : \mathscr{A}\mathsf{x} \in D(\mathscr{A})\}$. Then*

$$\forall \mathsf{x} \in D(\mathscr{A}^2): \quad \|\mathscr{A}^{[N]}\mathsf{x} - \mathscr{A}\mathsf{x}\|_{M_2} \to 0 \quad \text{as } N \to \infty. \tag{3.56}$$

This convergence of $\mathscr{A}^{[N]}$ on the dense subset $D(\mathscr{A}^2) \subset D(\mathscr{A})$ is used in [99] to prove strong convergence of the resulting solution operator $\mathscr{T}^{[N]}\colon M_2 \to M_2$. The exact solution operator maps an initial function to the exact state at time t,

$$\left[\begin{smallmatrix}x_t\\ x(t)\end{smallmatrix}\right] = \mathscr{T}(t)\left[\begin{smallmatrix}x_0\\ x(0)\end{smallmatrix}\right], \tag{3.57}$$

and, correspondingly, the approximated solution operator $\mathscr{T}^{[N]}(t)$ maps to the resulting approximation of that state. This operator $\mathscr{T}^{[N]}(t)$ is represented by the matrix exponential $\exp(A_c^L t)$ from the finite-dimensional ODE approximation in the

same manner as $\mathscr{A}^{[N]}$ is represented by the system matrix A_c^L. Clearly, x_t is a continuous function, which is why—in alignment with the made polynomial ansatz—the continuous interpretation of the resulting coordinates $c(t)$ should actually be preferred (amounting to $\mathrm{Proj}_{\mathrm{cont}}\mathscr{T}^{[N]}$ generated by $\mathscr{A}_{\mathrm{cc}}^{[N]} = \mathrm{Proj}_{\mathrm{cont}}\mathscr{A}_{\mathrm{dc}}^{[N]}$). Nevertheless, as x_t is continuous, it can be expected that the last coordinate $c^N(t)$ of its respective approximation vanishes as N increases, and thus the difference between both interpretations vanishes as well. The convergence statement for $\mathscr{T}^{[N]}$ is as follows.

Lemma 3.5.6 (Convergence of $\mathscr{T}^{[N]}$, [99, Thm. 3.6]). *Let $\{\mathscr{T}^{[N]}(t)\}_{t\geq 0}$, be the semigroup generated by $\mathscr{A}^{[N]}$. Then*

$$\forall \mathsf{x} \in M_2 : \quad \|\mathscr{T}^{[N]}(t)\mathsf{x} - \mathscr{T}(t)\mathsf{x}\|_{M_2} \to 0 \quad \textit{as } N \to \infty \tag{3.58}$$

uniformly on bounded intervals of $t \geq 0$.

Even more, the same statement of strong convergence also holds for the adjoint $\mathscr{T}^*(t)$ of the solution operator [97, Thm. 2.2]. Such a convergence result cannot be taken for granted. For other discretization schemes only weak convergence holds for the adjoint operator, which causes trouble when trying to compute control gain operators from operator-valued algebraic Riccati equations in optimal control problems [35, 13]. These operator equations, however, are related to what will be employed in the present thesis. That is why, the Legendre tau method is a particularly promising choice of discretization scheme.

3.6 Change of Basis for Polynomials

Given a polynomial of degree N, it can equivalently be represented in the Lagrange interpolation basis w.r.t. Chebyshev nodes, where a polynomial is written as $\phi(\theta) = \sum_{k=0}^{N} y^k \ell_k(\vartheta(\theta))$, see Figure 3.3j, or in the Legendre basis, where the same polynomial is written as $\phi(\theta) = \sum_{k=0}^{N} c^k p_k(\vartheta(\theta))$, see Figure 3.5j, or in

the monomial basis, where $\phi(\theta) = a_0 + a_1\theta + a_2\theta^2 + \ldots + a_N\theta^N = \sum_{k=0}^{N} a_k\theta^k$, or in any other basis of this $(N+1)$-dimensional space.

The coefficients are related by a linear transformation, e.g.,

$$y = T_{yc}\, c, \qquad \text{respectively} \quad c = T_{cy}\, y \quad \text{where } T_{cy} := T_{yc}^{-1}, \tag{3.59}$$

with a transformation matrix $T_{yc} \in \mathbb{R}^{n(N+1)\times n(N+1)}$. Consequently, if the evolution of $c(t)$ is described by a linear autonomous ODE, then the transformed coordinates $y(t)$ obey an ODE from a similarity transform of the system matrix according to

$$\dot{c} = A_c c \qquad \Leftrightarrow \qquad \dot{y} = \underbrace{T_{yc} A_c T_{cy}}_{A_y} y. \tag{3.60}$$

3.6.1 Transformation to Interpolation Coordinates

Consider the equivalence of a polynomial represented in Lagrange interpolation coordinates and in Legendre coordinates, $\phi(\theta) = \sum_{k=0}^{N} y^k \ell_k(\vartheta(\theta)) = \sum_{k=0}^{N} c^k p_k(\vartheta(\theta))$. As a consequence of that equivalence, the j-th interpolation coordinate, which is the pointwise evaluation $y^j = \phi(\tilde{\theta}_j)$, must be equal to $\phi(\tilde{\theta}_j) = \sum_{k=0}^{N} c^k p_k(\vartheta(\tilde{\theta}_j))$. Thus, since the searched transformation law amounts to $y^j = \sum_{k=0}^{N} T_{yc}^{jk} c_k$, the (j,k)-th (block-) entry of the transformation matrix is

$$T_{yc}^{jk} = p_k(\vartheta(\tilde{\theta}_j)) I_n. \tag{3.61}$$

The first ($j = 0$) and the last ($j = N$) block row of T_{yc} are particularly simple. By $p_k(-1) = (-1)^k$ and $p_k(1) = 1$ (see Figure 3.5), these become visible in

$$\begin{bmatrix} y^0 \\ y^N \end{bmatrix} = \begin{bmatrix} I_n & -I_n & \cdots & (-1)^N I_n \\ I_n & I_n & \cdots & I_n \end{bmatrix} c. \tag{3.62}$$

The remaining rows can numerically be derived from the following implementation.

Remark 3.6.1 (Implementation of A_y^L)*. *Efficient conversion algorithms [176] between Lagrange interpolation coordinates y and Legendre coordinates c are, e.g., available in the Chebfun toolbox [59]. Applying these to the identity matrix yields T_{yc}, respectively its inverse T_{cy}. As a result, $A_y^L = \mathtt{Ay}$ can be obtained by appending the lines*

```
Tyc=kron(legcoeffs2chebvals(eye(N+1)),eye(n));
Tcy=kron(chebvals2legcoeffs(eye(N+1)),eye(n));
Ay=Tyc*Ac*Tcy
```

to the code that is given in Remark 3.5.1.

Lagrange interpolation coordinates are most descriptive, but the price to pay is that, also for the Legendre tau approach, the resulting system matrix

$$A_y^L := T_{yc}\, A_c^L\, T_{cy}, \tag{3.63}$$

based on A_c^L from (3.27), becomes dense. See Table 3.2 for further comparisons of the coordinate choices. What will be important in the next chapter is that, in contrast to the original Legendre coordinates, the decisive boundary value $\phi(0)$ is explicitly given by the last coordinate y^N. This advantage, however, is also achieved by the following concept of mixed coordinates.

* The implementation hint is part of [S2].

Lagrange interpolation coordinates y in Chebyshev nodes	**Legendre coordinates c**	**Mixed coordinates χ**
⊕ Very descriptive, built from the pointwise evaluation of the polynomial in Chebyshev nodes, see Figure 3.3j	⊕ If the polynomial is of degree $K < N$ then only the first K coordinates are nonzero, see Figure 3.5j	
⊕ $\phi(0)$ is the coordinate y^N	⊖ $\phi(0) = \sum_{k=0}^{N} c^k$ is not explicitly among the coordinates	⊕ $\phi(0)$ is given by the coordinate $\chi^N = \hat{x}$
⊕ The pointwise evaluation of a non-polynomial function in the Chebyshev nodes ↓ immediately gives the Lagrange coordinates of its interpolating polynomial	⊕ Legendre series truncation of a non-polynomial function ↓ immediately gives the Legendre coordinates $\tilde{c}$ of its L_2-best approximation polynomial ⊖ The projection from (3.47) involves a correction step for the last coordinate $c^N \neq \tilde{c}^N$ that must be calculated additionally	⊕ The calculation of c^N in (3.47) is not required ⊕ Also appropriate for the representation of polynomials of degree $N-1$ with a discontinuous endpoint, see (3.48) ⊕ In that discontinuous interpretation (3.48), the coordinates are associated with a basis $\{\mathsf{h}_{\chi,k}\}_k$ that is an orthogonal basis in M_2, see Appendix A
	! Despite of $\{p_k(\vartheta(\cdot))\}_k$ being an orthogonal basis in L_2, the associated basis $\{\mathsf{g}_{c,k}\}_k$ of the coordinates c in M_2, which is described in Appendix A, is no orthogonal basis in M_2	! In the continuous interpretation, the associated basis $\{\mathsf{g}_{\chi,k}\}_k$ of the coordinates χ in M_2, which is described in Appendix A, is still no orthogonal basis in M_2
⊕ Natural coordinates in the Chebyshev collocation method	⊕ Natural coordinates in the Legendre tau method	
⊖ The system matrices A_y^C and A_y^L in Lagrange interpolation coordinates are dense with non-integer entries	⊕ All entries of A_c^L from (3.28) are integers (neglecting the factor $\frac{1}{h}$) and, for large N and $n = 1$, almost 3/4 of the entries are 0	⊖ A_χ^L in mixed coordinates is less simply structured and more dense than A_c^L ⊕ All entries of A_χ^L are integers (neglecting the factor $\frac{1}{h}$) and, for large N and $n = 1$, almost half of the entries are 0

Table 3.2: Advantages and disadvantages of the discussed coordinate choices.

3.6.2 Transformation to Mixed Coordinates

To take the special role of the boundary value $\hat{x} = \phi(0)$ into account without completely transforming to interpolation coordinates $y = \left[\begin{smallmatrix} z \\ \hat{x} \end{smallmatrix}\right] = T_{yc}c$, the combination of the first $(N-1)$ of the N Legendre coordinates and the boundary value $\hat{x} = \phi(0)$,

$$\chi = \begin{bmatrix} c^0 \\ \vdots \\ c^{N-1} \\ \hat{x} \end{bmatrix} = \underbrace{\begin{bmatrix} I_n & & & \\ & \ddots & & \\ & & I_n & \\ I_n & \cdots & I_n & I_n \end{bmatrix}}_{T_{\chi c}} c, \tag{3.64}$$

is an appropriate choice of coordinates for the result from the Legendre tau method[10]. In particular, when the discontinuous interpretation from Section 3.5.2.2 is of interest, these mixed coordinates χ are the coordinates of choice. The resulting system matrix is derived from

$$A_\chi^L = T_{\chi c}\, A_c^L\, T_{c\chi}, \qquad \text{with } T_{c\chi} = T_{\chi c}^{-1} = \begin{bmatrix} I_{nN} & 0_{nN\times n} \\ -1_N^\top \otimes I_n & I_n \end{bmatrix}. \tag{3.65}$$

A major advantage compared to (3.63) is that the system parameters A_0 and A_1, which before the similarity transform only occur in the last row of A_c^L, after the similarity transform still only affect the last (block-) row of A_χ^L in (3.65). The upper part of A_χ^L again only relies on the corresponding differentiation matrix D_χ, the structure of which is given in Figure 3.6, see the following remark. Thus, A_χ^L clearly has an advantageous structure compared to the transformation to a Lagrange interpolation coordinate representation in (3.63), which is as dense as A_y^C from the collocation method. Nevertheless, A_χ^L in mixed coordinates is

[10] For the sake of convenience: The Matlab implementation of (3.64) is simply `Txc=eye(N+1); Txc(end,:)=1; Txc=kron(Txc,eye(n)); Tcx=inv(Txc); Ax=Txc*Ac*Tcx`

slightly more dense and less simply structured than the original result A_c^L in Legendre coordinates from (3.28).

Remark 3.6.2 (Differentiation Matrices). *Let $\phi\colon [-h,0] \to \mathbb{R}$ be a polynomial of degree N or less, represented by the $N+1$ Lagrange interpolation coordinates $y \in \mathbb{R}^{N+1}$, or alternatively Legendre coordinates $c \in \mathbb{R}^{N+1}$, mixed coordinates $\chi \in \mathbb{R}^{N+1}$, or the coordinates associated to any other basis of that $N+1$ dimensional space. Its derivative ϕ' is a polynomial of degree $N-1$ and thus can still be represented in the same basis in an exact manner. In terms of the $N+1$ coordinates, the differentiation operator becomes a differentiation matrix. As a result, the representation of the derivative ϕ' in Lagrange interpolation coordinates is derived from $D_y y$, in Legendre coordinates it is derived from $D_c c$, and in mixed coordinates it is derived from $D_\chi \chi$, in each case relying on a differentiation matrix $D_y, D_c, D_\chi, \ldots \in \mathbb{R}^{(N+1)\times(N+1)}$. Figure 3.6 shows how these matrices look like.*

- *Since the $(N+1)$-th derivative of the polynomial of degree N is zero, the differentiation matrices are nilpotent matrices with $(D_y)^{N+1} = (D_c)^{N+1} = (D_\chi)^{N+1} = 0_{(N+1)\times(N+1)}$ and therefore have only zero eigenvalues.*
- *The differentiation matrices are related by the above discussed change of basis. For instance,*

$$D_y = T_{yc} D_c T_{cy}, \qquad T_{yc} = T_{cy}^{-1} \tag{3.66}$$

  describes the relation between D_y for Lagrange interpolation coordinates in Chebyshev nodes and D_c for Legendre coordinates.

- *The $(k+1)$-th column De_{k+1}, $k \in \{0,\ldots,N\}$, of the differentiation matrix $D \in \{D_y, D_c, D_\chi, \ldots\}$ describes the coordinates of the derivative of the k-th (including $k=0$) basis polynomial that is represented by the coordinate vector e_{k+1}. Since the k-th Legendre polynomial p_k is a polynomial of degree k, the Legendre differentiation matrix D_c must be triangular. (The same holds when Chebyshev polynomials T_k—see*

$$D_y = \vartheta' \begin{bmatrix} -21.5000 & 26.2741 & -6.8284 & 3.2398 & -2 & 1.4465 & -1.1716 & 1.0396 & -0.5000 \\ -6.5685 & 3.1543 & 4.6131 & -1.8478 & 1.0824 & -0.7654 & 0.6131 & -0.5412 & 0.2599 \\ 1.7071 & -4.6131 & 0.7071 & 3.0824 & -1.4142 & 0.9176 & -0.7071 & 0.6131 & -0.2929 \\ -0.8100 & 1.8478 & -3.0824 & 0.2242 & 2.6131 & -1.3066 & 0.9176 & -0.7654 & 0.3616 \\ 0.5000 & -1.0824 & 1.4142 & -2.6131 & 0 & 2.6131 & -1.4142 & 1.0824 & -0.5000 \\ -0.3616 & 0.7654 & -0.9176 & 1.3066 & -2.6131 & -0.2242 & 3.0824 & -1.8478 & 0.8100 \\ 0.2929 & -0.6131 & 0.7071 & -0.9176 & 1.4142 & -3.0824 & -0.7071 & 4.6131 & -1.7071 \\ -0.2599 & 0.5412 & -0.6131 & 0.7654 & -1.0824 & 1.8478 & -4.6131 & -3.1543 & 6.5685 \\ 0.5000 & -1.0396 & 1.1716 & -1.4465 & 2 & -3.2398 & 6.8284 & -26.2741 & 21.5000 \end{bmatrix}$$

(a) D_y for Lagrange interpolation coordinates w.r.t. Chebyshev nodes

$$D_c = \vartheta' \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 3 & 0 & 3 & 0 & 3 & 0 & 3 \\ 0 & 0 & 0 & 5 & 0 & 5 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 7 & 0 & 7 & 0 & 7 \\ 0 & 0 & 0 & 0 & 0 & 9 & 0 & 9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 11 & 0 & 11 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 13 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 15 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(b) D_c for Legendre coordinates

$$D_u = \vartheta' \begin{bmatrix} 0 & 1 & 0 & 3 & 0 & 5 & 0 & 7 & 0 \\ 0 & 0 & 4 & 0 & 8 & 0 & 12 & 0 & 16 \\ 0 & 0 & 0 & 6 & 0 & 10 & 0 & 14 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 & 12 & 0 & 16 \\ 0 & 0 & 0 & 0 & 0 & 10 & 0 & 14 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 12 & 0 & 16 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 14 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 16 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(c) D_u for Chebyshev coordinates

$$D_\chi = \vartheta' \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ -3 & -3 & 0 & -3 & 0 & -3 & 0 & -3 & 3 \\ 0 & 0 & 0 & 5 & 0 & 5 & 0 & 5 & 0 \\ -7 & -7 & -7 & -7 & 0 & -7 & 0 & -7 & 7 \\ 0 & 0 & 0 & 0 & 0 & 9 & 0 & 9 & 0 \\ -11 & -11 & -11 & -11 & -11 & -11 & 0 & -11 & 11 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 13 & 0 \\ -15 & -15 & -15 & -15 & -15 & -15 & -15 & -15 & 15 \\ -36 & -35 & -33 & -30 & -26 & -21 & -15 & -8 & 36 \end{bmatrix}$$

(d) D_χ for the mixed coordinates from Section 3.6.2

Figure 3.6: Differentiation matrices for different coordinate choices with $N = 8$. The constant scaling factor $\vartheta' = \frac{2}{h}$ arises due to (3.14).

Figure 3.2b—are used as basis, with the differentiation matrix for the associated Chebyshev coordinates[11] u being denoted by D_u in Figure 3.6.)

Since both in the Chebyshev collocation and in the Legendre tau method the last row of the corresponding differentiation matrix is replaced by the boundary condition (addressing the RFDE), the resulting system matrices A_y^C and A_c^L are, in contrast to D_y and D_c, no similar matrices.

Remark 3.6.3 (The τ in Lanczos' tau method). *The term tau method originates from a variable τ in Lanczos' original paper [122, eq. 8.15]. As discussed above, the overall differentiation matrix D_c indeed maps the coordinates c of an N-th order polynomial to the exact coordinates $D_c c$ of its derivative. Thus, on the restriction to polynomials of degree at most N, with x_t being represented by its Legendre coordinates $c(t)$, the exact coordinate representation of (3.6a) becomes*

$$\dot{c}(t) = (D_c \otimes I_n)c(t). \tag{3.67}$$

However, the boundary condition must somehow be incorporated, which is accomplished by replacing the last row of the above equation—see (3.28). Hence, although $r_c(t) = \dot{c}(t) - (D_c \otimes I_n)c(t)$ still represents the correct residual of (3.6a), the Legendre tau method only achieves that the first N of the $N+1$ rows of $r_c(t)$ vanish, and some error $\tau(t) \neq 0$ occurs in the last row of

$$r_c(t) = \dot{c}(t) - (D_c \otimes I_n)c(t) = \begin{bmatrix} 0_{nN\times 1} \\ \tau(t) \end{bmatrix} \tag{3.68}$$

(more explicitly, since the last row of D_c is zero, $\tau(t) = \dot{c}^N(t) = (A_c^L)^{(N,:)}c(t)$ from (3.38)). Note that $r_c(t)$ is a coordinate representation of the residual.

[11] Chebyshev coordinates u can efficiently be computed from Lagrange interpolation coordinates y based on the fast Fourier transform (FFT), see [179, 177, 88]. Similar to Remark 3.6.1, a ready to use Matlab implementation is available from `chebvals2chebcoeffs` in [59].

Therefore, in the conventional continuous interpretation (3.46), the Legendre coordinates (3.68) represent the polynomial residual

$$\sum_{k=0}^{N} r^k(t) p_k(\vartheta(\theta)) = \tau(t) p_N(\vartheta(\theta)). \tag{3.69}$$

Thus, instead of the desired $\frac{\partial x_t(\theta)}{\partial t} - \frac{\partial x_t(\theta)}{\partial \theta} = 0$ *in (3.6a), the tau method solves the modification of (3.6a)*

$$\frac{\partial x_t(\theta)}{\partial t} - \frac{\partial x_t(\theta)}{\partial \theta} = \tau(t) p_N(\vartheta(\theta)) \tag{3.70}$$

in an exact manner for the restriction to polynomials of degree at most N *(see also [38, Sec. 6.6], [154, Sec. 3.4.3],[73]). This point of view was taken by Lanczos and gave the tau method in [123] its name.*

Remark 3.6.4 (The τ in the discontinuous interpretation). *Noticeably, in the discontinuous interpretation from (3.53), the above arguments do not apply. Instead, the coordinates on the right-hand side of (3.68) yield, according to (3.48), the residual*

$$\frac{\partial x_t(\theta)}{\partial t} - \frac{\partial x_t(\theta)}{\partial \theta} = \begin{cases} 0_{n\times 1} & \text{if } \theta \in [-h, 0), \\ \tau(t) & \text{if } \theta = 0, \end{cases} \tag{3.71}$$

which vanishes in the overall domain $[-h, 0)$*. However, the vanishing original residual on the considered polynomial subspace does not yet mean that the original problem was solved exactly for initial functions that are polynomials of degree at most* N *(see also the discussion on* $\mathscr{A}_{\mathrm{dc}}^{[N]}$ *in Section 3.5.3)—these polynomials do not form an invariant subspace in the original problem, in contrast to the modified problem from Remark 3.6.3 that amounts to* $\mathscr{A}_{\mathrm{cc}}^{[N]}$ *from Section 3.5.3.*

3.7 Stability Properties of the Approximating ODEs*

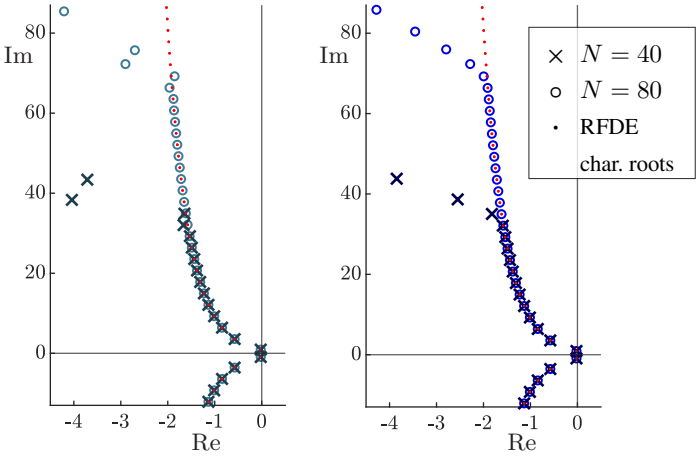
To conclude relevant stability statements for the RFDE from the ODE approximation, the ODE-approximation scheme should be stability preserving in the following sense.

Condition 3.7.1 (Stability Preservation Property). *Provided the discretization resolution N is chosen sufficiently large, the zero equilibrium in $\mathbb{R}^{n(N+1)}$ of the approximating ODE is exponentially stable if and only if the zero equilibrium in $C([-h,0],\mathbb{R}^n)$ of the RFDE is exponentially stable.*

Chebyshev collocation has successfully been applied in various fields [29, 28, 27, 192, 140] where Condition 3.7.1 is also desirable. It is known that eigenvalues of A_y^C converge to the characteristic roots of the RFDE, i.e., to the solutions s of $\det(sI_n - A_0 - \mathrm{e}^{-sh}A_1) = 0$, or, equivalently, to the eigenvalues of the infinitesimal generator of the C_0-semigroup of solution operators, see [29]. The red points in Figure 3.7a show typical eigenvalue chains in RFDEs, and the crosses and circles demonstrate how this chain is approached by the eigenvalues of A_y^C. There are also some additional spurious eigenvalues that do not match with RFDE characteristic roots. These, however, are easily identifiable as they do not persist when N changes [29, Prop. 3.7]. See, in Figure 3.7a, the crosses ($N = 40$) that do not match with circles ($N = 80$). Moreover, from numerical observations, they are not expected to hamper Condition 3.7.1, see also the discussions in [192, p. 361], [140, p. 853]. Thus, despite of not being proven, Condition 3.7.1, in practice, is a tenable assumption for the Chebyshev collocation method.

The Legendre tau method is similarly powerful in approximating eigenvalues, which is shown in Figure 3.7b. For the underlying rational approximation of $(sI - A_0 - \mathrm{e}^{-hs}A_1)^{-1}$, see Section 7.3 (choosing there $B = C_0 = I_n, p_1 = 0$).

* The author has prepublished Section 3.7 in [S2], ©2024 IEEE.



(a) Chebyshev collocation: eigenvalues of A_y^C in (3.16)

(b) Legendre tau: eigenvalues of A_c^L in (3.27) or, equivalently, A_y^L (3.63)

Figure 3.7*: Eigenvalues of the system matrix in the ODE approximation compared to the characteristic roots of the RFDE defined in (1.6). (The plotted example is a scalar system with the parameters $A_0 = -0.5, A_1 = -1, h = 2.2$.)

* The author has prepublished Figure 3.7 in [S1], ©2023 IEEE.

Stability preservation of the Legendre tau method (Condition 3.7.1) is proven in [97, Thm. 5.3].

3.8 Revisiting the Main Points of the Chapter

- A discretization of the abstract ODE (3.9) or (3.11), respectively a spatial discretization of the underlying PDE from Remark 3.2.1, yields the desired ODE approximation.
- Spectral methods like Chebyshev collocation or Legendre tau are based on a polynomial ansatz for $x_t(\theta)$. To be more precise, for any time $t \geq 0$, the state x_t, which represents the solution segment from Figure 3.1, is approximated by a polynomial on $\theta \in [-h, 0]$.
- The ODE approximation describes how the coordinates of that polynomial evolve with time.
- The resulting system matrices of the ODE approximations are given in (3.16) and (3.28). They can be implemented via few lines of code, see Remark 3.4.1 and Remark 3.5.1.
- The Chebyshev collocation method naturally uses Lagrange interpolation coordinates y to describe the polynomial, and the Legendre tau method naturally derives a description in Legendre coordinates c. Figure 3.3 and Figure 3.5 visualize the underlying polynomial bases.
- Let an initial function $x_0 = \phi \in C$ be given. The corresponding initial vector $y(0)$ for the Chebyshev-collocation-based ODE is obtained from pointwise evaluations of that function, see (3.23). The corresponding initial vector $c(0)$ for the Legendre-tau-based ODE is obtained from a Legendre series truncation with a modification in the last coordinate c^N, see (3.47).
- By a simple change of basis, results can be stated in the coordinates of any arbitrary basis of the polynomial space. In terms of the resulting ODE,

only a similarity transformation of the involved system matrix is required (Remark 3.6.1 gives a corresponding implementation hint).

- Table 3.2 shows advantages and disadvantages of the various coordinate choices.
- Lagrange interpolation coordinates y refer to pointwise values of the meant polynomial—see the red points in Figure 3.1c. Conversely, given a non-polynomial continuous function ϕ, only these pointwise evaluations (3.23) are needed to represent a polynomial approximation, namely the interpolation through these points. That these points are not equidistant, but Chebyshev nodes, helps to avoid the so-called Runge phenomenon from Figure 3.4.
- Concerning the coordinate choice for the result of the Legendre tau method the following can be recommended: It should be considered in
  - its original Legendre coordinates whenever a simple and relatively sparse structure of the ODE system matrix is important,
  - mixed coordinates (3.64) whenever the discontinuous reconstruction from Section 3.5.2.2 and the accompanying orthogonality property in M_2 become relevant (see also Appendix A),
  - Lagrange interpolation coordinates, whenever the simplicity of getting a coordinate representation from a given polynomial and the descriptiveness is important.
- Note the analogy in the structure of the following system matrices of the obtained ODEs:
  - the Chebyshev collocation system matrix A_y^C in its original Lagrange interpolation coordinates (3.16),
  - the Legendre tau system matrix A_c^L in its original Legendre coordinates (3.28), and

  - the Legendre tau system matrix A_χ^L when transformed to mixed coordinates (3.65).

  In each case, the lower (block-) row of the ODE system matrix addresses the RFDE (in terms of the polynomial ansatz for the state in the corresponding coordinates), whereas the upper part of the system matrix only relies on (all but the last row of) the differentiation matrix D_y, D_c, respectively D_χ, given in Figure 3.6.

- Stability preservation (Section 3.7) means that, for a sufficiently large discretization resolution, the equilibrium of the ODE approximation is exponentially stable if and only if the equilibrium of the RFDE is exponentially stable. For Chebyshev collocation, this is a tenable assumption. For the Legendre tau method, stability preservation is proven in the literature.

- Table 3.1 classifies the polynomial methods used in the present thesis. Besides of the spectral methods for differential equations, the underlying polynomial approximation of continuous functions also gives rise to powerful numerical integration methods that will be taken up in Section 4.2.

4 What ODE-Approximation Schemes Reveal about Lyapunov–Krasovskii Functionals

The present chapter proposes a numerical approach to complete-type and related LK functionals that is based on the ODE approximations discussed above. In contrast to existing approaches, the proposed procedure directly leads to an approximation of the overall LK functional—not making use of the integral formula (1.15)—thus not requiring the computation of the delay-Lyapunov matrix function—but only being based on the defining equation (1.14) in terms of Q_0, Q_1, Q_2. The objective does not lie in a stability criterion, but, as outlined in Chapter 1, rather in the functional itself and in its lower bound (1.16). Moreover, an interpretation of the results in terms of Lyapunov–Rumyantsev partial stability of the approximating ODE will hopefully provide an enlightening view on the Lyapunov–Krasovskii theory.

The chapter is organized as follows. **Section 4.1** introduces the numerical approach that is based on the ODE approximations discussed in the last chapter. For the sake of validation, in **Section 4.2**, numerical integration schemes are applied to the known semi-analytical result of the desired functionals. Then **Section 4.3** derives the formula for the quadratic lower bound. An example is discussed in **Section 4.4**. Additionally, **Section 4.5** interprets the approach in terms of partial stability of the approximating ODE. Finally, **Section 4.6** addresses the question of convergence. A summary of the chapter is provided in **Section 4.7**.

The following contributions of the present thesis are prepublished in

[S2] Scholl, T. H.; Hagenmeyer, V.; Gröll, L.: What ODE-approximation schemes of time-delay systems reveal about Lyapunov–Krasovskii functionals. IEEE Transactions on Automatic Control 69 (2024) 7, 4614–4629.

4.1 A Numerical Approach to Complete-Type and Related LK Functionals

The discussions shall initially not be confined to a specific discretization method. Therefore, and for simplicity, mainly Lagrange interpolation coordinates are used. Thus, the coordinates $y(t)$ approximate the state x_t in the pointwise manner (3.1), which is depicted in Figure 3.1. Nevertheless, all results are invariant under a change of coordinates, see Section 3.6 and Appendix A.

4.1.1 An Approximation Scheme for the LK Functional*

In the following, a Lyapunov function $V_y\colon \mathbb{R}^{n(N+1)} \to \mathbb{R}$ (to be more precise, a partial Lyapunov function, see Section 4.5) will be set up for the approximating ODE $\dot{y} = A_y\, y$ from (3.3). In view of the fact that the LK functional $V(x_t)$ is quadratic in terms of its argument $x_t \in C$, a quadratic ansatz

$$V_y(y) = y^\top P_y\, y \tag{4.1}$$

* The author has prepublished Section 4.1.1 in [S2], ©2024 IEEE.

in the polynomial coordinates y is made, with $P_y = P_y^\top \in \mathbb{R}^{n(N+1)\times n(N+1)}$ to be determined. The derivative of V_y along solutions of the ODE $\dot{y} = A_y y$ is set as $-y^\top Q_y y$ with a prescribed symmetric matrix Q_y

$$D^+_{(\dot{y}=A_y y)} V_y(y) = y^\top (P_y A_y + A_y^\top P_y) y \stackrel{!}{=} -y^\top Q_y y, \tag{4.2}$$

$\forall y \in \mathbb{R}^{n(N+1)}$. Thus, solving the Lyapunov equation

$$P_y A_y + A_y^\top P_y = -Q_y \tag{4.3}$$

yields the unknown matrix P_y. The right-hand side of (4.2) shall be constructed according to a discretization of the right-hand side of the defining equation of complete-type and related LK functionals from (1.14), i.e., of $D_f^+ V(x_t) = -x^\top(t)Q_0 x(t) - x^\top(t-h)Q_1 x(t-h) - \int_{-h}^0 x^\top(t+\theta)Q_2 x(t+\theta)\,\mathrm{d}\theta$, with freely chosen matrices $Q_0, Q_1 \succ 0_{n\times n}$, $Q_2 \succeq 0_{n\times n}$. Hence, in view of (3.1), a straightforward choice of Q_y in (4.3) becomes visible from

$$\begin{aligned} D^+_{(\dot{y}=A_y y)} V_y(y) &\stackrel{!}{=} -(y^N)^\top Q_0 y^N - (y^0)^\top Q_1 y^0 - \sum_{k=0}^{N} (y^k)^\top Q_2 y^k w_k \\ &= -y^\top \left(\begin{bmatrix} Q_1 & & & & \\ & 0_{n\times n} & & & \\ & & \ddots & & \\ & & & 0_{n\times n} & \\ & & & & Q_0 \end{bmatrix} + \begin{bmatrix} w_0 Q_2 & & & & \\ & w_1 Q_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & w_N Q_2 \end{bmatrix} \right) y \\ &=: -y^\top Q_y y\,, \end{aligned} \tag{4.4}$$

where $w_k \in \mathbb{R}$ are integration weights (see Section 4.2.1). Section 4.6.1 will present discretization-scheme-dependent modifications that aim at improved convergence properties. However, if the Legendre tau method is used, and $Q_2 = 0_{n\times n}$ (which, in view of the next chapters, is of prevalent interest), no modification is needed.

Altogether, since the above described approach amounts to solving a discretization of the original problem (1.14), $V_y(y)$ in (4.1) is intended to be an approximation of the LK functional $V(\phi)$. Convergence aspects will be addressed in Section 4.6, with a precise convergence statement for the Legendre-tau-based approach. As a result, given a prescribed argument $\phi \in C([-h, 0], \mathbb{R}^n)$, which might be $\phi = x_t$ for some $t \geq 0$, or, without loss of generality, $\phi = x_0$ at $t = 0$, a numerical approximation for the evaluation $V(\phi)$ can be obtained. To this end, the argument y in $V_y(y)$ must be chosen correspondingly. In the Chebyshev collocation method, this discretization y of ϕ is derived from stacking pointwise evaluations of ϕ as described in (3.23). If ϕ is a polynomial of degree at most N, (3.23) also agrees with the coordinate transform (3.59) of the discretization in the Legendre tau method (3.47). Otherwise, the latter might give a slightly deviating vector y (pointwise evaluations of the approximating polynomial $\phi^{[N]}$). Still, in practice, the simple construction of y from (3.23) can also be combined with the Legendre tau method since it represents, due to Remark 3.4.2, usually a very good polynomial approximation of ϕ (which can be inspected beforehand).

To sum up, only the Lyapunov equation (4.3) must be solved to obtain the approximation $V(\phi) \approx V_y(y)$.

Lyapunov Equation in Changed Coordinates

In the case of the Legendre tau approach, the transformed system matrix $A_y = A_y^L$ in Lagrange interpolation coordinates from (3.63) has to be used in the above Lyapunov equation (4.3). However, despite of being analytically equivalent, other coordinate choices such as Legendre coordinates or mixed coordinates, for which the system matrix is less dense, might be preferable from a numerical point of view.

For instance, in Legendre coordinates, the above approach becomes

$$V_c(c) := c^\top P_c c, \quad P_c = P_c^\top \in \mathbb{R}^{n(N+1)\times n(N+1)} \tag{4.5}$$
$$D^+_{(\dot c = A_c c)} V_c(c) = c^\top (P_c A_c + A_c^\top P_c) c \stackrel{!}{=} -c^\top \underbrace{(T_{yc}^\top Q_y T_{yc})}_{Q_c} c,$$

$\forall c \in \mathbb{R}^{n(N+1)}$ with T_{yc} from (3.59). That is,

$$P_c A_c + A_c^\top P_c = -T_{yc}^\top Q_y T_{yc} \tag{4.6}$$

is solved for P_c. If desired, the result can finally still be expressed in terms of Lagrange interpolation coordinates y

$$V_y(y) = V_c(c) = V_c(T_{cy} y) = y^\top \big(\underbrace{T_{cy}^\top P_c T_{cy}}_{=:P_y} \big) y \tag{4.7}$$

with $T_{cy} = T_{yc}^{-1}$. In (4.6), only the first and last block rows of T_{yc}, see (3.62), are required if Q_2 is zero or if the Legendre-tau-adapted treatment of Q_2 from (4.39) below is used. Analogous transformations hold for mixed coordinates from Section 3.6.2, in terms of which

$$P_\chi = T_{y\chi}^\top P_y T_{y\chi} = T_{c\chi}^\top P_c T_{c\chi}, \quad T_{y\chi} = T_{yc} T_{c\chi}. \tag{4.8}$$

In any case,

$$V_y(y) = y^\top P_y y = V_c(c) = c^\top P_c c = V_\chi(\chi) = \chi^\top P_\chi \chi \tag{4.9}$$

gives the desired approximation of $V(\phi)$.

Remark 4.1.1 (The role of mixed coordinates in the Legendre-tau-based result). *For the Legendre tau method, the coordinates c, respectively $y = T_{yc} c$ or, even more appropriate, $\chi = [(\tilde c^0)^\top, \ldots, (\tilde c^{(N-1)})^\top, \hat x]^\top$, are, strictly speaking, created from the argument ϕ in $V(\phi)$ via a projection that relies on taking integrals to obtain $(\tilde c^k)_{k\in\{0,\ldots,N-1\}}$ and on $\phi(0) = \hat x$. See (3.47). Therefore, $\chi^\top P_\chi \chi$ from*

(4.9), when written out in terms of the original function ϕ and the submatrices of $P_\chi = (P_{\chi,jk})_{j,k\in\{0,\dots,N\}}$, actually stands for an integral formula

$$\begin{aligned}\chi^\top P_\chi \chi &= \int_{-h}^{0}\int_{-h}^{0} \phi^\top(\xi) P^L_{\mathrm{zz}}(\xi,\theta)\phi(\theta)\,\mathrm{d}\theta\,\mathrm{d}\xi + 2\int_{-h}^{0} \phi^\top(0) P^L_{\mathrm{xz}}(\theta)\phi(\theta)\,\mathrm{d}\theta \\ &\quad + \phi^\top(0) P^L_{\mathrm{xx}}\phi(0), \\ P^L_{\mathrm{zz}}(\xi,\theta) &= \sum_{j=0}^{N-1}\sum_{k=0}^{N-1} P_{\chi,jk}\frac{2j+1}{h}\frac{2k+1}{h} p_j(\vartheta(\xi))p_k(\vartheta(\theta)), \\ P^L_{\mathrm{xz}}(\theta) &= \sum_{k=0}^{N-1} P_{\chi,Nk}\frac{2k+1}{h} p_k(\vartheta(\theta)), \qquad\qquad P^L_{\mathrm{xx}} = P_{\chi,NN}.\end{aligned}$$

However, a term of the form $\int_{-h}^{0} \phi^\top(\theta)\, P^L_{\mathrm{zz,diag}}(\theta)\, \phi(\theta)\,\mathrm{d}\theta$ as encountered in the LK functional formula (1.15) is missing. Therefore, convergence of the involved kernel function $P^L_{\mathrm{zz}}(\xi,\theta)$, as N increases, cannot be expected—which, however, is not necessary for the convergence of the overall result $\chi^\top P_\chi \chi$ anyway. A splitting approach in Section 4.6.1 will resolve that issue and establish a separation between the double integral term and the missing simple integral term, yielding the same $\chi^\top P_\chi \chi$. Still, there is no need to compute such explicit kernel functions since the finite-dimensional quadratic form $\chi^\top P_\chi \chi = c^\top P_c c = y^\top P_y y$ in terms of polynomial coordinates is much easier to handle than an integral formula.

4.1.2 Existence, Uniqueness, and Nonnegativity*

Note that Q_y in (4.4) is a positive semidefinite, but not necessarily positive definite matrix. Let us revisit some properties of the Lyapunov equation (4.3) in this rather uncommon semidefinite case, without further assumptions on the involved matrices. See [89, p. 284], and [39, Thm. 1] for Lemma 4.1.2c.

* The author has prepublished Section 4.1.2 in [S2], ©2024 IEEE.

Lemma 4.1.2. *For any matrix M, denote by $i_-(M)$ (respectively $i_0(M)$, $i_+(M)$) the number of eigenvalues with negative real part (zero real part, positive real part). Consider the Lyapunov equation $PA + A^\top P = -Q$ with given matrices $A, Q \in \mathbb{R}^{\nu\times\nu}$.*

(a) If $\sigma(A)\cap(-\sigma(A)) = \emptyset$, then a unique solution P of the Lyapunov equation exists.

(b) If $Q = Q^\top$ and P is a solution, then $P^\top$ is also a solution. If, additionally, (a) holds, then $P = P^\top$.

(c) If $Q \succeq 0_{\nu\times\nu}$, $P = P^\top$, and $i_0(A) = 0$, then $i_+(P) \leq i_-(A)$ and $i_-(P) \leq i_+(A)$.

Remark 4.1.3. *Existence of the LK functional V in (1.14) is analogously ensured by the time-delay counterpart of Lemma 4.1.2a, the so-called Lyapunov condition, see Proposition B.1.1 in the appendix.*

Proposition 4.1.4. *Let $Q_y \succeq 0_{n(N+1)\times n(N+1)}$ be given. If A_y is Hurwitz, then there exists a unique solution P_y in (4.3). Moreover, $P_y = P_y^\top$ is positive semidefinite.*

Proof. Lemma 4.1.2a with $\sigma(A) \subset \mathbb{C}^-$, Lemma 4.1.2b, and Lemma 4.1.2c with $i_0(A) = i_+(A) = 0$. □

Consequently, if the zero equilibrium of the ODE approximation (3.3) is asymptotically stable, and $D^+_{(\dot y = A_y y)} V_y(y)$ is chosen according to (4.4) and thus nonpositive, then existence, uniqueness, and nonnegativity of $V_y(y)$ in Section 4.1.1 are ensured.

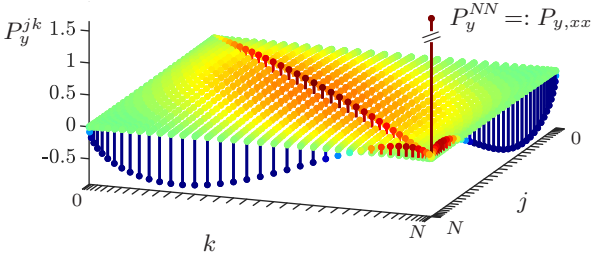


Figure 4.1*: Entries of the matrix P_y in Example 4.1.5 ($N = 40$).

4.1.3 Structure of the Result*

To get an impression of how the Lyapunov equation solution P_y from (4.3) looks like, an example with $n = 1$ is considered. As will be demonstrated, only little implementation effort is required.

Example 4.1.5. *Let $\dot{x}(t) = -0.5x(t) - x(t-2.2)$ and $Q_0 = Q_1 = 1$, $Q_2 = 0$ in (4.4). The solution P_y of (4.3) can be obtained via*

```
Q=blkdiag(Q1,zeros(n*(N-1)),Q0);    P=lyap(A',Q);
```

in Matlab[1], provided A_y is assigned to `A` *(see Remark 3.4.1 or Remark 3.6.1). The structure of P_y for $N = 40$ is depicted in Figure 4.1. It stems from the Legendre tau method, i.e., $A_y = A_y^L$ from (3.63) is used in the Lyapunov equation (4.3) (or, equivalently, $A_c = A_c^L$ from (3.27) in (4.6)). However, Chebyshev collocation with $A_y = A_y^C$ from (3.16) gives almost the same picture of P_y.*

In Figure 4.1, the combs on the last column, the last row, and the diagonal as well as the striking right lower element of the matrix P_y are also existent with a refined grid. Thus, $V_y(y) = y^\top P_y\, y = \sum_{j=0}^{N}\sum_{k=0}^{N}(y^j)^\top P_y^{jk} y^k$ is not the discretized

* The author has prepublished Section 4.1.3 in [S2], ©2024 IEEE.

1 If $Q_2 \neq 0_{n\times n}$ and $A_y = A_y^L$, then `Tcy'*Qc2*Tcy` from (4.39) is added to `Q`, with `Qc2=kron(delay*diag([1./(2*(0:N-1)+1),1]),Q2)`.

* The author has prepublished Figure 4.1 in [S2], ©2024 IEEE.

version of a Lebesque integral $\int_{-h}^{0}\int_{-h}^{0}\phi^\top(\xi)\,P(\xi,\theta)\,\phi(\theta)\,\mathrm{d}\theta\mathrm{d}\xi$. Instead, the combs suggest that

$$V_y(y) = y^\top P_y y = \begin{bmatrix} | \\ z \\ | \\ \hat{x} \end{bmatrix}^\top \begin{bmatrix} & & | \\ & P_{y,zz} & P_{y,xz}^\top \\ & & | \\ - & P_{y,xz} - & P_{y,xx} \end{bmatrix} \begin{bmatrix} | \\ z \\ | \\ \hat{x} \end{bmatrix} \tag{4.10}$$
$$= \hat{x}^\top P_{y,xx}\hat{x} + 2\sum_{k=0}^{N-1}\hat{x}^\top P_{y,xz}^k z^k + \sum_{j=0}^{N-1}\sum_{\substack{k=0\\k\neq j}}^{N-1}(z^j)^\top P_{y,zz}^{jk} z^k + \sum_{k=0}^{N-1}(z^k)^\top P_{y,zz}^{kk} z^k$$

describes, through the (discrete $\hookrightarrow$ continuous) correspondences indicated by (3.23) and by k vs. θ in Figure 3.1

$$\begin{aligned} z^k = \phi(\tilde{\theta}_k),\ k\in\{0,\dots,N-1\} &\quad\hookrightarrow\quad \phi(\theta),\ \theta\in[-h,0),\\ z^j = \phi(\tilde{\theta}_j),\ j\in\{0,\dots,N-1\} &\quad\hookrightarrow\quad \phi(\xi),\ \xi\in[-h,0),\\ \hat{x} = \phi(0) &\quad\hookrightarrow\quad \phi(0), \end{aligned}$$

the discrete version of some

$$\begin{aligned} V(\phi) = \phi^\top(0)P_{\mathrm{xx}}\phi(0) &+ 2\int_{-h}^{0}\phi^\top(0)P_{\mathrm{xz}}(\theta)\phi(\theta)\,\mathrm{d}\theta\\ &+ \int_{-h}^{0}\int_{-h}^{0}\phi^\top(\xi)P_{\mathrm{zz}}(\xi,\theta)\phi(\theta)\,\mathrm{d}\theta\,\mathrm{d}\xi\\ &+ \int_{-h}^{0}\phi^\top(\theta)\,P_{\mathrm{zz,diag}}(\theta)\,\phi(\theta)\,\mathrm{d}\theta. \end{aligned} \tag{4.11a}$$

Note that the latter exactly reflects the known structure of complete-type and related LK functionals given in (1.15).

4.2 Numerical Integration for Validation Purposes

In fact, complete-type and related LK functionals have the structure (4.11a), and the kernel functions can be identified in (1.15) as

$$\begin{aligned} P_{\mathrm{zz}}(\xi,\theta) &= A_1^\top \Psi(\xi-\theta;\tilde{Q})A_1, & P_{\mathrm{xz}}(\theta) &= \Psi(-h-\theta;\tilde{Q})A_1, \\ P_{\mathrm{zz,diag}}(\theta) &= Q_1 + (h+\theta)Q_2, & P_{\mathrm{xx}} &= \Psi(0;\tilde{Q}). \end{aligned} \tag{4.11b}$$

For the sake of validation (e.g., of the quadratic lower bound from the next section), the following considerations go the other way around and derive a discretization of the known formula of $V(\phi)$ by interpolatory quadrature rules. See Table 3.1. That is, the integrals in (4.11a) are replaced by weighted sums from evaluations at the grid points. These sums can again be rewritten as a quadratic form

$$V(\phi) \approx y^\top P_y^{\mathrm{quad}} y \tag{4.12}$$

like (4.10). Note that the involved kernel functions (4.11b) rely on the delay Lyapunov function Ψ, which is described in Appendix B.1.1. Two alternative quadrature rules are applied in the following: Clenshaw–Curtis and Gauss quadrature.

4.2.1 Clenshaw–Curtis Quadrature*

The Clenshaw–Curtis quadrature rule amounts to replacing the integral $\int_{-h}^{0} u(\theta)\,\mathrm{d}\theta$ of a given function $u \in C([-h,0],\mathbb{R})$ by a weighted sum $\sum_{k=0}^{N} u(\tilde{\theta}_k)w_k$ from values at the (Gauss–Lobatto) Chebyshev nodes $\{\tilde{\theta}_k\}_{k\in\{0,\dots,N\}}$. In fact, an interpolatory quadrature is nothing more than computing the integral of an interpolating polynomial instead of the original function u. In the case of Clenshaw–Curtis quadrature, the considered nodes for the interpolation are exactly the ones already

* The author has prepublished Section 4.2.1 in [S2], ©2024 IEEE.

encountered in (3.2), see Table 3.1 and Section 3.4.2. Thus, the approximation of the integral of some $u \in L_1([-h,0],\mathbb{R}^n)$ becomes

$$\int_{-h}^{0} u(\theta)\,\mathrm{d}\theta \approx \int_{-h}^{0} \sum_{k=0}^{N} u(\tilde{\theta}_k)\ell_k(\vartheta(\theta))\,\mathrm{d}\theta = \sum_{k=0}^{N} u(\tilde{\theta}_k)w_k, \tag{4.13}$$

where the required weights

$$w_k = \int_{-h}^{0} \ell_k(\vartheta(\theta))\,\mathrm{d}\tilde{\theta} \overset{(3.14)}{=} \frac{h}{2}\int_{-1}^{1} \ell_k(\tilde{\vartheta})\,\mathrm{d}\tilde{\vartheta} \tag{4.14}$$

are the integrals of the Lagrange basis polynomials, cf. Figure 3.3. The weights w_k defined by (4.14) are explicitly known, see [68, Sec. 3.7]. Ready to use implementations of the latter are, e.g., available[2] from the Chebfun toolbox [59].

Applying (4.13) to the integral terms in (4.11a) gives

$$\begin{aligned} V(\phi) \approx{}& \phi^\top(0)P_{\mathrm{xx}}\,\phi(0) + 2\sum_{k=0}^{N} w_k\phi^\top(0)P_{\mathrm{xz}}(\tilde{\theta}_k)\phi(\tilde{\theta}_k) \\ &+ \sum_{j=0}^{N} w_j \sum_{k=0}^{N} w_k\phi^\top(\tilde{\theta}_j)P_{\mathrm{zz}}(\tilde{\theta}_j,\tilde{\theta}_k)\phi(\tilde{\theta}_k) \\ &+ \sum_{k=0}^{N} w_k\phi^\top(\tilde{\theta}_k)P_{\mathrm{zz,diag}}(\tilde{\theta}_k)\phi(\tilde{\theta}_k). \end{aligned} \tag{4.15}$$

[2] implemented via `[theta,w]=chebpts(N+1,[-delay,0])`

As in (4.10), the result (4.15) can be written as a quadratic form. Let $y^k = \phi(\tilde{\theta}_k)$, $k \in \{0\dots,N\}$, where $y^N = \phi(\tilde{\theta}_N) = \phi(0)$. Then (with $p = \dim(z) := \dim([{y^0}^\top,\dots,{y^{N-1}}^\top]^\top) = nN$)

$$\begin{aligned} V(\phi) \;\approx\; y^\top P_y^{\mathrm{quad}} y \;=\; y^\top \Bigg(& \begin{bmatrix} 0_{p\times p} & 0_{p\times n} \\ 0_{n\times p} & P_{\mathrm{xx}} \end{bmatrix} \\ & + \begin{bmatrix} & 0_{p\times(p+n)} & \\ \text{—} & \left(P_{\mathrm{xz}}(\tilde{\theta}_k)w_k\right)_k & \text{—} \end{bmatrix} + \begin{bmatrix} & | & \\ 0_{(p+n)\times p} & & \left(w_j P_{\mathrm{xz}}^\top(\tilde{\theta}_j)\right)_j \\ & | & \end{bmatrix} \\ & + \Big(w_j P_{\mathrm{zz}}(\tilde{\theta}_j,\tilde{\theta}_k)w_k\Big)_{jk} + \mathrm{blkdiag}\Big(\left(w_k P_{\mathrm{zz,diag}}(\tilde{\theta}_k)\right)_k\Big) \Bigg)\, y. \end{aligned} \tag{4.16}$$

See Appendix B.1.2 for a factorization taking (4.11b) into account. Note that the right lower component of P_y^{quad} approximately becomes P_{xx} since the other contributions are weighted by w_N, which is quite small in the non-equidistant grid.

Since Clenshaw–Curtis quadrature relies on Chebyshev nodes (cf. Table 3.1), the involved vector y coincides with the one used above in the ODE-based approach. Indeed, the picture of the resulting P_y^{quad} turns out to be hardly distinguishable from Figure 4.1 when applied to Example 4.1.5. See Section 4.4 for further numerical comparisons.

4.2.2 Gauss Quadrature*

As an alternative, (Legendre) Gauss quadrature can be applied. In that case, the integral of a function is approximated by weighted sums from the function values at (Gauss) Legendre nodes. Being Gauss nodes, cf. Table 3.1, they do not contain the boundary points of the domain $[-h,0]$. That is why, in the following, a grid

* The author has prepublished Section 4.2.2 in [S2], ©2024 IEEE.

of only N (Gauss) Legendre nodes[3] $\tilde{\theta}_k^L$, is considered, to which the zero end point with zero weight is added. Thus, the $N+1$ nodes $[(\tilde{\theta}^L)^\top, 0]^\top$ and weights $[(w^L)^\top, 0]^\top$ are obtained. Therefore, contrary to the Gauss–Lobatto-node-based Clenshaw–Curtis quadrature, the contributions in (4.16) do not overlap, yielding for $y_G = [\phi^\top(\tilde{\theta}_0^L), \ldots, \phi^\top(\tilde{\theta}_{N-1}^L), \phi^\top(0)]^\top$

$$V(\phi) \approx y_G^\top \begin{bmatrix} \Big(w_j^L P_{\mathrm{zz}}(\tilde{\theta}_j^L, \tilde{\theta}_k^L) w_k^L\Big)_{jk} + D & (w_j^L P_{\mathrm{xz}}^\top(\tilde{\theta}_j^L))_j \\ — \ (P_{\mathrm{xz}}(\tilde{\theta}_k^L) w_k^L)_k \ — & P_{\mathrm{xx}} \end{bmatrix} y_G$$
$$\text{with } D = \mathrm{blkdiag}\Big((w_k^L P_{\mathrm{zz,diag}}(\tilde{\theta}_k^L))_k\Big). \tag{4.17}$$

To sum up, a finite-dimensional quadratic form that approximates $V(\phi)$ can also be obtained from a numerical integration of the known formula of complete-type and related LK functionals, e.g., relying on Clenshaw–Curtis or Gauss quadrature, which are proposed in the present section. However, in this numerical-integration-based approach, knowledge of the delay Lyapunov matrix function Ψ is needed, which is not the case in the ODE-based approach from Section 4.1. See Appendix B.1.3 for further remarks.

4.3 The Quadratic Lower Bound*

As a result of the preceding sections, we have an approximation of the LK functional. However, in applications, the quadratic lower bound (1.16) is needed

[3] via `[thetaL,wL]=legpts(N,[-delay,0])` using the toolbox [59]

* The author has prepublished Section 4.3 in [S2], ©2024 IEEE.

as well. If $Q_0, Q_1 \succ 0_{n\times n}, Q_2 \succeq 0_{n\times n}$, existence of a nonzero[4] coefficient $k_1 > 0$ is proven in [110, Lem. 2.10], given the RFDE equilibrium is exponentially stable. In a discrete version for the approximation V_y, the bound (1.16) becomes $\forall y = [z^\top, \hat{x}^\top]^\top, z \in \mathbb{R}^{nN}, \hat{x} \in \mathbb{R}^n$:

$$k_1 \|\hat{x}\|_2^2 \leq V_y(y). \tag{4.18}$$

Since solely $\hat{x} = y^N$ is considered, (4.18) does not refer to the common $\lambda_{\min}(P_y)\|y\|_2^2 \leq V_y(y)$ encountered in the template ii from Section 1.2. Why this discrete version of (1.16) still also makes sense in a Lyapunov analysis of the approximating ODE, will be explained in Section 4.5.

The main contribution of the present section, Lemma 4.3.1, immediately leads to the searched bound (4.18) in Theorem 4.3.2. For the sake of readability, the lemma is expressed in terms of a general positive semidefinite matrix P with a left upper submatrix Z, instead of P_y and $P_{y,zz}$ introduced in (4.10). The result is based on the generalized Schur complement (4.20), cf. [91], where Z^- is a generalized matrix inverse of Z, e.g., the Moore-Penrose inverse. If Z is nonsingular, then $Z^- = Z^{-1}$.

Lemma 4.3.1. *Let* $P = \left[\begin{smallmatrix} Z & B \\ B^\top & X \end{smallmatrix}\right]$ *with* $Z = Z^\top \in \mathbb{R}^{p\times p}, B \in \mathbb{R}^{p\times n}, X = X^\top \in \mathbb{R}^{n\times n}$. *If* P *is positive semidefinite, then*

$$\min_{\substack{z\in\mathbb{R}^p \\ x\in\mathbb{R}^n\setminus\{0_n\}}} \frac{1}{\|x\|_2^2} \begin{bmatrix} z \\ x \end{bmatrix}^\top \begin{bmatrix} Z & B \\ B^\top & X \end{bmatrix} \begin{bmatrix} z \\ x \end{bmatrix} = \lambda_{\min}(P/Z), \tag{4.19}$$

$$\textit{where} \quad P/Z = X - B^\top Z^- B. \tag{4.20}$$

[4] In contrast to quadratic forms from finite-dimensional matrices, in infinite dimensions coercivity of the associated bilinear form (existence of a quadratic lower bound) is a stronger concept than positive definiteness (positivity for any nonzero element). The same holds for the partial concepts. Consequently, despite of V_y being partially positive definite w.r.t. $\hat{x}$ (Definition 4.5.8), the largest possible coefficient in (4.18) as $N \to \infty$ could become $k_1 \to 0$, see Remark 4.4.3.

The minimum is attained by $\left[\begin{smallmatrix} z \\ x \end{smallmatrix}\right] = \left[\begin{smallmatrix} -Z^- B v \\ v \end{smallmatrix}\right]$*, with* v *being an arbitrary eigenvector in* $(P/Z)\, v = v\, \lambda_{\min}(P/Z)$ *and* Z^- *is an arbitrary generalized inverse of* Z*.*

Proof. Let us replace z by $w := z + Z^- B\, x$, which amounts to the coordinate transformation

$$\begin{bmatrix} z \\ x \end{bmatrix} = \begin{bmatrix} I_p & -Z^- B \\ 0_{n\times p} & I_n \end{bmatrix} \begin{bmatrix} w \\ x \end{bmatrix} =: T_{yq} \begin{bmatrix} w \\ x \end{bmatrix}. \tag{4.21}$$

As a result, the so-called generalized Aitken block-diagonalization of P is obtained in

$$\begin{aligned} \begin{bmatrix} z \\ x \end{bmatrix}^\top \begin{bmatrix} Z & B \\ B^\top & X \end{bmatrix} \begin{bmatrix} z \\ x \end{bmatrix} &= \begin{bmatrix} w \\ x \end{bmatrix}^\top T_{yq}^\top \begin{bmatrix} Z & B \\ B^\top & X \end{bmatrix} T_{yq} \begin{bmatrix} w \\ x \end{bmatrix} \\ &= \begin{bmatrix} w \\ x \end{bmatrix}^\top \begin{bmatrix} Z & (I_p - ZZ^-)B \\ B^\top(I_p - Z^-Z) & X - B^\top Z^- B \end{bmatrix} \begin{bmatrix} w \\ x \end{bmatrix} \\ &= w^\top Z w + x^\top (P/Z)\, x, \end{aligned} \tag{4.22}$$

with the last step being based on $(I_p - ZZ^-)B = 0_{p\times n}$, which holds if P is positive semidefinite [91, Thm. 1.19]. The submatrix Z of P is also positive semidefinite due to Cauchy's interlacing theorem, and thus (4.22) is lower bounded by

$$w^\top Z\, w + x^\top (P/Z)\, x \;\geq\; x^\top (P/Z)\, x \;\geq\; \lambda_{\min}(P/Z) \|x\|_2^2.$$

The bound is attained for $w = 0_p$ and $x = v$. □

The following theorem is not only useful for the ODE-based approach from Section 4.1.1. It is as well applicable to the numerical-integration-based results from Section 4.2.

Theorem 4.3.2. *If P_y in $V_y(y) = y^\top P_y y$ is positive semidefinite, then the largest possible coefficient in (4.18) is*

$$k_1 = \lambda_{\min}(P_y / P_{y,zz}), \tag{4.23}$$

where $P_{y,zz}$ denotes the left upper $nN \times nN$ submatrix of P_y and $(\cdot/\cdot)$ is the generalized Schur complement (4.20).

Proof. Lemma 4.3.1 applied to $P = P_y$ with $Z = P_{y,zz}$ as in (4.10). □

Testing whether P_y is positive semidefinite is not even required if V_y originates from the ODE-based approach in Section 4.1.1. If A_y is Hurwitz, only (4.23) must be evaluated.

Corollary 4.3.3. *Let $V_y(y) = y^\top P_y y$, where P_y is a solution of (4.3) for a given positive semidefinite matrix Q_y. If A_y is Hurwitz, then (4.18) holds with k_1 from (4.23).*

Proof. By Proposition 4.1.4, P_y is positive semidefinite. Consequently, Theorem 4.3.2 applies. □

Finally, it should be noted that Lemma 4.3.1 can also be applied to P_χ from (4.8), thus obtaining the quadratic lower bound as in Corollary 4.3.3, but without the need of the dense transformation matrix T_{cy}.

4.4 Example and Comparison*

In the following example, the thus obtained bound will be compared with known quadratic lower bounds (1.16) on the LK functional (4.11). These known formulas for the coefficient in $k_1\|x(t)\|_2^2 \leq V(x_t)$ are

$$\begin{aligned} k_1 &= \max \alpha && \text{[110, Lem. 2.10]} \\ &\quad \text{s.t.} \quad \begin{bmatrix} Q_0 & 0_{n\times n} \\ 0_{n\times n} & Q_1 \end{bmatrix} + \alpha \begin{bmatrix} A_0^\top + A_0 & A_1 \\ A_1^\top & 0_{n\times n} \end{bmatrix} \succeq 0_{2n\times 2n}, \\ k_1 &= \min \left\{ \frac{\lambda_{\min}(Q_0)}{2\|A_0\|_2 + \|A_1\|_2}, \frac{\lambda_{\min}(Q_1)}{\|A_1\|_2} \right\}, && \text{[137, Prop. 1]} \end{aligned}$$

provided the equilibrium is exponentially[5] stable and $Q_0, Q_1 \succ 0_{n\times n}, Q_2 \succeq 0_{n\times n}$. Two issues should be noted.

First, the LK functional satisfies the monotonicity condition of the common LK theorem by construction, cf. (4.29). Thus, the existence of a quadratic lower bound with $k_1 > 0$ (or actually even $k_1 \geq 0$, cf. Theorem 4.5.13) is also the crucial missing step that proves asymptotic stability via the LK functional. However, the above formulas are only valid if exponential (equivalently, asymptotic) stability has been proven beforehand. Hence, the stability analysis must already be done by other means in a separate step. For instance, this can be achieved via the eigenvalues of A_y. Having thus A_y already at hand, the proposed approach becomes even more convenient.

Remark 4.4.1. *As a consequence of the above issue, how at all to conclude stability from the LK functional (4.11) or the involved delay Lyapunov matrix function Ψ has long been an open question. It has only recently been resolved*

* The author has prepublished Section 4.4 in [S2], ©2024 IEEE.

5 equivalently, asymptotically since (1.4) is a linear autonomous RFDE. In linear RFDEs with bounded delays, uniform asymptotic stability and uniform exponential stability are equivalent [82, Thm. 5.3 in Ch. 6]. Moreover, in autonomous or periodic RFDEs (in contrast to neutral FDEs), asymptotic stability is always uniform [82, Lemma 1.1 in Ch. 5].

by Egorov et al. [60] and Gomez et al. [72]. Their criterion is equivalent to requiring that, for some $\tilde{Q} \succ 0_{n\times n}$*,*

$$\tilde{P}_{\mathrm{zz}}(\xi,\theta) := \Psi(\xi-\theta;\tilde{Q}) \tag{4.24}$$

is a positive definite kernel, in the sense that the block matrix $(\tilde{P}_{\mathrm{zz}}(\theta_j,\theta_k))_{jk}$ *must be positive semidefinite, with an a priori bound on the discretization resolution of the grid* $(\xi,\theta) \in \{(\theta_j,\theta_k)\}_{jk} \subset [-h,0]\times[-h,0]$*. Note that (4.24) is related to* $P_{\mathrm{zz}}(\xi,\theta)$ *from (4.11b), and* $(P_{\mathrm{zz}}(\theta_j,\theta_k))_{jk}$ *is a central part of* P_y^{quad} *in (4.16). In the end, despite of a completely different framework, the result can be brought in relation to the numerical integration from Section 4.2, see Appendix B.1.2.*

Second, of course the LK functional changes as the delay changes. Note that, however, the above stated formulas for k_1 do not depend on the value of the delay.

Example 4.4.2. *For all delay values* h *that are smaller than the critical delay* $h_c := \arccos(-0.9)/\sqrt{1-0.9^2} \approx 6.17$*, the equilibrium of*

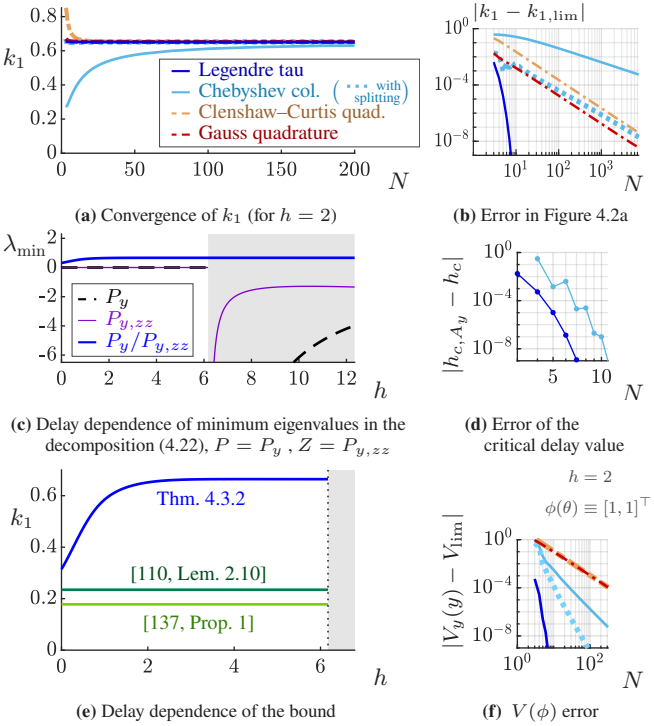
$$\dot{x}(t) = \begin{bmatrix} -2 & 0 \\ 0 & -0.9 \end{bmatrix} x(t) + \begin{bmatrix} -1 & 0 \\ -1 & -1 \end{bmatrix} x(t-h) \tag{4.25}$$

is asymptotically stable (see Example 2.2.5). Let $Q_0 = Q_1 = I_2, Q_2 = 0_{2\times 2}$*. For any given* $h > 0$ *(affecting* A_y*), the Lyapunov equation for* P_y *can be implemented as in Example 4.1.5. The resulting* k_1 *from (4.23) is obtained via the additional lines*

```
p=mat2cell(P,n*[N,1],n*[N,1]);
k1=min(eig(p{2,2}-p{2,1}*(p{1,1}\p{1,2})))
```

in Matlab (as $P_{y,zz}$ *is nonsingular). For the sake of comparison, the numerical-integration-based matrices* P_y^{quad} *from the Clenshaw–Curtis quadrature in (4.16)*

* The author has prepublished Figure 4.2 in [S2], ©2024 IEEE.



(a) Convergence of k_1 (for $h = 2$)

(b) Error in Figure 4.2a

(c) Delay dependence of minimum eigenvalues in the decomposition (4.22), $P = P_y$, $Z = P_{y,zz}$

(d) Error of the critical delay value

(e) Delay dependence of the bound

(f) $V(\phi)$ error

Figure 4.2*: Example 4.4.2. In particular, Figure 4.2e shows the improved quadratic lower bound. (Figures a,b,d,f share the same legend).

and Gauss quadrature in (4.17) are considered as well. Figure 4.2a shows the convergence of the lower-bound coefficient k_1 for all considered approaches.

Figure 4.2 also gives some further insights. Figure 4.2b certificates a surprisingly fast convergence for the Legendre tau method. This is also confirmed by other examples (if $Q_2 \neq 0_{n\times n}$, the Lyapunov equation right-hand side from Section 4.6.1 below should be used). For the Chebyshev collocation method, a splitting approach is introduced in Section 4.6.1 which gives an improved rate of convergence, shown by a dashed line in Figure 4.2b.

Figure 4.2c (Legendre tau, $N = 1000$) shows the interplay of the matrices in (4.22), once the asymptotic stability is lost for delays larger than $h \approx 6.17$. Theorem 4.5.13 will prove that positive semidefiniteness of P_y is indeed necessary and sufficient for A_y being Hurwitz.

Let us consider the boundary h_{c,A_y} between the white and gray delay region in Figure 4.2c. It marks the smallest delay at which the matrix A_y is no longer Hurwitz (equivalently, where no longer a positive semidefinite solution P_y exists), which can, e.g., be fine estimated by a bisection method. Already with a rough discretization resolution N, this boundary reflects the analytically known critical delay h_c of (4.25) quite precisely, and its rapid convergence is shown in Figure 4.2d for both the Legendre tau and the Chebyshev collocation method.

Most importantly, Figure 4.2e reveals that the largest possible quadratic lower bound depends on the value of the delay. Theorem 4.3.2 clearly gives a less conservative value of k_1 than the known formulas (green lines). For non-small delays, the bound is even improved by a multiple.

Moreover, Figure 4.2f shows the rapid convergence of the numerical result for $V(\phi)$ with an exemplary argument ϕ.

A remark on non-complete-type functionals is in order.

Remark 4.4.3. *If $Q_1 = Q_2 = 0_{n\times n}$, only a local cubic lower bound on V is known to exist, and non-existence of a positive quadratic one (see the comments on (4.18)) is proven for [110, Example 2.1]. Indeed, for this example, k_1 from (4.23) converges to zero as N increases.*

Finally, the reduced conservatism of k_1, already indicated by Figure 4.2e, is confirmed by other examples in Table 4.1.

* The author has prepublished Table 4.1 in [S2], ©2024 IEEE.

	k_1	[137, Prop. 1]	[110, Lem. 2.10]	Thm. 4.3.2
[137], Example 5.1		0.7500	0.8229	1.4596
[137], Example 5.2		0.6000	2.3238	3.8660
[137], Example 5.4		0.1464	0.1978	0.5229

Table 4.1*: Improvements of the quadratic lower bound for three physical examples from the literature.

4.5 Interpretation in Terms of Partial Stability

Section 4.1.1 provides a function V_y that is derived similarly to the standard construction of Lyapunov functions discussed in the ODE template ii from Section 1.2 in the introduction of this thesis. However, there is a major difference: the matrix Q_y in the derivative (4.4) is not chosen positive definite if $Q_2 = 0_{n\times n}$, which is a choice of interest in this thesis (moreover, note that, even if $Q_2 \succ 0_{n\times n}$, the involved Q_y is theoretically positive definite for any finite N, but the smallest eigenvalue of Q_y converges to zero as N increases since the integration weights diminish). As a result, V_y obtained in Section 4.1.1 does not necessarily qualify as a Lyapunov function for the ODE (3.3). Remarkably, the lower bound (4.18) of interest on V_y also does not fit with the classical Lyapunov theory for ODEs.

The objective of the present section is to explain why V_y is still meaningful for a stability analysis of the approximating ODE. In particular, within the presented approach, the lower bound (4.18) is exactly what is required. First, the section clarifies what we are actually looking for when we target stability in an RFDE.

4.5.1 Stability in RFDEs*

Having in mind the classical Lyapunov theorem for ODEs, one might wonder why the lower bound in (1.16) relies on $\|x(t)\|$ and not on the norm of the RFDE state x_t. The latter addresses the norm in $C([-h,0],\mathbb{R}^n)$ defined by

$$\|x_t\|_C = \max_{\theta\in[-h,0]} \|x_t(\theta)\|. \tag{4.26}$$

For Lyapunov functions in ODEs, both the positive-definiteness bound ($\kappa_1(\|x\|) \leq V(x)$, $\kappa_1 \in \mathcal{K}$) and the monotonicity requirement ($D^+_{(\dot{x}=f(x))}V(x) \leq -\kappa_3(\|x\|)$, $\kappa_3 \in \mathcal{K}$) refer to the norm of the ODE state. Thus, one would expect (4.26) at these places when transferring Lyapunov's results from $x(t) \in \mathbb{R}^n$ to $x_t = \phi \in C([-h,0],\mathbb{R}^n)$. However, this is not the case in the following common LK theorem—neither in the left inequality of (4.28) nor in (4.29). Instead of $\|\phi\|_C = \|x_t\|_C$, only $\|\phi(0)\| = \|x_t(0)\| = \|x(t)\|$ occurs. As usual, the theorem refers to general autonomous RFDEs

$$\dot{x}(t) = f(x_t), \tag{4.27}$$

with $f(0_{n_{[-h,0]}}) = 0_n$ and f locally Lipschitz.

Theorem 4.5.1 (Lyapunov–Krasovskii theorem [82, Thm. 5-2.1]). *If there is a continuous $V\colon C([-h,0],\mathbb{R}^n) \to \mathbb{R}_{\geq 0}$ such that, for all ϕ in a domain $G \subseteq C([-h,0],\mathbb{R}^n)$, $0_{n_{[-h,0]}} \in G$, it holds*

$$\kappa_1(\|\phi(0)\|) \leq V(\phi) \leq \kappa_2(\|\phi\|_C) \tag{4.28}$$

$$D^+_{(4.27)}V(\phi) \leq -\kappa_3(\|\phi(0)\|), \tag{4.29}$$

with some class-K functions $\kappa_{1,2,3} \in \mathcal{K}$, then the zero equilibrium of (4.27) is asymptotically stable.

* The author has prepublished Section 4.5.1 in [S2], ©2024 IEEE.

The key to the above question is that there are two, obviously equivalent, definitions of asymptotic stability in the RFDE. Starting from the same norm ball for the initial function x_0, they differ in the condition on the implication side: Either the state x_t with the norm (4.26) is taken into account (Definition 4.5.2a), or the pointwise solution $x(t) \in \mathbb{R}^n$ is considered (Definition 4.5.2b).

Definition 4.5.2 (Lyapunov stability in RFDEs). *The zero equilibrium of (4.27) is asymptotically stable if*

a)
$$\begin{aligned} &\forall \varepsilon > 0, \exists \delta > 0: \ \|x_0\|_C < \delta \implies \forall t \geq 0 : \|x_t\|_C < \varepsilon \\ &\textit{and} \quad \exists r > 0: \quad \|x_0\|_C < r \implies \|x_t\|_C \to 0 \textit{ as } t \to \infty \end{aligned}$$

or, equivalently,

b)
$$\begin{aligned} &\forall \varepsilon > 0, \exists \delta > 0: \ \|x_0\|_C < \delta \implies \forall t \geq 0 : \|x(t)\| < \varepsilon \\ &\textit{and} \quad \exists r > 0: \quad \|x_0\|_C < r \implies \|x(t)\| \to 0 \textit{ as } t \to \infty. \end{aligned}$$

In terms of the whole state x_t, the pointwise consideration in Definition 4.5.2b refers only to the boundary value $x(t) = x_t(0)$ in Figure 3.1b. The classical LK theorem, Theorem 4.5.1, addresses Definition 4.5.2b since, $\forall t \geq 0$,

$$\kappa_1(\|x(t)\|) \overset{(4.28)}{\leq} V(x_t) \overset{(4.29)}{\leq} V(x_0) \overset{(4.28)}{\leq} \kappa_2(\|x_0\|_C) \tag{4.30}$$

gives a pointwise estimation $\|x(t)\| \leq \kappa_1^{-1}(\kappa_2(\|x_0\|_C))$ to indicate stability. A theorem that addresses Definition 4.5.2a would instead rely on $\kappa_1(\|\phi\|_C)$ in (4.28) and $\kappa_3(\|\phi\|_C)$ in (4.29), as has been expected above. Such a theorem is also valid [116, Thm. 30.1], but these bounds are quite restrictive and not satisfied by common LK functionals.

4.5.2 Partial Stability in ODEs*

In the approximating ODE, see Figure 3.1c, the state $y(t) \in \mathbb{R}^{n(N+1)}$ represents the RFDE state $x_t \in C([-h, 0], \mathbb{R}^n)$, and its last vector-valued component

* The author has prepublished Section 4.5.2 in [S2], ©2024 IEEE.

$y^N(t) = \hat{x}(t) \in \mathbb{R}^n$ represents the pointwise solution value $x(t) \in \mathbb{R}^n$. While Definition 4.5.2a translates to the usual[6] definition of asymptotic stability in the ODE, Definition 4.5.2b amounts to the concept of partial asymptotic stability with respect to (w.r.t.) $\hat{x}$. Again, the definition will be given for a general class of systems. These are ODEs where $y(t)$ is partitioned into two parts, $z(t) \in \mathbb{R}^p$ and $\hat{x}(t) \in \mathbb{R}^n$, with $\dim(y(t)) = p + n$, and the latter part $\hat{x}(t)$ is of special interest. That is, autonomous ODEs

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} z(t) \\ \hat{x}(t) \end{bmatrix} = \begin{bmatrix} f^z(z(t), \hat{x}(t)) \\ f^x(z(t), \hat{x}(t)) \end{bmatrix} \tag{4.31}$$

with $\begin{bmatrix} f^z(0_p, 0_n) \\ f^x(0_p, 0_n) \end{bmatrix} = 0_{p+n}$ and $f^{z,x}$ locally Lipschitz are considered.

Definition 4.5.3 (Lyapunov–Rumyantsev partial stability).
The zero equilibrium of (4.31) is partially stable w.r.t. $\hat{x}$ if

$$\forall \varepsilon > 0, \exists \delta > 0 : \qquad \left\| \begin{bmatrix} z(0) \\ \hat{x}(0) \end{bmatrix} \right\| < \delta \implies \forall t \geq 0 : \|\hat{x}(t)\| < \varepsilon.$$

It is partially asymptotically stable w.r.t. $\hat{x}$ if, additionally,

$$\exists r > 0 : \qquad \left\| \begin{bmatrix} z(0) \\ \hat{x}(0) \end{bmatrix} \right\| < r \implies \|\hat{x}(t)\| \to 0 \text{ as } t \to \infty.$$

For an in-depth discussion of this stability concept, see [184]. As in Definition 4.5.2b for stability in RFDEs, the initial value deviations consider the whole state, but the implications address only the part $\hat{x}$ that is of special interest.

The following partial stability theorem fits well with Theorem 4.5.1 (note that an upper bound $V_y(\left[\begin{smallmatrix} z \\ \hat{x} \end{smallmatrix}\right]) \leq \kappa_2(\| \left[\begin{smallmatrix} z \\ \hat{x} \end{smallmatrix}\right] \|)$ always exists).

[6] The choice of the norm $\|y\|_\infty = \max_{k \in \{0,\dots,N\}} \|y^k\|$ is irrelevant due to the equivalence of norms in finite-dimensional spaces.

Theorem 4.5.4 (Peiffer and Rouche 1969 [152, Thm. II]). *If there is a continuous $V_y\colon \mathbb{R}^{p+n} \to \mathbb{R}_{\geq 0}$, $V_y(0_{p+n}) = 0$, such that, for all $\left[\begin{smallmatrix} z \\ \hat{x} \end{smallmatrix}\right]$ in a domain $G \subseteq \mathbb{R}^{p+n}$, $0_{p+n} \in G$, it holds*

$$\kappa_1(\|\hat{x}\|) \leq V_y(\left[\begin{smallmatrix} z \\ \hat{x} \end{smallmatrix}\right]), \tag{4.32}$$

with $\kappa_1 \in \mathcal{K}$, and $D^+_{(4.31)}V_y(\left[\begin{smallmatrix} z \\ \hat{x} \end{smallmatrix}\right]) \leq 0$, then the zero equilibrium of (4.31) is partially stable w.r.t. $\hat{x}$. If, additionally, $\forall \left[\begin{smallmatrix} z \\ \hat{x} \end{smallmatrix}\right] \in G$:

$$D^+_{(4.31)}V_y(\left[\begin{smallmatrix} z \\ \hat{x} \end{smallmatrix}\right]) \leq -\kappa_3(\|\hat{x}\|) \tag{4.33}$$

with $\kappa_3 \in \mathcal{K}$, and there exists $r > 0$ such that $\left\| \left[\begin{smallmatrix} z(0) \\ \hat{x}(0) \end{smallmatrix}\right] \right\| < r$ implies that $\|f^x(z(t), \hat{x}(t))\|$ is bounded for all $t \geq 0$, then it is partially asymptotically stable w.r.t. $\hat{x}$.

As in the classical LK theorem for RFDEs (Theorem 4.5.1), both the (partial) positive-definiteness condition (4.32) and the monotonicity requirement[7] (4.33) consider only the part of special interest $\hat{x}(t) = y^N(t) \approx x(t) = x_t(0)$. Henceforth, V in Theorem 4.5.4 will be called a *partial Lyapunov function*.

To sum up, the discretization of Definition 4.5.2b for RFDE stability is exactly the definition of Lyapunov–Rumyantsev partial stability w.r.t. $\hat{x}$ (Definition 4.5.3). Moreover, the Lyapunov–Krasovskii theorem for stability in the RFDE (Theorem 4.5.1) becomes Peiffer and Rouche's theorem for partial stability (Theorem 4.5.4).

[7] Criteria that come without the boundedness condition below (4.33) require a full monotonicity condition $D^+_{(4.31)}V_y(y) \leq -\kappa_3(\|y\|), \kappa_3 \in \mathcal{K}$, cf. [119].

4.5.3 Equivalence of Stability and Partial Stability in the Approximating ODE*

In general ODEs, the concept of partial stability is a weaker concept than stability. We can still focus without doubt on partial stability if the equivalence between Definition 4.5.2a and 4.5.2b is reflected by the ODE approximation, so that proving partial stability w.r.t. $\hat{x}$ is already sufficient for proving stability.

Condition 4.5.5. *The zero equilibrium of the ODE approximation (3.3) is (asymptotically) stable if and only if it is partially (asymptotically) stable w.r.t. $\hat{x}$.*

To verify this condition for the discretization schemes at hand, the following result from the realm of total stability is expedient.

Lemma 4.5.6 ([118, Thm. 3.11.3]). *If the zero equilibrium of the auxiliary system*

$$\dot{z} = f^z(z, 0_n) \tag{4.34}$$

is asymptotically stable, then, in (4.31), partial (asymptotic) stability w.r.t. $\hat{x}$ of the zero equilibrium implies (asymptotic) stability of the zero equilibrium.

Loosely speaking, for reasonable approximations the latter seems to be a matter of course since, if $x(t)$ for $t \geq 0$ could be forced to remain zero, then, for $t \geq h$, the solution segment x_t is zero, which should at least asymptotically be reflected by $z(t) \to 0_p$ as $t \to \infty$. In terms of the linear ODE (3.3), Lemma 4.5.6 only refers to the submatrix $A_{y,zz} := ({A_y}^{jk})_{j,k\in\{0,\dots,N-1\}}$. For collocation schemes like $A_y = A_y^C$ in Section 3.4, stability of this submatrix is clearly neither affected by the RFDE coefficient matrices A_0, A_1 (occurrence only in the last block-row), nor the delay h (scalar factor), nor the dimension n (Kronecker product with I_n). For tau methods, an analogous independence can be achieved by first applying a change of basis w.r.t. the z-coordinates. Using the more natural mixed coordinates (3.64)

* The author has prepublished Section 4.5.3 in [S2], ©2024 IEEE.

rather than interpolation coordinates still incorporates $\hat{x}$ as the last component of the coordinate vector but the upper part of the system matrix does not rely on A_0, A_1. Consequently, in these coordinates, setting, e.g., $A_0 = A_1 = 0_{n\times n}$ does not alter the submatrix eigenvalues. The next lemma formulates the thus motivated coordinate invariant form of Lemma 4.5.6 for the linear ODE. Whether it applies is, consequently, no question of A_0, A_1, h, but it is rather a question of the discretization scheme.

Corollary 4.5.7. *Consider the linear ODE (3.3). If there exists a change of coordinates w.r.t. z, where $[z^\top, \hat{x}^\top]^\top = T\,[v^\top, \hat{x}^\top]^\top$, such that the left upper $nN \times nN$ submatrix of $T^{-1}A_yT$ is Hurwitz, then Condition 4.5.5 holds.*

Proof. Lemma 4.5.6 with (4.31) given by $\frac{\mathrm{d}}{\mathrm{d}t}\left[\begin{smallmatrix} v \\ \hat{x}\end{smallmatrix}\right] = T^{-1}A_yT\left[\begin{smallmatrix} v \\ \hat{x}\end{smallmatrix}\right]$. □

For $A_y = A_y^C$ from the Chebyshev collocation method (3.4), the submatrix $A_{y,zz}$ can indeed proven to be Hurwitz for any discretization resolution N [56, Prop. 2], [170]. Thus, by Corollary 4.5.7 (with $v = z$), Condition 4.5.5 holds. For other discretization schemes, see [38, Sec. 4.3.2]. For the Legendre tau system matrix in mixed coordinates[8] (3.64), Lemma 4.5.6 can numerically shown to be true for relevant values of N.

Consequently, Condition 4.5.5 is not only a reasonable assumption for an ODE that approximates an RFDE, but, regarding the discretization of an RFDE, it can even be confirmed as a property of the underlying discretization schemes.

4.5.4 Proving Stability in the ODE via the Result from the Numerical Approach*

The main result of this section, Theorem 4.5.12, will show that V_y from Section 4.1.1 indeed always qualifies as a partial Lyapunov function for (3.3) if the

[8] In terms of Corollary 4.5.7, $v = [(c^0)^\top, \dots, (c^{N-1})^\top]^\top$.

* The author has prepublished Section 4.5.4 in [S2], ©2024 IEEE.

equilibrium is asymptotically stable. As a side effect, Theorem 4.5.13 will give a necessary and sufficient stability criterion in terms of P_y. The following wording is used.

Definition 4.5.8. *Let $\hat{x}$-pd be the abbreviation for "partially positive definite w.r.t. the components $\hat{x}$".*

(a) A function $U\colon \mathbb{R}^{p+n} \to \mathbb{R};\ y = \left[\begin{smallmatrix} z \\ \hat{x} \end{smallmatrix}\right] \mapsto U(y)$ is called $\hat{x}$-pd on $\Omega \subseteq \mathbb{R}^{p+n}$ if it is positive semidefinite, i.e., $\forall y \in \Omega : U(y) \geq 0$, $U(0_{p+n}) = 0$, and $\forall y = \left[\begin{smallmatrix} z \\ \hat{x} \end{smallmatrix}\right] \in \Omega$ with $\|\hat{x}\| \neq 0 : U(y) > 0$.

(b) A symmetric matrix $M = M^\top \in \mathbb{R}^{(p+n)\times(p+n)}$ is called $\hat{x}$-pd if $U(y) = y^\top M y$ is $\hat{x}$-pd on $\mathbb{R}^{p+n}$.

Analogously to local, or in terms of $U(y) = y^\top M y$ even global, positive definiteness, cf. [108, Lemma 4.3], partial positive definiteness can be expressed via a class-K function.

Lemma 4.5.9. *$M = M^\top \in \mathbb{R}^{(p+n)\times(p+n)}$ is $\hat{x}$-pd if and only if $\exists \kappa \in \mathcal{K}$ such that $\forall \left[\begin{smallmatrix} z \\ \hat{x} \end{smallmatrix}\right] \in \mathbb{R}^{p+n} : \kappa(\|\hat{x}\|) \leq \left[\begin{smallmatrix} z \\ \hat{x} \end{smallmatrix}\right]^\top M \left[\begin{smallmatrix} z \\ \hat{x} \end{smallmatrix}\right]$.*

Regarding $y^\top Q_y y = -D^+_{(4.31)} V_y(y)$, Lemma 4.5.9 refers to the class-K function in (4.33). For Q_y in (4.4) or (4.39),

$$\kappa_3(\|\hat{x}\|_2) := \lambda_{\min}(Q_0)\, \|\hat{x}\|_2^2 \leq y^\top Q_y y. \tag{4.35}$$

can be chosen. Rather decisive is whether the Lyapunov equation solution P_y is also $\hat{x}$-pd, as it is required in (4.32) with $V_y(y) = y^\top P_y\, y$.

Lemma 4.5.10. *Let $P_y = P_y^\top$ be a solution of (4.3) for a $\hat{x}$-pd Q_y. If P_y is positive semidefinite, then it is even $\hat{x}$-pd.*

Proof. The result is shown by contradiction. Assume there exists a $y_c = \left[\begin{smallmatrix} z_c \\ \hat{x}_c \end{smallmatrix}\right]$ with $\|\hat{x}_c\| \neq 0$ such that $y_c^\top P_y y_c = 0$. Then $P_y y_c = 0_{n(N+1)}$ (cf. a decomposition

$P_y = C^\top C$ in $y_c^\top P_y y_c = \|Cy_c\|_2^2 = 0$, $C^\top C y_c = 0_{n(N+1)}$), which leads by (4.3) to $y_c^\top Q_y y_c = 0$, contradicting that Q_y is $\hat{x}$-pd. □

Lemma 4.5.11. *Let $P_y = P_y^\top$ be a solution of (4.3) for a $\hat{x}$-pd Q_y. Consider Theorem 4.5.4 in terms of partial asymptotic stability w.r.t. $y^N = \hat{x}$ for the zero equilibrium in (3.3). If P_y is positive semidefinite, then, assuming Condition 4.5.5, $V_y(y) = y^\top P_y\, y$ satisfies the conditions on a partial Lyapunov function in Theorem 4.5.4.*

Proof. In Theorem 4.5.4, (4.32) and (4.33) hold by Lemma 4.5.10 and Lemma 4.5.9. The boundedness condition on $\|f^x(z(t), \hat{x}(t))\|$ in Theorem 4.5.4 is also ensured: due to Condition 4.5.5, the already provable partial stability implies stability, which is accompanied by compactness of the trajectories, and the image under the continuous mapping f^x remains compact. □

As the above lemma is applicable whenever the ODE equilibrium is asymptotically stable, the desired interpretation of the function V_y is obtained.

Theorem 4.5.12. *If A_y is Hurwitz and Condition 4.5.5 applies, then V_y from Section 4.1.1 is a partial Lyapunov function for (3.3).*

Proof. If A_y is Hurwitz, P_y is positive semidefinite by Proposition 4.1.4. As Q_y in Section 4.1.1 is $\hat{x}$-pd, Lemma 4.5.11 applies. □

The focus of this chapter is not preliminary on a stability criterion in terms of P_y because stability of the linear ODE can simply be concluded from the eigenvalues of A_y. Nevertheless, the following result might still be of interest since it shows that V_y must only be tested for positive semidefiniteness. Proving existence of κ_1 in (4.32) is not required due to Lemma 4.5.10.

Theorem 4.5.13. *Assume Condition 4.5.5 holds. Let $P_y = P_y^\top$ be a solution of (4.3) for a $\hat{x}$-pd matrix Q_y (e.g., (4.4) or (4.39)). The zero equilibrium of the*

approximating ODE (3.3) is asymptotically stable if and only if P_y is positive semidefinite.

Proof. If $P_y \succeq 0_{n(N+1)\times n(N+1)}$, then Lemma 4.5.11 applies. Thus, partial asymptotic stability w.r.t. $\hat{x}$ can be proven by Theorem 4.5.4. The latter implies asymptotic stability by Condition 4.5.5. Conversely, if A_y is Hurwitz, then $P_y \succeq 0_{n(N+1)\times n(N+1)}$ because of Proposition 4.1.4. □

4.5.5 Meaning for the RFDE*

As a consequence, if the chosen discretization scheme is stability preserving as discussed in Section 3.7, then the stability-dependent characterization of P_y derived above in Section 4.5.4 is also meaningful for the RFDE.

Corollary 4.5.14. *Assume the discretization scheme satisfies Condition 3.7.1 and 4.5.5. Provided N is sufficiently large, then P_y from Section 4.1.1 is positive semidefinite if and only if the zero equilibrium of the RFDE is asymptotically stable.*

Proof. Theorem 4.5.13 combined with Condition 3.7.1. □

The overall Section 4.5 can be concluded as follows. The function V_y obtained in Section 4.1.1 does not necessarily qualify as a classical Lyapunov function. Instead, it is a partial Lyapunov function for a system in which proving partial stability is already sufficient for proving stability.

* The author has prepublished Section 4.5.5 in [S2], ©2024 IEEE.

4.6 Convergence

Finally, it remains to discuss convergence aspects of the proposed approach for complete-type LK functionals in terms of an increasing discretization resolution N of the underlying ODE approximation. In any case, refined results with various enlarged values of N should be considered to get an impression of the possible reliability of obtained results.

4.6.1 Scheme-Dependent Improvements

A splitting of the LK functional is introduced below in order to provide improved results for the Lyapunov equation from the Chebyshev collocation system matrix. Although such a splitting approach is not required in the practical implementation of the Legendre-tau-based approach, the proposed splitting will still be helpful when it comes to the proof of convergence. Moreover, for the Legendre-tau-based approach an improved treatment of Q_2 from (1.14) is introduced in Section 4.6.1.2.

4.6.1.1 Chebyshev collocation*

Consider the ODE-based approach with the Chebyshev collocation method. To improve the convergence properties (indicated in Figure 4.2), the present section proposes to transform the problem of approximating $V(\phi)$ to a problem of approximating a modified $V_0(\phi)$ with Q_1 and Q_2 being zero. To this end, the shift matrices in the following splitting lemma are chosen as

$$\tilde{Q}_1 = Q_1 \quad \text{and} \quad \tilde{Q}_2 = Q_2. \tag{4.36}$$

The idea is closely related to the derivation of complete-type functionals in [110, Thm. 2.11].

* The author has prepublished Section 4.6.1.1 in [S2], ©2024 IEEE.

Lemma 4.6.1 (Splitting). *For $Q_0, Q_1, Q_2 \in \mathbb{R}^{n\times n}$, let $V(\phi) = V(\phi;\ Q_0,\ Q_1,\ Q_2)$ denote a solution of (1.14). Then*

$$
\begin{aligned}
V(\phi;\ Q_0,\ Q_1,\ Q_2\ ) &= V_0(\phi) + V_1(\phi) + V_2(\phi) \qquad (4.37)\\
\textit{with}\qquad V_0(\phi) &= V\big(\phi;\ (Q_0+\tilde{Q}_1+h\tilde{Q}_2),\ (Q_1-\tilde{Q}_1),\ (Q_2-\tilde{Q}_2)\ \big)\\
V_1(\phi) &= V(\phi; -\tilde{Q}_1, \tilde{Q}_1, 0_{n\times n})\\
&= \int_{-h}^{0} \phi^\top(\eta)\tilde{Q}_1\phi(\eta)\,\mathrm{d}\eta\\
V_2(\phi) &= V(\phi; -h\tilde{Q}_2, 0_{n\times n}, \tilde{Q}_2)\\
&= \int_{-h}^{0} \phi^\top(\eta)(h+\eta)\tilde{Q}_2\phi(\eta)\mathrm{d}\eta
\end{aligned}
$$

for arbitrarily chosen shift matrices $\tilde{Q}_1, \tilde{Q}_2 \in \mathbb{R}^{n\times n}$.

Proof. Using $\phi(\eta) = x_t(\eta) = x(t+\eta)$, the derivatives of V_1 and V_2 are

$$
\begin{aligned}
D_f^+ V_1(x_t) &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{t-h}^{t} x^\top(\xi)\tilde{Q}_1 x(\xi)\,\mathrm{d}\xi\\
&= x^\top(t)\tilde{Q}_1 x(t) - x^\top(t-h)\tilde{Q}_1 x(t-h),\\
D_f^+ V_2(x_t) &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{t-h}^{t} x^\top(\xi)\big((h+\xi-t)\tilde{Q}_2\big)x(\xi)\,\mathrm{d}\xi\\
&= hx^\top(t)\tilde{Q}_2 x(t) - \int_{t-h}^{t} x^\top(\xi)\tilde{Q}_2 x(\xi)\,\mathrm{d}\xi.
\end{aligned}
$$

They compensate in (4.37) the difference between $D_f^+ V_0(x_t)$ and the desired $D_f^+ V(x_t)$ from (1.14). □

The first term $V_0(\phi)$ in (4.37) with $\tilde{Q}_1 = Q_1, \tilde{Q}_2 = Q_2$ can be approximated by $y^\top P_{y,0}\, y$ from a Lyapunov equation with Q_0 in (4.4) being replaced by $Q_0+Q_1+hQ_2$, and Q_1 and Q_2 being replaced by zero. Since $V_1(\phi)$ and $V_2(\phi)$ in (4.37) are

analytically known, these terms can be treated by a numerical integration. Their contributions are added on the (block-)diagonal of $P_{y,0}$, i.e. $V(\phi) \approx y^\top P_y\, y$,

$$P_y = P_{y,0} + \mathrm{diag}((w_k)_k) \otimes Q_1 + \mathrm{diag}((w_k(h+\tilde{\theta}_k))_k) \otimes Q_2, \tag{4.38}$$

where w_k are integration weights, see Section 4.2.1.

4.6.1.2 Legendre tau*

A separate numerical treatment of V_1 and V_2 in (4.37) is not required if the Legendre-tau-based approach is used. However, if Q_2 is nonzero, the following modification of Q_y in (4.4) should be used in (4.6) or (4.3)

$$\begin{aligned} &Q_y = \mathrm{blkdiag}(Q_1, 0_{n(N-1)\times n(N-1)}, Q_0) + T_{cy}^\top Q_{c,2} T_{cy} \\ &\text{with } Q_{c,2} := \mathrm{diag}([\ (\tfrac{h}{2}\tfrac{2}{2k+1})_{k\in\{0,\dots,N-1\}},\ h\ ]) \otimes Q_2 \end{aligned} \tag{4.39}$$

(the right lower component hQ_2 in $Q_{c,2}$ is motivated by Lemma 4.6.3 below).

Despite of not being treated separately in the numerical approach, the arising contributions for $V_1(\phi)$ and $V_2(\phi)$ within the approximation of $V(\phi)$ are still of interest for the proofs in the next sections. They can be obtained by solving Lyapunov equations with $Q_{0,1,2}$ being replaced by the matrices behind the semicolon in $V_1(\phi) = V(\phi;\dots)$ and $V_2(\phi) = V(\phi;\dots)$ from Lemma 4.6.1. In the following two lemmas, the corresponding solutions are explicitly derived by solving the Lyapunov equation (4.6) in Legendre coordinates analytically. The results show that the Legendre-tau-based approximations of $V_1(\phi)$ and $V_2(\phi)$ give even the exact value for any ϕ that is a polynomial of degree $N-1$ or less.

Lemma 4.6.2. *The Legendre-tau-based approximation of $V_1(\phi)$ in Lemma 4.6.1 becomes $V_1(\tilde{\phi}^{(N-1)})$, where $\tilde{\phi}^{(N-1)}(\theta) = \sum_{k=0}^{N-1} \tilde{c}^k p_k(\vartheta(\theta))$ is the $(N-1)$-th order Legendre series truncation of $\phi(\theta)$, see (3.44).*

* The author has prepublished Section 4.6.1.2 in [S2], ©2024 IEEE.

Proof. For $Q_y = \operatorname{diag}([1, 0_{1\times n(N-1)}, -1]) \otimes \tilde{Q}_1$, and thus $Q_c^{jk} = (-1 + (-1)^{j+k})\tilde{Q}_1$, it can be verified that

$$P_c = \operatorname{diag}([(\tfrac{h}{2}\tfrac{2}{2k+1})_{k\in\{0,\dots,N-1\}}, 0]) \otimes \tilde{Q}_1 \tag{4.40}$$

is a solution of (4.6). Hence, $V_c(c) = c^\top P_c c = \sum_{k=0}^{N-1} \frac{h}{2}\frac{2}{2k+1} c_k^\top \tilde{Q}_1 c_k$. Equivalence with $V_1(\tilde{\phi}^{(N+1)}) = \int_{-h}^{0} (\sum_{j=0}^{N-1} \tilde{c}^j p_j(\vartheta(\theta)))^\top \tilde{Q}_1 (\sum_{k=0}^{N-1} \tilde{c}^k p_k(\vartheta(\theta)))\,\mathrm{d}\theta$ follows from (3.47) and $\int_{-1}^{1} p_j(\vartheta)p_k(\vartheta)\,\mathrm{d}\vartheta = \frac{2}{2k+1}\delta_{jk}$, see Lemma 3.5.3. □

Lemma 4.6.3. *Provided (4.39) is used, the Legendre-tau-based approximation of $V_2(\phi)$ in Lemma 4.6.1 becomes $V_2(\tilde{\phi}^{(N-1)})$ with $\tilde{\phi}^{(N-1)}$ as in Lemma 4.6.2.*

Proof. In view of the definition of $V_2(\phi)$, let $Q_0 := -h\tilde{Q}_2$ and $Q_2 := \tilde{Q}_2$. Consider $Q_c = Q_{c,0} + Q_{c,2}$ with $Q_{c,0} = T_{yc}^\top(\operatorname{diag}([0_{1\times nN}, 1]) \otimes Q_0)T_{yc} = 1_{(N+1)\times(N+1)} \otimes Q_0$ and $Q_{c,2} = \operatorname{diag}([(\frac{1}{2k+1})_{k\in\{0,\dots,N-1\}}, 1]) \otimes hQ_2$ from (4.39). It can be verified that P_c with

$$P_c^{jk} = \begin{cases} (\frac{h}{2})^2 \frac{2}{2j+1}\frac{k}{2k+1}\tilde{Q}_2 & \text{if } j = k-1 < N-1, \\ (\frac{h}{2})^2 \frac{2}{2j+1}\tilde{Q}_2 & \text{if } j = k < N, \\ (\frac{h}{2})^2 \frac{2}{2j+1}\frac{k+1}{2k+1}\tilde{Q}_2 & \text{if } j = k+1 < N, \\ 0_{n\times n} & \text{otherwise,} \end{cases} \tag{4.41}$$

solves (4.6). The equality $c^\top P_c c = V_2(\sum_{k=0}^{N-1} c^k p_k(\vartheta(\theta)))$ is shown by using the three-term recurrence relation [88, (4.17)] $\vartheta p_k(\vartheta) = \frac{k}{2k+1}p_{k-1}(\vartheta) + \frac{k+1}{2k+1}p_{k+1}(\vartheta)$ in V_2. □

4.6.2 Convergence Towards the Functional

For the numerical approach that is presented Section 4.1, the following convergence statement is desirable.

Condition 4.6.4. *For any given $\phi \in C([-h,0],\mathbb{R}^n)$, the scalar value $V_y(y)$ converges to $V(\phi)$ as N increases.*

Henceforth, the notation $y = \pi_y(\phi)$ intends to emphasize that the discretization $y \in \mathbb{R}^{n(N+1)}$ is uniquely determined from $\phi \in C$ (depending on the discretization scheme). Moreover, to keep track of the discretization resolution N, a superscript $^{[N]}$ is added, e.g., in $V_y^{[N]}(\cdot) = V_y(\cdot)$ and $\pi_y^{[N]}(\cdot) = \pi_y(\cdot)$. Consequently, the desired convergence statement from Condition 4.6.4 can be rewritten as

$$\forall \phi \in C: \quad V_y^{[N]}(\pi_y^{[N]}(\phi)) \to V(\phi), \quad (N \to \infty). \tag{4.42}$$

To prove this statement is the objective of the present section. Motivated by the numerical results in Section 4.4, it focuses on the Legendre tau method. Moreover this choice benefits from existing convergence proofs for the approximation of algebraic Riccati equations from the context of optimal control [70, 99, 97].

4.6.2.1 Operator-based description*

The following considerations rely on the fact that any argument $\phi \in C$ for $V(\phi)$ gives rise to an element

$$\left[\begin{smallmatrix} \phi \\ \phi(0) \end{smallmatrix}\right] \in C \times \mathbb{R}^n \ \subset \ L_2 \times \mathbb{R}^n = M_2 \tag{4.43}$$

in the product space $M_2 = L_2([-h,0],\mathbb{R}^n) \times \mathbb{R}^n$ already encountered in Section 3.2.2. Note that $(M_2, \langle\cdot,\cdot\rangle_{M_2})$ is a Hilbert space with the natural inner product

$$\left\langle \left[\begin{smallmatrix} \phi_1 \\ r_1 \end{smallmatrix}\right], \left[\begin{smallmatrix} \phi_2 \\ r_2 \end{smallmatrix}\right] \right\rangle_{M_2} = \int_{-h}^{0} \phi_1^\top(\theta)\, \phi_2(\theta)\, \mathrm{d}\theta + r_1^\top r_2, \tag{4.44}$$

* The author has prepublished Section 4.6.2.1 in [S2], ©2024 IEEE.

$\phi_{1,2} \in L_2$, $r_{1,2} \in \mathbb{R}^n$. Similarly to the well-known $V_{\mathbb{R}^n}(x) = \langle x, Px\rangle_{\mathbb{R}^n} = x^\top Px$ in the finite-dimensional ODE setting for $x \in \mathbb{R}^n$, a complete-type LK functional can be written as

$$V(\phi) = V_{M_2}(\left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right]) = \left\langle \left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right], \mathscr{P}\left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right]\right\rangle_{M_2} \tag{4.45}$$

with a self-adjoint operator $\mathscr{P}\colon M_2 \to M_2$. Consider the splitting $V = V_0 + V_{12}$ with $V_{12} = V_1 + V_2$ from Lemma 4.6.1 ($\tilde{Q}_1 = Q_1, \tilde{Q}_2 = Q_2$). For the first part, which becomes

$$V_0(\phi) = \langle \left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right], \mathscr{P}_0\left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right]\rangle_{M_2}, \tag{4.46}$$

the self-adjoint operator $\mathscr{P}_0\colon M_2 \to M_2$ is described by suboperators on L_2 and $\mathbb{R}^n$ according to

$$\mathscr{P}_0 \begin{bmatrix}\phi\\ r\end{bmatrix} = \begin{bmatrix}\mathscr{P}_{zz}\phi + \mathscr{P}_{zx} r\\ \mathscr{P}_{xz}\phi + \mathscr{P}_{xx} r\end{bmatrix} = \begin{bmatrix}v\\ w\end{bmatrix}, \text{ with} \tag{4.47}$$

$$\begin{bmatrix}v(\theta)\\ w\end{bmatrix} = \begin{bmatrix}\int_{-h}^0 P_{\mathrm{zz}}(\theta,\eta)\,\phi(\eta)\,\mathrm{d}\eta + P_{\mathrm{zx}}(\theta)\,r\\ \int_{-h}^0 P_{\mathrm{xz}}(\eta)\,\phi(\eta)\,\mathrm{d}\eta + P_{\mathrm{xx}}\,r\end{bmatrix}. \tag{4.48}$$

Thus, (4.11a) is regained by (4.46), using (4.47) with $r = \phi(0)$,

$$V_0(\phi) \overset{(4.44)}{=} \int_{-h}^0 \phi^\top(\theta)\,v(\theta)\,\mathrm{d}\theta + \phi^\top(0)\,w \tag{4.49}$$

$$\begin{aligned}\overset{(4.48)}{=} &\int_{-h}^0 \phi^\top(\theta)\Big(\int_{-h}^0 P_{\mathrm{zz}}(\theta,\eta)\,\phi(\eta)\,\mathrm{d}\eta + P_{\mathrm{zx}}(\theta)\,\phi(0)\Big)\,\mathrm{d}\theta\\ &+ \phi^\top(0)\Big(\int_{-h}^0 P_{\mathrm{xz}}(\eta)\,\phi(\eta)\,\mathrm{d}\eta + P_{\mathrm{xx}}\,\phi(0)\Big)\end{aligned} \tag{4.50}$$

(to be more precise, (4.11a) with $P_{\mathrm{zz,diag}}(\theta) \equiv 0_{n\times n}$). The missing part $V_{12} = V_1 + V_2$ in (4.37) can also be written as

$$V_{12}(\phi) = \int_{-h}^{0} \phi^\top(\theta) \left(Q_1 + (h+\theta)Q_2\right) \phi(\theta)\,\mathrm{d}\theta \tag{4.51}$$

$$= \langle \left[\begin{smallmatrix} \phi \\ \phi(0) \end{smallmatrix}\right], \mathscr{P}_{12} \left[\begin{smallmatrix} \phi \\ \phi(0) \end{smallmatrix}\right] \rangle_{M_2} \tag{4.52}$$

based on the multiplication operator

$$\mathscr{P}_{12} \begin{bmatrix} \phi \\ r \end{bmatrix} = \begin{bmatrix} \mathscr{P}_{zz,\mathrm{diag}}\phi \\ 0_n \end{bmatrix} = \begin{bmatrix} v \\ 0_n \end{bmatrix}, \text{ with} \tag{4.53}$$

$$v(\theta) = P_{\mathrm{zz,diag}}(\theta)\phi(\theta) = \left(Q_1 + (h+\theta)Q_2\right)\phi(\theta).$$

Nevertheless, V_{12} will be treated separately[9].

4.6.2.2 Convergence of the First Part*

In this section, convergence towards V_0 shall be shown. The underlying operator $\mathscr{P}_0$ in (4.46) satisfies an operator-valued Lyapunov equation, cf. [47, 156]. The right-hand side of that Lyapunov-equation is based on the right-hand side of (1.14). Because of the splitting approach, the latter is $D^+_{(1.4)}V_0(x_t) = x^\top(t)\tilde{Q}x(t)$ with $\tilde{Q} = Q_0 + Q_1 + hQ_2$, or, for $x_t = \phi$,

$$D^+_{(1.4)}V_0(\phi) = -\phi^\top(0)\,\tilde{Q}\,\phi(0) = -\langle \left[\begin{smallmatrix} \phi \\ \phi(0) \end{smallmatrix}\right], \mathscr{Q} \left[\begin{smallmatrix} \phi \\ \phi(0) \end{smallmatrix}\right] \rangle_{M_2}, \tag{4.54}$$

[9] The term $\phi^\top(-h)Q_1\phi(-h)$ would require an unbounded operator $\mathscr{Q}$ in the Lyapunov equation (4.55). Moreover, $\mathscr{P}_{12}$ is not compact.

* The author has prepublished Section 4.6.2.2 in [S2], ©2024 IEEE.

$\mathscr{Q}\left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right] = \left[\begin{smallmatrix}0_{n_{[-h,0]}}\\ \tilde{Q}\,\phi(0)\end{smallmatrix}\right]$. Therefore, the operator-valued Lyapunov equation for the self-adjoint operator $\mathscr{P}_0 = \mathscr{P}_0^*$ reads

$$\underbrace{\langle \mathsf{x}, \mathscr{P}_0\mathscr{A}\mathsf{x}\rangle_{M_2} + \langle \mathsf{x}, \mathscr{A}^*\mathscr{P}_0\mathsf{x}\rangle_{M_2}}_{=2\langle \mathsf{x}, \mathscr{P}_0\mathscr{A}\mathsf{x}\rangle_{M_2}} = -\langle \mathsf{x}, \mathscr{Q}\mathsf{x}\rangle_{M_2}, \tag{4.55}$$

$\forall \mathsf{x} \in D(\mathscr{A}) \subset M_2$, cf. [47, 156], where $\mathscr{A}$ is given by (3.12) and and $\mathscr{A}^*$ is its adjoint[10,11].

The ODE-based approach in Section 4.1.1 yields an approximation $V_0(\phi) \approx V_{y,0}(y) = y^\top P_{y,0}\, y$, or, in the notation of (4.42), $V_{y,0}^{[N]}(\pi_y^{[N]}(\phi))$. Similarly to the exact $V_0(\phi)$ in (4.46), this approximation can be described via

$$V_{y,0}^{[N]}(\pi_y^{[N]}(\phi)) = \langle \left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right], \mathscr{P}_0^{[N]} \left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right] \rangle_{M_2} \tag{4.56}$$

with an approximated operator $\mathscr{P}_0^{[N]}$, for which the matrix $P_{0,y}$, respectively $P_{0,c}$ or $P_{0,\chi}$, is a coordinate representation, see Appendix A. Moreover, similarly to the exact operator $\mathscr{P}_0$ from (4.55), this approximated operator $\mathscr{P}_0^{[N]}$ also satisfies an operator-valued Lyapunov equation,

$$2\langle \mathsf{x},\ \mathscr{P}_0^{[N]}\mathscr{A}^{[N]}\mathsf{x}\rangle_{M_2} = -\langle \mathsf{x},\ \mathscr{Q}\mathsf{x}\rangle_{M_2}. \tag{4.57}$$

The latter, however, only relies on the approximation $\mathscr{A}^{[N]}$ from the Legendre tau method, see Section 3.5.3, instead of the original operator $\mathscr{A}$. See Appendix A for the relation between the operator-valued and the matrix-valued Lyapunov equation.

[10] See (A.92) for an explicit description of the adjoint of $\mathscr{A}$.

[11] For (4.55) to make sense, (v, w) in (4.47) should be an element of $D(\mathscr{A}^*)$ from (A.92). Indeed, in (4.11b), $P_{\mathrm{zz}}(-h,\eta) = A_1^\top P_{\mathrm{xz}}(\eta)$ and $P_{\mathrm{zx}}(-h) = A_1^\top P_{\mathrm{xx}}$.

It has to be shown that, $\forall \phi \in C$, the scalar value $V_0(\phi)$ in (4.46) is indeed the limit of its approximations in (4.56) as $N \to \infty$. In terms of the operators, weak[12] operator convergence $\mathscr{P}_0^{[N]} \overset{\text{weakly}}{\to} \mathscr{P}_0$ suffices for that objective.

Lemma 4.6.5. *Let (4.56) describe a Legendre-tau-based result for $V_0(\phi)$. Assume $\{\|\mathscr{P}_0^{[N]}\|\}_N$ is bounded[13], and the existence and uniqueness conditions from Lemma 4.1.2 and Remark 4.1.3 hold. Then $\mathscr{P}_0^{[N]}$ converges weakly to $\mathscr{P}_0$ as $N \to \infty$.*

Proof. See [99, Thm. 5.1 (i)] with zero input operator and the uniqueness conditions from Section 4.1.2. For the sake of plausibility, the proof is briefly sketched: As outlined above, the operator $\mathscr{P}_0$ is a self-adjoint solution of the operator-valued Lyapunov equation (4.55), whereas the approximation $\mathscr{P}_0^{[N]}$ only satisfies (4.57) from the approximated system dynamics. According to [70, Thm. 6.5], due to the boundedness of $\{\|\mathscr{P}_0^{[N]}\|\}_N$, there exists a subsequence $\{\mathscr{P}_0^{[N_k]}\}_k$ that has a weak limit $\mathscr{P}_0^{[N_k]} \overset{\text{weakly}}{\to} \mathscr{P}_0^{[\infty]}$, as $k \to \infty$. Moreover, according to Lemma 3.5.5, $\mathscr{A}^{[N]} \overset{\text{strongly}}{\to} \mathscr{A}$ on the restriction to $D(\mathscr{A}^2)$. The product of a weakly convergent operator and a strongly convergent operator converges weakly and thus the left-hand side in (4.57) converges according to

$$\forall \mathsf{x} \in D(\mathscr{A}^2): \quad \left\langle \mathsf{x}, \mathscr{P}_0^{[N_k]} \mathscr{A}^{[N_k]} \mathsf{x} \right\rangle_{M_2} \to \left\langle \mathsf{x}, \mathscr{P}_0^{[\infty]} \mathscr{A} \mathsf{x} \right\rangle_{M_2}, \quad \text{as } k \to \infty,$$

[12] The operator sequence $\{\mathscr{P}^{[N]}\}_N$ converges weakly to $\mathscr{P}$ if
$\forall \mathsf{z}, \mathsf{x} \in M_2$: $\lim_{N\to\infty} \langle \mathsf{z}, \mathscr{P}^{[N]} \mathsf{x} \rangle_{M_2} = \langle \mathsf{z}, \mathscr{P} \mathsf{x} \rangle_{M_2}$ (i.e., $\forall \mathsf{x} \in M_2 : \mathscr{P}^{[N]} \mathsf{x} \overset{\text{weakly}}{\to} \mathscr{P} \mathsf{x}$).
It converges strongly if $\forall \mathsf{x} \in M_2 : \lim_{N\to\infty} \|\mathscr{P}^{[N]} \mathsf{x} - \mathscr{P} \mathsf{x}\|_{M_2} = 0$.
The implications 'operator norm conv.' $\Rightarrow$ 'strong conv.' $\Rightarrow$ 'weak conv.' hold.

[13] The operator norm of $\mathscr{P}_0^{[N]}$ can be computed from its coordinate matrix (considering, e.g., mixed coordinates) $P_{0,\chi}$ via $\|\mathscr{P}_0^{[N]}\| = \|(H_\chi^{-1/2} \otimes I_n) P_{0,\chi} (H_\chi^{-1/2} \otimes I_n)\|_2$, with H_χ given by (A.21). See (A.112) (where $n = 1$).

if it is restricted to $D(\mathscr{A}^2)$. However, $D(\mathscr{A}^2)$ is a core for $\mathscr{A}$ (i.e., dense in $D(\mathscr{A})$ for the graph norm, see [63, Def. 1.6]). Therefore, analogously to [99, Thm. 5.1 (i)], it can be concluded that the limit of (4.57)

$$2\langle \mathsf{x}, \mathscr{P}_0^{[\infty]} \mathscr{A} \mathsf{x} \rangle_{M_2} = -\langle \mathsf{x}, \mathscr{Q} \mathsf{x} \rangle_{M_2} \tag{4.58}$$

does not only hold for $\mathsf{x} \in D(\mathscr{A}^2)$ but for all $\mathsf{x} \in D(\mathscr{A})$. Hence, the weak limit $\mathscr{P}_0^{[\infty]}$ indeed satisfies the original Lyapunov equation (4.55), and thus $\mathscr{P}_0^{[\infty]} = \mathscr{P}_0$. Since all weakly convergent subsequences $\{\mathscr{P}_0^{[N_k]}\}_k \subseteq \{\mathscr{P}_0^{[N]}\}_N$ have the same weak limit $\mathscr{P}_0^{[\infty]} = \mathscr{P}_0$, and since for any subsequence of $\{\mathscr{P}_0^{[N]}\}_N$ again a convergent subsequence exists, it is argued in [70, Thm. 6.7] that the whole sequence $\{\mathscr{P}_0^{[N]}\}_N$ must converge weakly to $\mathscr{P}_0^{[\infty]} = \mathscr{P}_0$. □

In fact, this result is not at all special to the Legendre tau method. An alternative proof from [70, Thm. 6.7] applies to any discretization scheme that satisfies standard conditions proving convergence of numerical solutions for $(x_t, x(t))$ in M_2. Lemma 4.6.5 relies on uniform boundedness and existence assumptions. In the following, it is shown that these can be ignored in the case of an exponentially stable RFDE equilibrium. Nevertheless, while simplifying the considerations, stability of the equilibrium is no necessary condition in the above derivations.

Lemma 4.6.6. *If the RFDE equilibrium is exponentially stable, then the assumptions in Lemma 4.6.5 hold.*

Proof. Let $\mathscr{T}(t)\colon M_2 \to M_2; \left[\begin{smallmatrix} x_0 \\ x_0(0) \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} x_t \\ x(t) \end{smallmatrix}\right] = \mathscr{T}(t)\left[\begin{smallmatrix} x_0 \\ x_0(0) \end{smallmatrix}\right]$ be the solution operator, and $\mathscr{T}^{[N]}(t)$ its approximation (represented by $\mathrm{e}^{A_y^{[N]}t}$). Due to the stability preservation property from [97, Thm. 5.3], $\exists M \geq 1, \beta > 0, \bar{N} \in \mathbb{N}$, such that $\forall N \geq \bar{N} : \|\mathscr{T}^{[N]}(t)\| \leq M\mathrm{e}^{-\beta t}$. Therefore, the improper integral formula

$$\mathscr{P}^{[N]}\mathsf{x} = \int_0^\infty (\mathscr{T}^{[N]})^*(s)\, \mathscr{Q}\, \mathscr{T}^{[N]}(s)\, \mathsf{x}\, \mathrm{d}s \tag{4.59}$$

is applicable, see, e.g., [70]. Thus, with $\|\mathscr{Q}\| = \|\tilde{Q}\|_2$, the operators $\mathscr{P}^{[N]}$ are uniformly bounded by $\|\mathscr{P}^{[N]}\| \leq \int_0^\infty \|\tilde{Q}\|_2 \|\mathscr{T}^{[N]}(s)\|^2 \,\mathrm{d}s \leq \|\tilde{Q}\|_2 \frac{M^2}{2\beta}$. Moreover, the existence and uniqueness assumptions hold by Proposition 4.1.4. □

The convergence towards $V_0(\phi)$ does not require more than the thus established weak convergence $\mathscr{P}_0^{[N]} \overset{\text{weakly}}{\to} \mathscr{P}_0$. However, the following stronger result will become helpful in Section 4.6.3.

Lemma 4.6.7. *Let (4.56) describe a Legendre-tau-based result for $V_0(\phi)$. If the RFDE equilibrium is exponentially stable, then $\mathscr{P}_0^{[N]}$ converges in operator norm to $\mathscr{P}_0$, i.e., it holds $\|\mathscr{P}_0^{[N]} - \mathscr{P}_0\| \to 0$ as $N \to \infty$.*

Proof. See [70, Thm. 6.9], where even convergence in the trace norm [70, p. 111] is proven. The result requires that not only the approximations of the solution operator $\mathscr{T}(t)$ converge strongly[12], but also those of its adjoint $\mathscr{T}^*(t)$, which for the Legendre tau method is proven in [97, Thm. 2.2]. □

4.6.2.3 Convergence of the Overall Result*

To prove Condition 4.6.4 on convergence towards $V = V_0 + V_{12}$, it only remains to include V_{12}.

Theorem 4.6.8. *If the RFDE equilibrium is exponentially stable or, more generally, if the assumptions of Lemma 4.6.5 hold, then Condition 4.6.4 applies for the Legendre-tau-based approach (provided the Lyapunov equation right-hand side from (4.39) is used).*

Proof. Since P_y depends linearly on Q_y in the Lyapunov equation (4.3), the approximation of V is the superposition of the approximations of V_0 and $V_{12} = V_1 + V_2$ from Lemma 4.6.1. For the first one, the convergence, $\forall \phi \in C$:

* The author has prepublished Section 4.6.2.3 in [S2], ©2024 IEEE.

$V_{y,0}^{[N]}(\pi_y^{[N]}(\phi)) \to V_0(\phi)$ as $N \to \infty$, is a consequence of the weak convergence of $\mathscr{P}_0^{[N]}$ proven in Lemma 4.6.5. Concerning the second one, the lemmas in Section 4.6.1.2 show that $V_{12}(\phi)$ is approximated by $V_{y,12}^{[N]}(\pi_y^{[N]}(\phi)) = V_{12}(\tilde{\phi}^{(N-1)})$, where $\tilde{\phi}^{(N-1)}$ is a Legendre series truncation of ϕ. The convergence, $\forall \phi \in C: \quad V_{12}(\tilde{\phi}^{(N-1)}) \to V_{12}(\phi)$ as $N \to \infty$, follows from the L_2-convergence of the involved Legendre series truncation, $\|\phi - \tilde{\phi}^{(N-1)}\|_{L_2} \to 0$ as $N \to \infty$ [68, Thm. 6.2.3], combined with the continuity[14] of V_{12} in L_2. □

4.6.3 Quadratic Lower Bound on the Functional*

In the following, it will be proven that, for $N \to \infty$, the quadratic lower bound on the approximation gives also a valid quadratic lower bound on the functional. This holds for any discretization scheme satisfying Condition 4.6.4. Moreover, for the Legendre tau method, the thus obtained bound will be shown to be tight, meaning that the largest possible coefficient k_1 in (1.16) is obtained.

For any discretization resolution N, the largest possible coefficient $k_1^{[N]}$ for the bound (4.18) on the approximation $V_y^{[N]}$ is given by (4.23). Note that $k_1^{[N]}$ and, similarly, the largest possible coefficient $k_1 = k_1^{\mathrm{opt}}$ for the bound (1.16) on the functional V are defined by

$$k_1^{[N]} = \min_{\substack{z \in \mathbb{R}^{nN} \\ \hat{x} \in \mathbb{R}^n \setminus \{0_n\}}} \frac{1}{\|\hat{x}\|_2^2} V_y^{[N]}\left(\begin{bmatrix} z \\ \hat{x} \end{bmatrix}\right), \qquad k_1^{\mathrm{opt}} = \inf_{\substack{\phi \in C \\ \phi(0) \neq 0_n}} \frac{1}{\|\phi(0)\|_2^2} V(\phi). \tag{4.60}$$

[14] A quadratic form $V(x) = \langle x, \mathscr{P}x \rangle_X$, $\mathscr{P} = \mathscr{P}^*$, in a Hilbert space X is continuous if $\exists k > 0 : \langle x, \mathscr{P}x \rangle_X \leq k\|x\|_X^2$, which by $\inf k = \|\mathscr{P}\|$ holds if $\mathscr{P}$ is bounded. Note that $V_{12}(\phi) = \langle \phi, \mathscr{P}_{zz,\mathrm{diag}}\phi \rangle_{L_2} \leq (\|Q_1\|_2 + h\|Q_2\|_2)\|\phi\|_{L_2}^2$. For $V_{M_2}(\mathrm{x}) = \langle \mathrm{x}, \mathscr{P}\mathrm{x} \rangle_{M_2} \leq k\|\mathrm{x}\|_{M_2}^2$ see [110, p. 65].

* The author has prepublished Section 4.6.3 in [S2], ©2024 IEEE.

However, since both the functional and its approximation are quadratic, with $V(c\phi) = c^2 V(\phi)$ for any $c \in \mathbb{R}$ in (4.11a) and $V_y^{[N]}(cy) = c^2 V_y^{[N]}(y)$ in (4.1), definition (4.60) simplifies to

$$k_1^{[N]} = \min_{\substack{z\in\mathbb{R}^{nN} \\ \hat{x}\in\mathbb{R}^n\setminus\{0_n\}}} V_y^{[N]}\big(\tfrac{1}{\|\hat{x}\|_2}\left[\begin{smallmatrix} z \\ \hat{x}\end{smallmatrix}\right]\big), \qquad k_1^{\mathrm{opt}} = \inf_{\substack{\phi\in C \\ \phi(0)\neq 0_n}} V\big(\tfrac{1}{\|\phi(0)\|_2}\phi\big). \tag{4.61}$$

Theorem 4.6.9. *If Condition 4.6.4 holds, then* $k_1 = \limsup_{N\to\infty} k_1^{[N]}$ *is a valid quadratic-lower-bound coefficient in (1.16).*

Proof. Let ϕ_δ give a $V(\phi_\delta)$ that is arbitrarily close to the infimum in (4.61) according to

$$\forall \delta > 0, \exists \phi_\delta \in C, \|\phi_\delta(0)\|_2 = 1 : \quad V(\phi_\delta) < k_1^{\mathrm{opt}} + \delta. \tag{4.62}$$

The assumed convergence (4.42), i.e., $\forall \phi \in C$, $\forall \varepsilon > 0$, $\exists \bar{N}(\varepsilon, \phi) \in \mathbb{N}$, $\forall N \geq \bar{N}(\varepsilon, \phi) : |V_y^{[N]}(\pi_y^{[N]}(\phi)) - V(\phi)| < \varepsilon$, shows that

$$\forall N \geq \bar{N}(\tfrac{\varepsilon}{2}, \phi_\delta) : \quad |V_y^{[N]}(\pi_y^{[N]}(\phi_\delta)) - \underbrace{V(\phi_\delta)}_{\overset{(4.62)}{<} k_1^{\mathrm{opt}}+\delta} | < \tfrac{\varepsilon}{2}, \tag{4.63}$$

and thus, $\forall N \geq \bar{N}(\frac{\varepsilon}{2}, \phi_\delta)$:

$$k_1^{[N]} \overset{(4.61)}{=} \min_{\substack{z\in\mathbb{R}^{nN}, \\ \hat{x}\in\mathbb{R}^n\setminus\{0_n\}}} V_y^{[N]}(\tfrac{1}{\|\hat{x}\|_2}\left[\begin{smallmatrix} z \\ \hat{x}\end{smallmatrix}\right]) \leq V_y^{[N]}(\pi_y^{[N]}(\phi_\delta)) \overset{(4.63)}{<} V(\phi_\delta) + \tfrac{\varepsilon}{2} \overset{(4.62)}{<} k_1^{\mathrm{opt}} + \delta + \tfrac{\varepsilon}{2}. \tag{4.64}$$

Choosing $\delta = \frac{\varepsilon}{2}$, (4.64) becomes $k_1^{[N]} < k_1^{\mathrm{opt}} + \varepsilon$. Hence, $\limsup_{N\to\infty} k_1^{[N]} \leq k_1^{\mathrm{opt}}$. Any $k_1 \leq k_1^{\mathrm{opt}}$ is admissible in (1.16). □

For the Legendre tau method, it will be proven below that $k_1^{[N]}$ converges to the largest admissible coefficient k_1^{opt}. The proof involves the following assumption on the arguments of the minimum in (4.61): For any N, consider a vector

$y^{[N]} = \left[\begin{smallmatrix} z^{[N]} \\ \hat{x}^{[N]} \end{smallmatrix}\right]$, with $\|\hat{x}^{[N]}\|_2 = 1$, such that $V_y^{[N]}(y^{[N]}) = k_1^{[N]}$. By (3.48), any $y^{[N]}$, respectively $c = T_{cy} y^{[N]}$, represents a function $\phi_{\mathrm{d}}^{[N]}$ (not expected to be continuous at $\theta = 0$). The assumption below is that $\phi_{\mathrm{d}}^{[N]}$ remains uniformly bounded in L_2, which, however, could numerically[15] be confirmed for all tested examples that give a nonzero k_1.

Theorem 4.6.10. *Consider the Legendre tau method with (4.39). As described above, for $\phi_{\mathrm{d}}^{[N]}$ being related to $k_1^{[N]}$, assume that $\exists \beta > 0$, $\forall N : \|\phi_{\mathrm{d}}^{[N]}\|_{L_2} < \beta$. Then the quadratic-lower-bound coefficient $k_1^{[N]}$ from Corollary 4.3.3 converges to the largest possible quadratic-lower-bound coefficient on the functional in (1.16).*

Proof. Henceforth, C_{d} denotes the set of functions $\phi\colon [-h, 0] \to \mathbb{R}^n$ that are continuous on $[-h, 0)$ and possibly have a jump discontinuity at the end point $\phi(0^-) \neq \phi(0)$. Note that $\phi_{\mathrm{d}}^{[N]} \in C_{\mathrm{d}}$. The functional $V\colon C \to \mathbb{R}$ can straightforwardly be extended to arguments in C_{d} since $V(\phi) = V_{M_2}((\phi, \phi(0)))$ holds by (4.45), which, in fact, is defined for all $(\phi, \phi(0)) \in L_2 \times \mathbb{R}^n$. Also on this extended set of arguments, the value of interest from (4.61) is still the infimum $k_1^{\mathrm{opt}} = \inf_{\substack{\phi \in C_{\mathrm{d}} \\ \phi(0) \neq 0_n}} V\big(\frac{1}{\|\phi(0)\|_2}\phi\big)$ (even on $L_2 \times \mathbb{R}^n$ it would be since V_{M_2} is continuous[14] in $M_2 = L_2 \times \mathbb{R}^n$ and C is dense in L_2). With a slight abuse of notation the name V is not altered for the extension on C_{d}. By construction, the discretization $\pi_y^{[N]}(\phi_{\mathrm{d}}^{[N]}) = \left[\begin{smallmatrix} z^{[N]} \\ \hat{x}^{[N]} \end{smallmatrix}\right]$ yields an argument of the minimum in (4.61). First, it will be shown that $\forall \varepsilon > 0, \exists \bar{N}_1(\varepsilon) \in \mathbb{N}$, such that

$$\forall N \geq \bar{N}_1(\varepsilon) : \quad |\underbrace{V_y^{[N]}(\pi_y^{[N]}(\phi_{\mathrm{d}}^{[N]}))}_{k_1^{[N]}} - V(\phi_{\mathrm{d}}^{[N]})| < \varepsilon. \tag{4.65}$$

According to the splitting approach (Lemma 4.6.1 with $\tilde{Q}_1 = Q_1, \tilde{Q}_2 = Q_2$), V is decomposed into three parts $V(\phi_{\mathrm{d}}^{[N]}) = V_0(\phi_{\mathrm{d}}^{[N]}) + V_1(\phi_{\mathrm{d}}^{[N]}) + V_2(\phi_{\mathrm{d}}^{[N]})$

[15] The L_2 norm of (3.48) can be computed from $\|\phi_{\mathrm{d}}^{[N]}\|_{L_2}^2 = \sum_{k=0}^{N-1} \frac{h}{2} \frac{2}{2k+1} \|c^k\|_2^2$ using the first $N-1$ of the N subvectors in c. These are either derived via $c = T_{cy}\left[\begin{smallmatrix} z \\ \hat{x} \end{smallmatrix}\right]$, cf. Remark 3.6.1, where $z = -P_{y,zz}^{-1} P_{y,xz}^{\top} \hat{x}$ and $\hat{x} = v/\|v\|_2$, see Lemma 4.3.1, or are directly available if Lemma 4.3.1 is applied to the mixed coordinates from (3.64).

and its approximation correspondingly. The second and third term, $V_1(\phi_{\mathrm{d}}^{[N]})$ and $V_2(\phi_{\mathrm{d}}^{[N]})$, do not contribute to the error in (4.65) since $\phi_{\mathrm{d}}^{[N]}(\theta)$ is on $\theta \in [-h, 0)$ a polynomial of degree $N-1$ for which the approximation is exact, according to the lemmas of Section 4.6.1.2. Therefore, it suffices to show uniform convergence on $\cup_N \{\phi_{\mathrm{d}}^{[N]}\}$ for the approximations of V_0. Let $\mathsf{x}^{[N]} = (\phi_{\mathrm{d}}^{[N]}, \phi_{\mathrm{d}}^{[N]}(0)) \in M_2$. By assumption, $\|\mathsf{x}^{[N]}\|_{M_2}^2 = \|\phi_{\mathrm{d}}^{[N]}\|_{L_2}^2 + \|\phi_{\mathrm{d}}^{[N]}(0)\|_2^2 \leq \beta^2 + 1$. Thus, using (4.46) and (4.56), the error in (4.65) becomes $|\langle \mathsf{x}^{[N]}, \mathscr{P}_0^{[N]} \mathsf{x}^{[N]} \rangle_{M_2} - \langle \mathsf{x}^{[N]}, \mathscr{P}_0 \mathsf{x}^{[N]} \rangle_{M_2}| \leq \|\mathscr{P}_0^{[N]} - \mathscr{P}_0\| \, (\beta^2 + 1)$. By Lemma 4.6.7, the latter converges to zero, and thus (4.65) holds. Consequently, $\forall N \geq \bar{N}_1(\varepsilon)$:

$$k_1^{[N]} \overset{(4.65)}{>} V(\phi_{\mathrm{d}}^{[N]}) - \varepsilon \geq \inf_{\substack{\phi \in C_{\mathrm{d}} \\ \phi(0) \neq 0_n}} V\big(\tfrac{1}{\|\phi(0)\|_2}\phi\big) - \varepsilon \overset{(4.61)}{=} k_1^{\mathrm{opt}} - \varepsilon. \tag{4.66}$$

With $\bar{N}_0(\varepsilon) := \bar{N}(\frac{\varepsilon}{2}, \phi_{\delta=\varepsilon/2})$ from Theorem 4.6.9, it can be concluded that

$$\forall N \geq \max\{\bar{N}_0(\varepsilon), \bar{N}_1(\varepsilon)\}: \quad k_1^{\mathrm{opt}} - \varepsilon \overset{(4.66)}{<} k_1^{[N]} \overset{(4.64)}{<} k_1^{\mathrm{opt}} + \varepsilon,$$

completing the proof of $|k_1^{[N]} - k_1^{\mathrm{opt}}| \to 0$ $(N \to \infty)$. □

4.7 Revisiting the Main Points of the Chapter

- The proposed numerical approach to complete-type and related LK functionals is based on the ODE approximation of time-delay systems discussed in Chapter 3.
  - The approach neither requires knowledge of the delay-Lyapunov matrix function, nor does it involve linear matrix inequalities.
  - Simply the Lyapunov equation (4.3) relying on the system matrix A_y of the ODE approximation is solved for the matrix P_y. (Equivalently to the indicated formulation in Lagrange interpolation coordinates,

other coordinates can be used, e.g., mixed coordinates or Legendre coordinates, see (4.6).)

  - The result $V_y(y) = y^\top P_y y$ approximates the LK functional value $V(\phi)$, given y represents the discretization of ϕ.
  - For the Legendre-tau-based result with an increasing discretization resolution N, convergence towards $V(\phi)$ is proven in Section 4.6.2.

- The combs in Figure 4.1, which shows the entries of P_y, indicate the structure (4.11) of complete-type LK functionals without any prior knowledge.
- In the Lyapunov–Krasovskii theorem (Theorem 4.5.1) for time-delay systems, the pointwise norm of the solution value $\|x(t)\|$ occurs, where, in view of the classical Lyapunov theorem for ODEs, actually the norm of the state $\|x_t\|_C$ should be expected. Due to that special role of $\|x(t)\|$ (see the black point in Figure 3.1), the counterpart to the concept of LK functionals for RFDEs is not the concept of classical Lyapunov functions for ODEs.
  - Instead, the Lyapunov–Krasovskii theorem is recognized to parallel Theorem 4.5.4, which only proves partial asymptotic stability in the ODE (Section 4.5.2).
  - Partial asymptotic stability in the approximating ODE already implies asymptotic stability (Section 4.5.3). The latter reflects that, in the RFDE, convergence of $x(t)$ as $t \to \infty$ implies convergence of the adhering solution segment x_t.
- For an appropriate ODE approximation with a sufficiently large discretization resolution N, the involved matrix P_y is positive semidefinite if and only if the RFDE equilibrium is asymptotically stable (Corollary 4.5.14).
- A formula for a partial positive-definiteness bound on the functional approximation is derived (Section 4.3).

  - With the Legendre-tau ODE-based result, a rapid convergence of the obtained quadratic-lower-bound coefficient is observed as N increases (Figure 4.2b).
  - The resulting limit is proven to be the best possible quadratic-lower-bound coefficient k_1 on the LK functional (Section 4.6.3), cf. (1.16) in the introduction.
  - Examples demonstrate that the latter significantly improves known results (Table 4.1). In particular, the obtained k_1 depends on the delay, which is not the case in existing formulas (Figure 4.2e).

- For the sake of validation, a numerical integration of the LK functional formula by Clenshaw–Curtis and Gauss quadrature rules is also proposed. The result can as well be written as a quadratic form $y^\top P_y^{\mathrm{quad}} y$.
  - In contrast to the ODE-based approach, the numerical-integration-based approach requires knowledge of the delay Lyapunov matrix function Ψ.
  - In Appendix B.1.2, the relation between P_y^{quad} and a known stability criterion in terms of Ψ is shown (Remark 4.4.1).

- The ODE-based approach is expected to provide approximations of LK functionals even in more general cases where the LK functional is not known analytically and thus the numerical-integration-based approach cannot be applied. In fact, the ODE-based approach paves the way to a more adapted construction of the functional in the next chapter.

5 Introduction of Lyapunov–Krasovskii Functionals of Robust Type

Inspired by the widespread theory of complete-type LK functionals, this chapter introduces an alternative class of LK functionals that intends to achieve less conservative robustness statements. As outlined in Section 1.4, the robustness analysis is an important application of complete-type and related LK functionals. Still, the achievable robustness bounds by complete-type and related LK functionals turn out to be rather conservative. The proposed LK functionals in the present chapter share the same structure as the LK functionals of complete type, and also they share to be defined via their derivative along solutions of the nominal system. The defining equation for the derivative, however, is chosen differently: the Lyapunov equation, which forms the template for the defining equation of complete-type LK functionals, is replaced by the template of an algebraic Riccati equation.

The chapter is organized as follows. **Section 5.1** states some leading questions, motivated by which **Section 5.2** introduces the concept of LK functionals of robust type. Afterwards, important properties of these LK functionals are discussed: **Section 5.3** is devoted to the monotonicity properties along solutions of the perturbed equation, and **Section 5.4** tackles the partial positive-definiteness property. In **Section 5.5**, existence conditions are derived from the infinite-dimensional Kalman–Yakubovich–Popov lemma combined with a splitting approach. These

considerations finally lead to the desired robustness bounds. An example is discussed in **Section 5.6**, before **Section 5.7** summarizes the most important points of the overall chapter.

The chapter is part of

[S3] Scholl, T. H.: Lyapunov–Krasovskii functionals of robust type for the stability analysis in time-delay systems. Submitted for publication, arXiv preprint available (2023). arXiv:2312.16738.

5.1 Leading Questions*

The robustness approach via complete-type LK functionals outlined in Section 1.4 relies on the rather arbitrary choice of three positive definite matrices Q_0, Q_1, Q_2 in (1.14), i.e., in $D_f^+ V(x_t) = -x^\top(t)Q_0x(t) - x^\top(t-h)Q_1x(t-h) - \int_{-h}^{0} x^\top(t+\theta)Q_2x(t+\theta)\,\mathrm{d}\theta$. The following questions motivate a different construction of the LK functional.

1. Why should the desirable LK-functional derivative be restricted to the structure in (1.14). Why not choosing for instance

$$\begin{aligned} D_f^+ V(x_t) = &-x^\top(t)Q_{\mathrm{xx}}x(t) - x^\top(t-h)Q_{\bar{\mathrm{x}}\bar{\mathrm{x}}}x(t-h) \\ &- 2x^\top(t-h)Q_{\bar{\mathrm{x}}\mathrm{x}}x(t) - 2\int_{-h}^{0} x^\top(t)Q_{\mathrm{xz}}(\eta)x(t+\eta)\,\mathrm{d}\eta \\ &- 2\int_{-h}^{0} x^\top(t-h)Q_{\bar{\mathrm{x}}\mathrm{z}}(\eta)x(t+\eta)\,\mathrm{d}\eta \\ &- \int_{-h}^{0}\int_{-h}^{0} x^\top(t+\xi)Q_{\mathrm{zz}}(\xi,\eta)x(t+\eta)\,\mathrm{d}\eta\mathrm{d}\xi \end{aligned} \tag{5.1}$$

   in the place of (1.14)? For that choice, according to [115, Thm. 5.1] (regarding all terms but $Q_{\bar{\mathrm{x}}\bar{\mathrm{x}}}$) and [111] (regarding $Q_{\bar{\mathrm{x}}\bar{\mathrm{x}}}$), as well a unique

* Chapter 5 is part of [S3].

functional V exists, which has the same form as complete-type LK functionals (see (5.5) below). Unfortunately, it is already an open question how to choose Q_0, Q_1, Q_2 in (1.14) in an optimal manner—with the above more general structure (5.1), we are even more spoiled for choice on how to specify the kernel functions.

2. Can the construction of $D_f^+V(x_t)$ be aligned to the outcome of the derivative $D_{(f+g)}^+V(x_t)$ along solutions of the perturbed equation? In (1.14), the derivative $D_f^+V(x_t)$ of complete-type LK functionals along solutions of the unperturbed system is, by construction, chosen partially negative definite (in the sense required for the LK theorem, see (4.29)). However, that unperturbed linear system $\dot{x}(t) = A_0x(t) + A_1x(t-h) = f(x_t)$ from (1.4) is actually not what shall be analyzed, see Chapter 1. Rather, the perturbed system (1.2) and thus $D_{(f+g)}^+V(x_t)$ is of interest. The restrictive linear norm bound (1.17) from complete-type LK functionals ensures that the perturbation cannot turn the partially negative definite chosen $D_f^+V(x_t)$ into a $D_{(f+g)}^+V(x_t)$ that is no longer partially negative definite. Thus, it ensures that the LK theorem still applies despite of the perturbation g. In the end, various inequality estimations of $D_{(f+g)}^+V(x_t)$ are employed to obtain that linear norm bound $\gamma_{\max}$ on g, which is why (1.17) is very restrictive. This might be improved by focusing on the outcome of $D_{(f+g)}^+V$ when constructing D_f^+V.

3. Can some information on the structure of the perturbation $g(x_t)$ in (1.2) be incorporated in the construction of V? It might be highly relevant for the achievable linear norm bound γ in (1.17) whether the perturbation affects only certain components of $\dot{x}(t)$, or whether it depends only on certain parts of $x(t)$ and $x(t-h)$.

Example 5.1.1 (Delay-free motivation, structure). *Consider the delay-free example*

$$\dot{x} = \begin{bmatrix} -0.1 & 0 \\ 0 & -10 \end{bmatrix} x + g(x). \tag{5.2}$$

Compared to $g(x) = \left[\begin{smallmatrix} g_1(x_1) \\ 0 \end{smallmatrix}\right]$*, a perturbation* $g(x) = \left[\begin{smallmatrix} 0 \\ g_2(x_2) \end{smallmatrix}\right]$ *is far less critical. Therefore, if this structure information is included, less restrictive bounds on the perturbation should become possible. The decisive structure information can be made visible by a Lur'e description in terms of a fictive feedback law* $g(x) = -B\,a(Cx) = -\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right] a(\left[\begin{smallmatrix} 0 & 1 \end{smallmatrix}\right] x)$ *with* $a(\zeta) = -g_2(\zeta)$*. The latter is a well-known means, see, e.g., concepts from the realm of absolute stability [108], the structured stability radius [89], or other robustness results. Once the structure describing matrices* $B = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ *and* $C = \left[\begin{smallmatrix} 0 & 1 \end{smallmatrix}\right]$ *are fixed, only the specific restriction on* $a(\zeta)$ *is of interest.*

4. Why striving exclusively for a linear norm bound?

Example 5.1.2 (Delay-free motivation, asymmetric bound). *Consider the delay-free scalar ODE*

$$\dot{x} = -x + g(x) \quad \textit{with} \quad g(x) = -x^3. \tag{5.3}$$

Actually, $g(x) = -x^3$ *is even a helpful perturbation of* $\dot{x} = -x$*, not at all hampering the global asymptotic stability of the origin. However, a linear norm bound* $|g(x)| \leq \gamma|x|$ *cannot distinguish between* $g(x) = -x^3$ *and the globally harmful* $g(x) = x^3$*. Hence, the approach from Section 1.2 iii, which relies on a set* G *where the linear norm bound is satisfied, can only provide a small estimation of the domain of attraction. However, from the*

one-dimensional phase portrait, it is obvious that stability is preserved for any

$$g(x) = -a(x) \quad \textit{where} \quad \begin{cases} a(x) \geq 0, & x > 0, \\ a(x) \leq 0, & x < 0, \end{cases} \tag{5.4}$$

which, in fact, is satisfied globally by $a(x) = x^3$.

5.2 LK Functionals of Robust Type*

Before being able to introduce the proposed LK functionals of robust type in Definition 5.2.4, some ingredients of the defining equation have to be discussed. Similarly to complete-type LK functionals in (1.14), LK functionals of robust type will be defined in terms of their derivative $D_f^+ V$. The resulting functionals $V\colon C([-h,0],\mathbb{R}^n) \to \mathbb{R}; \phi \mapsto V(\phi)$ will be quadratic, time-invariant, and have the form

$$\begin{aligned} V(\phi) = {} & \phi^\top(0)\, P_{\mathrm{xx}}\, \phi(0) + 2 \int_{-h}^{0} \phi^\top(0)\, P_{\mathrm{xz}}(\eta) \phi(\eta)\, \mathrm{d}\eta \\ & + \int_{-h}^{0} \int_{-h}^{0} \phi^\top(\xi) P_{\mathrm{zz}}(\xi,\eta) \phi(\eta)\, \mathrm{d}\eta\, \mathrm{d}\xi \\ & + \int_{-h}^{0} \phi^\top(\eta) P_{\mathrm{zz,diag}}(\eta) \phi(\eta)\, \mathrm{d}\eta, \end{aligned} \tag{5.5}$$

$P_{\mathrm{xx}} \in \mathbb{R}^{n\times n}$, $P_{\mathrm{xz}} \in L_2([-h,0],\mathbb{R}^{n\times n})$, $P_{\mathrm{zz}} \in L_2([-h,0]\times[-h,0],\mathbb{R}^{n\times n})$ and $P_{\mathrm{zz,diag}}(\eta) \equiv \overline{P}_{\mathrm{zz,diag}} \in \mathbb{R}^{n\times n}$. Thus, they have the same form as the LK functionals of complete type, cf. (4.11)—respectively even a simpler one since $P_{\mathrm{zz,diag}}(\eta)$ is constant, which in (1.15) is only the case if $Q_2 = 0_{n\times n}$.

* Chapter 5 is part of [S3].

The defining equation for the functional, which will be given in (5.17) below, takes the leading questions from Section 5.1 into account in the following respects.

1. **More general structure of the nominal derivative:** The proposed LK-functional derivative D_f^+V, in fact, will have the general structure (5.1). However, the matrices and kernel functions need not explicitly be chosen. In particular, they are not arbitrarily chosen. Instead, the overall construction of the desired derivative D_f^+V relies on a characterization of the class of perturbations g for which a robustness statement shall be achieved. Additionally, however, the defining equation for D_f^+V, also depends on the solution V itself. To be more precise, based on the first two terms in the LK functional (5.5), the expression

$$v(\phi) := P_{\mathrm{xx}}\,\phi(0) + \int_{-h}^{0} P_{\mathrm{xz}}(\eta)\phi(\eta)\,\mathrm{d}\eta \tag{5.6}$$

   will be encountered in the defining equation (5.17). Due to that implicit definition of D_f^+V, the problem might seem rather involved. However, in terms of the numerical approach from Chapter 4, the Lyapunov equation required for complete-type LK functionals is simply replaced by an algebraic Riccati equation, which will be discussed in Section 6.1.

2. **Aligning the Construction of D_f^+V to the outcome of $D_{(f+g)}^+V$:** The defining equation (5.17) proposed below is only constructed having the outcome of $D_{(f+g)}^+V$ in mind. Therefore, the sense of the chosen D_f^+V will become clear from the result of $D_{(f+g)}^+V$ in Section 5.3.

3. **Perturbation structure:** To take the structure of the perturbation g into account, the perturbation is decomposed into three mappings, namely

$$g(x_t) = -B\,a(\mathcal{C}x_t). \tag{5.7}$$

That is, first, a linear operator $\mathcal{C}\colon C([-h,0],\mathbb{R}^n) \to \mathbb{R}^p$, to confine what the perturbation is based upon,

$$\mathcal{C}\phi = \begin{bmatrix} C_1\phi(-h) \\ C_0\,\phi(0) \end{bmatrix}, \quad \text{i.e., } \mathcal{C}x_t = \begin{bmatrix} C_1\,x(t-h) \\ C_0\,x(t) \end{bmatrix} \tag{5.8}$$

with $C_1 \in \mathbb{R}^{p_1\times n}$, $C_0 \in \mathbb{R}^{p_0\times n}$, $p_0 + p_1 = p$ (where C_0 or C_1 vanish if $p_0 = 0$ or $p_1 = 0$); second, a possibly nonlinear continuous map

$$a\colon \mathbb{R}^p \to \mathbb{R}^m \tag{5.9}$$

with $a(0_p) = 0_m$; and, third, a matrix

$$B \in \mathbb{R}^{n\times m} \tag{5.10}$$

that indicates which components of $\dot{x}(t)$ in (1.2) are affected by the perturbation. The negative sign in (5.7) intends to resemble a negative feedback.

Example 5.2.1. *Let $g(x_t)$ in the time-delay system (1.2) be given by*

$$g(x_t) = \begin{bmatrix} 0 \\ -x_1^3(t-h) \end{bmatrix}.$$

The latter can be expressed via $B = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$, $C_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}$, and $p_0 = 0$, with the core nonlinearity $a(\zeta) = \zeta^3$ in (5.7). An alternative choice is $B = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$, full-rank matrices $C_1 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & \varepsilon \end{smallmatrix}\right]$, $C_0 = \varepsilon I_2$, $\varepsilon \in \mathbb{R}$, and $a([\zeta_1,\zeta_2,\zeta_3,\zeta_4]^\top) = \zeta_1^3$. Such full-rank matrices for C_0 and/or C_1 might be helpful to establish some desired properties in the resulting LK functional (e.g., certain lower bounds on $V(\phi)$ or upper bounds on $D^+_{(f+g)}V(\phi)$).

As a consequence, the perturbation restriction must only refer to the mapping a in (5.7). Thus, $\|\tilde{g}(x(t), x(t-h))\|_2 \leq \gamma \left\| \left[\begin{smallmatrix} x(t) \\ x(t-h) \end{smallmatrix} \right] \right\|_2$ from (1.17) is replaced by

$$\|a(\mathcal{C}x_t)\|_2 \leq \gamma \|\mathcal{C}x_t\|_2. \tag{5.11}$$

If, however, identity matrices

$$B = C_0 = C_1 = I_n \tag{5.12}$$

are chosen, which is always possible, then, by

$$\tilde{g}(x(t), x(t-h)) \overset{(5.7)}{=} -Ba\left(\left[\begin{smallmatrix} C_1 x(t-h) \\ C_0 x(t) \end{smallmatrix} \right] \right) \overset{(5.12)}{=} -a\left(\left[\begin{smallmatrix} x(t-h) \\ x(t) \end{smallmatrix} \right] \right),$$

a statement as in (1.17) is recovered, which does not incorporate structural information.

4. **Perturbation restriction:** There might be more appropriate types of restrictions than (5.11) on the possibly nonlinear map $\zeta \mapsto a(\zeta)$ in (5.7). Note that the linear norm bound (5.11) can equivalently be written as $a^\top(\mathcal{C}x_t)\, a(\mathcal{C}x_t) \leq \gamma^2\, (\mathcal{C}x_t)^\top (\mathcal{C}x_t)$ or

$$w\big(\mathcal{C}x_t,\, a(\mathcal{C}x_t)\big) \geq 0 \quad \text{with} \tag{5.13a}$$
$$w(\zeta, \alpha) = \gamma^2 \zeta^\top \zeta - \alpha^\top \alpha. \tag{5.13b}$$

To establish other types of perturbation restrictions, indefinite quadratic forms in $[\zeta^\top, \alpha^\top]^\top$ that are more general than (5.13b) are taken into account. That is,

$$w\big(\mathcal{C}x_t,\, a(\mathcal{C}x_t)\big) \geq 0 \quad \text{with} \tag{5.14a}$$
$$w(\zeta, \alpha) = \zeta^\top \Pi_{\zeta\zeta} \zeta + 2\zeta^\top \Pi_{\zeta a} \alpha + \alpha^\top \Pi_{aa} \alpha, \tag{5.14b}$$

where the matrices $\Pi_{\zeta\zeta} = \Pi_{\zeta\zeta}^\top \in \mathbb{R}^{p\times p}$, $\Pi_{\zeta a} \in \mathbb{R}^{p\times m}$, $\Pi_{aa} = \Pi_{aa}^\top \in \mathbb{R}^{m\times m}$ should depend on a parameter, in terms of which the robustness

statement can be expressed: for instance, the linear norm bound γ parameterizes $\Pi_{\zeta\zeta} = \gamma^2 I_p$ in (5.13b). The third matrix is henceforth assumed to be negative definite,

$$\Pi_{aa} \prec 0_{m\times m}. \tag{5.15}$$

As is customary with a linear norm bound, the perturbation restriction

$$w(\zeta, a(\zeta)) \geq 0 \tag{5.16}$$

from (5.14) should either locally, i.e., for sufficiently small $\zeta \in \mathbb{R}^p$, or even globally be satisfied by the function $\zeta \mapsto a(\zeta)$ from the perturbation $g(x_t) = -Ba(\mathcal{C}x_t)$. Table 5.1 provides an overview of possible choices for $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$ in (5.14) and the associated permitted sector for $a(\zeta)$.

Example 5.2.2. *For simplicity, consider $p = m = 1$, i.e. $\zeta \mapsto \alpha = a(\zeta)$ is a scalar map. Then $w(\zeta, \alpha) \geq 0$ defined in (5.13b) describes a sector in the (ζ, α) plane. That is, $w(\zeta, \alpha) \geq 0$ holds for the (ζ, α)-combinations in the non-gray-region of the (ζ, α)-plots from the last column of Table 5.1. A possible type of perturbation restriction $w(\zeta, a(\zeta)) \geq 0$ is a linear norm bound, see Table 5.1, row (I\|a). It is represented by $\Pi_{\zeta\zeta} = \gamma^2, \Pi_{\zeta a} = 0, \Pi_{aa} = -1$ since $w(\zeta, \alpha) = \gamma^2\zeta^2 - \alpha^2 \geq 0$ is equivalent to the desired $|\alpha| \leq \gamma|\zeta|$. The robustness statement then provides the maximum slope γ similarly to the statement known from complete-type LK functionals in (1.17). However, for a nonlinearity like $a(\zeta) = \zeta^3$, the sector in Table 5.1, row (II\|a), fits much better. The robustness statement then describes the admissible upper slope $\frac{1}{\rho}$ which might be considerably larger than γ. As a result, the range of ζ, for which $a(\zeta) = \zeta^3$ resides within the sector, is also accordingly larger. For a saturation nonlinearity, rather Table 5.1, row (III), with a fixed upper slope K_2 should be chosen, and the smallest admissible lower slope K_1 be determined.*

Remark 5.2.3 (Negative definiteness of Π_{aa}). *Assumption (5.15) forbids to use $\rho = 0$ in row (II\|a) of Table 5.1, which would amount to a pure passivity*

restriction $w(\zeta, a(\zeta)) = \zeta a(\zeta) \geq 0$ (cf. Example 5.1.2). Nevertheless, as $\rho > 0$ can be chosen arbitrarily small, still an arbitrarily large domain of ζ can be obtained on which the perturbation $a(\zeta) = \zeta^3$ resides in the sector that is sketched in the right column of row (II|a).

Definition 5.2.4 (LK functional of robust type).
A functional $V\colon C([-h,0],\mathbb{R}^n) \to \mathbb{R}_{\geq 0}$ that has the structure (5.5) is called a Lyapunov–Krasovskii functional of robust type w.r.t.

- *the nominal linear system $\dot{x}(t) = f(x_t)$,*
- *the perturbation structure $(B, \mathcal{C})$, and*
- *the perturbation restriction matrices $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$*

if for all $\phi \in C([-h,0],\mathbb{R}^n)$ it holds

$$\begin{aligned} D_f^+ V(\phi) = &-(\mathcal{C}\phi)^\top \Pi_{\zeta\zeta}(\mathcal{C}\phi) \\ &- \Big[v^\top(\phi)B - (\mathcal{C}\phi)^\top \Pi_{\zeta a}\Big](-\Pi_{aa})^{-1}\Big[B^\top v(\phi) - \Pi_{\zeta a}^\top \mathcal{C}\phi\Big] - e(\phi), \end{aligned} \tag{5.17}$$

with $v\colon C([-h,0],\mathbb{R}^n) \to \mathbb{R}^n$ given by (5.6) and $e(\phi) \equiv 0$. Moreover, if $e(\phi) \geq 0$ is some arbitrary nonnegative discrepancy between the left- and the remaining right-hand side in (5.17), V is called an inequality-based Lyapunov–Krasovskii functional of robust type.

In the case of a linear norm bound, i.e., $\Pi_{\zeta\zeta} = \gamma^2 I_p,\ \Pi_{\zeta a} = 0_{p\times m},\ \Pi_{aa} = -I_m$ (see row (I|a) of Table 5.1), the defining equation (5.17) simplifies to

$$D_f^+ V(\phi) = -\gamma^2(\mathcal{C}\phi)^\top \mathcal{C}\phi - v^\top(\phi)BB^\top v(\phi). \tag{5.18}$$

Because of (5.8), $-(\mathcal{C}\phi)^\top\mathcal{C}\phi = -\phi^\top(0)C_0^\top C_0\phi(0) - \phi^\top(-h)C_1^\top C_1\phi(-h)$. Note that these terms resemble the first two terms in $D_f^+V(\phi)$ from the defining equation (1.14) for complete-type LK functionals, which are given by $-\phi^\top(0)Q_0\phi(0) - \phi^\top(-h)Q_1\phi(-h)$.

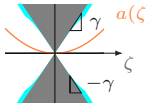
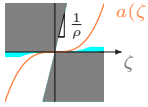
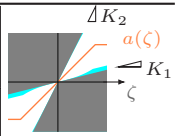
Perturbation restriction $\boldsymbol{w(\zeta, a(\zeta)) \geq 0}$, $w(\zeta,\alpha) = \zeta^\top \Pi_{\zeta\zeta}\zeta + 2\zeta^\top \Pi_{\zeta a}\alpha + \alpha^\top \Pi_{aa}\alpha$	Sector notation	$\Pi_{\zeta\zeta}$	$2\Pi_{\zeta a}$	Π_{aa}	Restriction to the graph of a if $p = m = 1$
(I|a) $\|a(\zeta)\|_2 \leq \gamma\|\zeta\|_2$ ($\Leftrightarrow \quad \gamma^2\zeta^\top\zeta - a^\top(\zeta)\, a(\zeta) \geq 0 \quad$) **linear norm bound** on $a(\zeta)$, restriction of the **small-gain theorem**	$[-\gamma, \gamma]$	$\gamma^2 I_p$	$0_{p\times m}$	$-I_m$	
(I|b) $\|Wa(\zeta)\|_2 \leq \gamma\|L\zeta\|_2$, (I|a) in elliptic norms		$\gamma^2 L^\top L$	$0_{p\times m}$	$-W^\top W$	
(II|a) $a^\top(\zeta)\,\zeta \geq \rho\|a(\zeta)\|_2^2$, ($\Leftrightarrow \quad a^\top(\zeta)\,[\zeta - \rho a(\zeta)] \geq 0 \quad$), **strict output passivity** of $\zeta \mapsto a(\zeta)$, **passivity-theorem-like** restriction	$[0, \frac{1}{\rho}]$	$0_{m\times m}$	I_m	$-\rho I_m$	
(II|b) $-a^\top(\zeta)\,\zeta \geq \hat{\rho}\|a(\zeta)\|_2^2$, ($\Leftrightarrow \quad -a^\top(\zeta)\,[\zeta + \hat{\rho}a(\zeta)] \geq 0 \quad$), strict output passivity of $\zeta \mapsto -a(\zeta)$,	$[-\frac{1}{\hat{\rho}}, 0]$	$0_{m\times m}$	$-I_m$	$-\hat{\rho} I_m$	
	alternatively, use (II|a) with $\hat{a}(\zeta) = -a(\zeta)$ in $g(x_t) = \hat{B}\hat{a}(\mathcal{C}x_t)$, $\hat{B} = -B$				
(III|a) $-[a(\zeta) - K_1\zeta]^\top[a(\zeta) - K_2\zeta] \geq 0$ **general sector bound**, restriction of the **circle criterion**	$[K_1, K_2]$	$-\mathrm{sym}(K_1^\top K_2)$	$K_1^\top + K_2^\top$	$-I_m$	
(III|b) $[a(\zeta) - K_1\zeta]^\top[\zeta - K_2^{-1}a(\zeta)] \geq 0, \quad K_2 \succ 0$ passivity $\hat{a}^\top\hat{\zeta} \geq 0$ after loop transformations $\hat{\zeta} = \zeta - K_2^{-1}a(\zeta); \quad \hat{a} = a(\zeta) - K_1\zeta$		$-K_1^\top$	$I_m + K_1^\top K_2^{-1}$	$-K_2^{-1}$	
(III|c) $\mp[K_1^{-1}a(\zeta) - \zeta]^\top[\zeta - K_2^{-1}a(\zeta)] \geq 0$ assuming $\mathrm{sym}(\mp K_1^{-\top}K_2^{-1}) \succ 0$		$\pm I_m$	$\mp K_1^{-1} \mp K_2^{-1}$	$-\mathrm{sym}(\mp K_1^{-\top}K_2^{-1})$	

Table 5.1: Since the perturbation structure $(B, \mathcal{C})$ in $g(x_t) = -Ba(\mathcal{C}x_t)$ from (5.7) is fixed, the perturbation restriction only refers to the (usually nonlinear) map $a(\cdot)$. This perturbation restriction is fully described via the matrices $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$ given in the third column. In row (II) and (III), it is assumed that $p = m$. As indicated in the last column, in the scalar case, the perturbation restriction is simply representable via a sector, the graph of $a(\cdot)$ is allowed to reside within. The turquoise shading amounts to a tightening from Theorem 5.3.2 (choosing a non-quadratic offset function ℓ results in a deviation from the sector form). The sector slopes up to which an LK functional of robust type exists, only depend on the robustness of the nominal system under the given perturbation structure. The corresponding bounds on γ, $\frac{1}{\rho}$, and k_1 (if $K_1 = k_1 I_m$) are listed in Table 5.2.

The nominal linear system $\dot{x}(t) = f(x_t)$ determines the left-hand side of (5.17), respectively (5.18). The existence of a Lyapunov–Krasovskii functional of robust type (or of its generalization with $e \geq 0$) is only ensured if the chosen perturbation restriction $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$ fits with the robustness of this nominal system under the chosen perturbation structure. If a parameter is incorporated in $\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa}$ that controls the size of the sector, like the slope γ of the linear norm bound, then the existence condition of V can be expressed in terms of that parameter. Consequently, permissible values of the involved parameter are characterized by the fact that a solution V of (5.17) exists. Corresponding solvability conditions will be derived from the Kalman–Yakubovich–Popov lemma in Section 5.5.3. This leads, e.g., to an explicit bound $\gamma_{\max}$ on γ in (5.18) that will be derived in Section 5.5.5. If $\gamma < \gamma_{\max}$ a solution of (5.18) exists, whereas if $\gamma > \gamma_{\max}$ no solution exists.

First, however, properties of the functional that are decisive for its usability will be discussed: monotonicity along solutions (Section 5.3) and partial positive definiteness (Section 5.4). The results will show that, if C_0 and C_1 in (5.8) are chosen as full-rank matrices, then (provided a fundamental requirement described in Section 5.4 holds) the conditions imposed by the classical LK theorem (Theorem 4.5.1) hold in any case. Although not further detailed in this thesis, weaker properties than those in the classical LK theorem can also be expedient to prove stability, e.g., relying on LaSalle's invariance principle [124] or on other methods [136]. Once it is ensured that the resulting functional indeed satisfies suitable properties that prove stability of the zero equilibrium in the overall system, the searched robustness statement boils down to the found existence condition of the functional.

Remark 5.2.5 (Existing results from the realm of absolute stability). *In terms of the class of perturbations that satisfy the perturbation restriction not only locally but globally, the proposed approach addresses the problem of absolute stability [5], which means to tackle a whole family of sector-bounded nonlinearities at once. Modern robustness theory might be dominated by frequency-domain methods. However, its foundation can be found in the theory of absolute stability,*

which relied on a parallel development of frequency-domain (Lyapunov-function-free) and Lyapunov-function-based approaches, as well as the proof of their equivalence by the Kalman–Yakubovich–Popov lemma. See the extensive literature reviews in [78, 127]. The frequency-domain (Lyapunov-function-free) approaches have already in the very beginning been extended to time-delay systems [158, 80, 107, 186, 17, 20]. Still, in view of possibly non-global stability results for nonlinear perturbations, an LK-functional-based approach can be seen to be preferable—provided the functional is explicitly computable. However, concerning Lyapunov-function-based considerations, there is no satisfactory counterpart for time-delay systems in terms of a computable LK functional without adding conservatism. The required computability is not satisfied by generalizations to abstract differential equations on Hilbert spaces [129, 131, 130, 7] that merely provide existence statements. Computable LK functionals so far are only encountered in semidefinite-programming-based approaches for absolute stability in time-delay systems [19, 22, 162, 83, 84, 125, 174], where, however, the inherent limited number of degrees of freedom comes along with an additional conservatism—see the discussion on LMI-based stability criteria in Section 1.3.

To sum up, the construction of the LK functional of robust type incorporates a characterization of the perturbation in two respects: First, the concept allows to incorporate the perturbation structure. If, however, $B = C_0 = C_1 = I_n$ is chosen, then an unstructured robustness consideration as the one known from complete-type LK functionals is recovered. Second, the concept allows to incorporate the type of perturbation restriction that fits best with the nonlinearity. If, however, $\Pi_{\zeta\zeta} = \gamma^2 I_p, \quad \Pi_{\zeta a} = 0_{p\times m}, \quad \Pi_{aa} = -I_m$ is chosen, then a linear norm bound as known from complete-type LK functionals is recovered. In the latter case, the defining equation simplifies to (5.18). The defining equation only has a solution if the perturbation restriction is compatible with the given nominal system under the given perturbation structure. For instance, (5.18) is solvable if $\gamma < \gamma_{\max}$, with $\gamma_{\max}$ to be discussed in Section 5.5.5.

5.3 The LK-Functional Derivative Along Solutions of the Perturbed RFDE*

Along solutions of the unperturbed RFDE, the LK-functional derivative $D_f^+ V(x_t)$ is given by the right-hand side of the defining equation (5.17). Hence, $D_f^+ V(x_t)$ is exactly known once V and thus v have been determined. However, rather of interest is $D_{(f+g)}^+ V(x_t)$. The following lemma is valid for any functional having the structure (5.5), independently from the defining equation. When applied to the special case of complete-type functionals, it leads to a result known from [110, Lem. 2.14].

Lemma 5.3.1 (Perturbation effect on the derivative). *For a functional given by (5.5), it holds*

$$D_{(f+g)}^+ V(\phi) = D_f^+ V(\phi) + 2\, v^\top(\phi)\, g(\phi) \tag{5.19}$$

with v being defined in (5.6).

Proof. For $\phi = x_t$, the LK functional (5.5) becomes

$$\begin{aligned} V(x_t) = {} & x^\top(t) P_{\mathrm{xx}}\, x(t) + 2x^\top(t) \int_{t-h}^{t} P_{\mathrm{xz}}(\eta - t) x(\eta)\, \mathrm{d}\eta \\ & + \int_{t-h}^{t} \int_{t-h}^{t} x^\top(\xi) P_{\mathrm{zz}}(\xi - t, \eta - t) x(\eta)\, \mathrm{d}\eta\, \mathrm{d}\xi \\ & + \int_{t-h}^{t} x^\top(\eta) P_{\mathrm{zz,diag}}(\eta - t) x(\eta)\, \mathrm{d}\eta. \end{aligned} \tag{5.20}$$

* Chapter 5 is part of [S3].

Compare $D_f^+V(x_t)$ with $D_{(f+g)}^+V(x_t)$, i.e., the derivative along trajectories of $\dot{x}(t) = f(x_t)$ with the derivative along trajectories of $\dot{x}(t) = f(x_t) + g(x_t)$. A difference can only occur in terms that involve $\dot{x}(t)$ in

$$\begin{aligned} D_{(f+g)}^+V(x_t) = 2 \quad & \underbrace{\dot{x}^\top(t)}_{(f(x_t)+g(x_t))^\top} \quad P_{\mathrm{xx}}\, x(t) \\ & + 2\Big(\underbrace{\dot{x}^\top(t)}_{(f(x_t)+g(x_t))^\top} \int_{t-h}^{t} P_{\mathrm{xz}}(\eta - t)x(\eta)\,\mathrm{d}\eta + x^\top(t)\frac{\mathrm{d}}{\mathrm{d}t}\int_{t-h}^{t}(\ldots)\,\mathrm{d}\eta\Big) \\ & + \frac{\mathrm{d}}{\mathrm{d}t}\int_{t-h}^{t}\int_{t-h}^{t}(\ldots)\,\mathrm{d}\eta\,\mathrm{d}\xi + \frac{\mathrm{d}}{\mathrm{d}t}\int_{t-h}^{t}(\ldots)\,\mathrm{d}\eta \end{aligned} \tag{5.21}$$

(the Leibniz integral rule applied to the abbreviated terms cannot give rise to $\dot{x}(t)$). Thus, the scalar difference is $2g^\top(x_t)\Big(P_{\mathrm{xx}}\, x(t) + \int_{t-h}^{t} P_{\mathrm{xz}}(\eta - t)x(\eta)\,\mathrm{d}\eta\Big) = 2g^\top(x_t)v(x_t)$. □

The proposed defining equation (5.17) is tailored to the objective that the above derivative (5.19) shall easily be proven to be nonpositive. In fact, a desired result $D_{(f+g)}^+V(\phi) \leq -\ell(\mathcal{C}\phi)$ with a chosen offset function ℓ can be prescribed. To this end, the perturbation restriction $w(\zeta, a(\zeta)) \geq 0$ introduced in (5.14), is strengthened to $w(\zeta, a(\zeta)) \geq \ell(\zeta)$ in (5.22) below. Being chosen as some (not necessarily quadratic) nonnegative function with $\ell(0) = 0$, a small offset ℓ comes along with a slight reduction of the permissible region for the graph of a. The latter is indicated by the turquoise shading in the last column of Table 5.1. Of course, a vanishing offset $\ell(\zeta) \equiv 0$ suffices if only $D_{(f+g)}^+V(\phi) \leq 0$ is desired, in which case (5.22) is merely the original perturbation restriction from (5.14).

Theorem 5.3.2 (Main result of Section 5.3). *Let V be an LK functional of robust type described by Def. 5.2.4. Then for any $\phi \in C$ for which the perturbation restriction (5.14) is exceeded by a given offset function $\ell\colon \mathbb{R}^p \to \mathbb{R}$ in the sense of*

$$w(\mathcal{C}\phi, a(\mathcal{C}\phi)) \geq \ell(\mathcal{C}\phi), \tag{5.22}$$

the LK-functional derivative along solutions of the perturbed equation satisfies

$$D^+_{(f+g)}V(\phi) \le -\ell(\mathcal{C}\phi). \tag{5.23}$$

If $e(\phi) \not\equiv 0$ *in (5.17), then* $D^+_{(f+g)}V(\phi) \le -\ell(\mathcal{C}\phi) - e(\phi)$.

Proof. Consider (5.19) with $g(\phi)$ from (5.7),

$$D^+_{(f+g)}V(\phi) = D^+_f V(\phi) - 2v^\top(\phi)Ba(\mathcal{C}\phi). \tag{5.24}$$

The defining equation (5.17) for $D^+_f V(\phi)$ involves the term $\hat{b}^\top\hat{b}$ when abbreviating

$$\hat{b}^\top := \big[v^\top(\phi)B - (\mathcal{C}\phi)^\top\Pi_{\zeta a}\big](-\Pi_{aa})^{-1/2}. \tag{5.25}$$

Thus, (5.24) (assuming $\mathrm{e}(\phi) \equiv 0$) can be written as

$$D^+_{(f+g)}V(\phi) = -(\mathcal{C}\phi)^\top\Pi_{\zeta\zeta}(\mathcal{C}\phi) - \hat{b}^\top\hat{b} - 2v^\top(\phi)Ba(\mathcal{C}\phi). \tag{5.26}$$

Adding $0 = -\|\hat{b} + \hat{a}\|_2^2 + \hat{b}^\top\hat{b} + 2\hat{b}^\top\hat{a} + \hat{a}^\top\hat{a}$ with

$$\hat{a} := (-\Pi_{aa})^{1/2}a(\mathcal{C}\phi), \tag{5.27}$$

and noting that a part of the mixed term

$$2\hat{b}^\top\hat{a} = 2\big[v^\top(\phi)B - (\mathcal{C}\phi)^\top\Pi_{\zeta a}\big]a(\mathcal{C}\phi) \tag{5.28}$$

eliminates the perturbation term from (5.26), leads to

$$D^+_{(f+g)}V(\phi) = -(\mathcal{C}\phi)^\top\Pi_{\zeta\zeta}(\mathcal{C}\phi) - \|\hat{b} + \hat{a}\|_2^2 - 2(\mathcal{C}\phi)^\top\Pi_{\zeta a}a(\mathcal{C}\phi) + \hat{a}^\top\hat{a}.$$

Due to $\hat{a}^\top\hat{a} = a^\top(\mathcal{C}x_t)(-\Pi_{aa})a(\mathcal{C}\phi)$, the resulting

$$\begin{aligned} D^+_{(f+g)}V(\phi) = &-(\mathcal{C}\phi)^\top\Pi_{\zeta\zeta}(\mathcal{C}\phi) - \|\hat{b} + \hat{a}\|_2^2 \\ &- 2(\mathcal{C}\phi)^\top\Pi_{\zeta a}a(\mathcal{C}\phi) - a^\top(\mathcal{C}\phi)\Pi_{aa}a(\mathcal{C}\phi) \end{aligned}$$

explicitly involves the perturbation restriction (5.14b) in

$$D^+_{(f+g)}V(\phi) = -w(\mathcal{C}\phi, a(\mathcal{C}\phi)) - \|\hat{b} + \hat{a}\|_2^2. \tag{5.29}$$

Hence, (5.22) immediately leads to the estimation (5.23). If $-e(\phi) \neq 0$ in (5.17), this term also occurs in (5.29). □

Remark 5.3.3 (Time-varying perturbation). *The results can straightforwardly be extended to a time-varying $a(\zeta, t)$. Requiring $w(\mathcal{C}\phi, a(\mathcal{C}\phi, t)) \geq \ell(\mathcal{C}\phi)$ leads as well to $D^+_{(f+g)}V(\phi) \leq -\ell(\mathcal{C}\phi)$.*

The following corollary focuses on a perturbation restriction in form of a linear norm bound, comparable to (1.17) known from complete-type LK functionals.

Corollary 5.3.4 (Functional for a linear norm bound). *Assume $\gamma > 0$ is chosen such that an LK functional $V(\phi)$ having the form (5.5) exists that solves*

$$D^+_f V(\phi) = -\gamma^2(\mathcal{C}\phi)^\top \mathcal{C}\phi - v^\top(\phi)BB^\top v(\phi), \tag{5.30}$$

where v is given by (5.6) (see Cor. 5.5.13 in Section 5.5.5 for a respective range of γ). If

$$\|a(\mathcal{C}\phi)\|_2 \leq \sqrt{\gamma^2 - k_3}\|\mathcal{C}\phi\|_2, \tag{5.31}$$

with some $k_3 \in [0, \gamma^2)$, then the derivative of $V(\phi)$ along solutions of the perturbed RFDE (1.2) is nonpositive with

$$D^+_{(f+g)}V(\phi) \leq -k_3\|\mathcal{C}\phi\|_2^2. \tag{5.32}$$

Proof. Consider $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$ from Table 5.1, row (I|a). The defining equation (5.17) becomes (5.30). Choosing $\ell(\mathcal{C}x_t) = k_3\|\mathcal{C}x_t\|_2^2$, the strengthened perturbation restriction (5.22) becomes (5.31), and (5.23) is (5.32). □

A quadratic offset ℓ, as it has been chosen in the above corollary for simplicity, results in tightened sector slopes like (5.31). However, in the example sketched in row (II|a) of Table 5.1, $a(\zeta)$ at $\zeta = 0$ is already tangent to the original sector bound with $\ell(\zeta) \equiv 0$. Thus, tightened sector slopes are inappropriate in this case. Choosing rather

$$\ell(\zeta) = \kappa(\|\zeta\|) \tag{5.33}$$

with a not specified class-K function $\kappa \in \mathcal{K}$ is less demanding in terms of the perturbation restriction, and simultaneously amounts to what is usually the desired estimation for $D^+_{(f+g)}V(\phi)$ in (5.23). Provided ζ is considered on a bounded set, a tightening via (5.33) only results in an open rather than a closed sector condition.

Lemma 5.3.5. *Let $\Omega \subset \mathbb{R}^p$ be a bounded set. Then the existence of a class-K function $\kappa \in \mathcal{K}$ such that $\forall \zeta \in \Omega : w(\zeta, a(\zeta)) \geq \kappa(\|\zeta\|)$ is equivalent to the open sector restriction $w(\zeta, a(\zeta)) > 0$ for all $\zeta \in \Omega \setminus \{0_p\}$.*

Proof. Note that $\zeta \mapsto \beta(\zeta) = w(\zeta, a(\zeta))$ is a continuous function $\beta \colon \mathbb{R}^p \to \mathbb{R}$ with $\beta(0_p) = 0$. □

Finally, in terms of the classical LK theorem, the following can be concluded: Choosing ℓ according to (5.33), the functional satisfies

$$D^+_{(f+g)}V(\phi) \overset{(5.23)}{\leq} -\ell(\mathcal{C}\phi) = -\kappa\left(\left\|\begin{bmatrix} C_1\phi(-h) \\ C_0\phi(0) \end{bmatrix}\right\|\right). \tag{5.34}$$

Thus, it meets the well-known monotonicity condition $\exists \kappa_3 \in \mathcal{K}$, $\forall \phi \in C$: $D^+_{(f+g)}V(\phi) \leq -\kappa_3(\|\phi(0)\|)$ from Theorem 4.5.1 whenever C_0 in (5.8) is chosen as a full-rank matrix.

To sum up, in complete-type LK functionals the restrictive linear norm bound on the perturbation is required to ensure that the perturbation cannot turn the, by construction, nonpositive $D^+_f V(x_t)$ into to a positive $D^+_{(f+g)}V(x_t)$. In contrast, LK functionals of robust type are designed in such a way that any

$a(\cdot)$ residing within the sector from the prescribed perturbation restriction gives rise to a nonpositive derivative $D^+_{(f+g)}V(x_t) \leq 0$. Moreover, a strengthened $D^+_{(f+g)}V(x_t) \leq -\ell(\mathcal{C}x_t)$ with some offset function ℓ is simply accomplished by a respective tightening of the perturbation restriction.

5.4 Positive-Definiteness Properties*

Besides of the partial negative definiteness of $D_{(f+g)}V(\phi)$ discussed above, also the partial positive definiteness of $V(\phi)$ is of interest for the usability of the LK functional (see (4.28)).

Complete-type and related LK functionals are known to satisfy the partial positive definiteness condition demanded in the classical LK theorem if and only if the equilibrium of the nominal system is exponentially stable [110]. (In fact, the situation is analogous to the delay-free template from Section 1.2, where positive definiteness of P from a Lyapunov equation with $Q \succ 0_{n\times n}$ holds if and only if A is Hurwitz.)

Similar applies for LK functionals of robust type. The nominal exponential stability is again already sufficient for the nonnegativity of $V(\phi)$, at least in those cases in which the perturbation sector described by $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$ contains the zero perturbation $a(\zeta) \equiv 0_m$ in its inner. The latter applies for the linear norm bound, and, in fact, for any perturbation restriction with $\Pi_{\zeta\zeta} \succ 0$ (choosing $K = 0_{m\times p}$ in (5.35) below).

However, the subsequent theorem can also be applied in more general cases (the equilibrium of the nominal system might even be unstable if $a(\zeta) \equiv 0_m$ does not belong to the sector). The only condition to be imposed is that a stabilizing linear control law $a(\zeta) = K\zeta$, $K \in \mathbb{R}^{m\times p}$ (see, e.g., Remark 5.5.10), can be found in the inner of the sector of allowed perturbations. Note that the latter is a necessary condition for the simultaneous exponential stability under all considered

* Chapter 5 is part of [S3].

perturbations anyway since the linear control law is itself part of that perturbation family.

Theorem 5.4.1 (Partial positive definiteness). *If there exists a* $K \in \mathbb{R}^{m\times p}$ *such that the linear control law* $a(\mathcal{C}x_t) = K\mathcal{C}x_t$

(a) *belongs to the interior of the considered perturbation family (5.14), i.e.,* K *satisfies*

$$\Pi_{\zeta\zeta} + \Pi_{\zeta a}K + K^\top \Pi_{\zeta a}^\top + K^\top \Pi_{aa} K \succ 0_{p\times p}, \tag{5.35}$$

and

(b) *renders the zero equilibrium of*

$$\dot{x}(t) = f(x_t) - BK\mathcal{C}x_t \tag{5.36}$$

exponentially stable,

then $\exists k_{1,0} > 0, \exists k_{1,1} > 0, \forall \phi \in C$:

$$k_{1,0}\frac{\|C_0\phi(0)\|^3}{\|\phi\|_C} + k_{1,1}\frac{\|C_1\phi(0)\|^3}{\|\phi\|_C} \le V(\phi). \tag{5.37}$$

Proof. The argument in $V(\phi)$ is an arbitrary function $\phi \in C([-h,0],\mathbb{R}^n)$. The latter is taken as an initial condition $x_0 = \phi$ for the stabilized problem (5.36). Knowing that the resulting state x_t decays with increasing time t exponentially to $0_{n_{[-h,0]}}$, where $V(0_{n_{[-h,0]}}) = 0$, and knowing that V is quadratic, $V(\phi)$ becomes

$$\begin{aligned} V(x_0) &= -\Big(\lim_{t_1\to\infty} \underbrace{V(x_{t_1})}_{\to 0,\ \text{exp.}} - V(x_0)\Big) \\ &= -\int_0^\infty D^+_{(f-BK\mathcal{C})}V(x_t)\,\mathrm{d}t \overset{(5.23)}{\ge} \int_0^\infty \ell(\mathcal{C}x_t)\,\mathrm{d}t, \end{aligned} \tag{5.38}$$

given ℓ is chosen such that (5.22) holds for the involved $a(\mathcal{C}x_t) = K\mathcal{C}x_t$. Using $\alpha = K\zeta$ in (5.14) shows that a possible function ℓ in (5.22) is

$$\ell(\zeta) = k\|\zeta\|_2^2 \quad \text{with} \tag{5.39}$$
$$k = \lambda_{\min}(\Pi_{\zeta\zeta} + \Pi_{\zeta a}K + K^\top \Pi_{\zeta a}^\top + K^\top \Pi_{aa} K) \overset{(5.35)}{>} 0.$$

Hence, (5.38), where $\mathcal{C}$ is defined in (5.8), becomes

$$V(x_0) \geq \int_0^\infty k\|\mathcal{C}x_t\|_2^2 \,\mathrm{d}t \tag{5.40}$$
$$= \int_0^\infty k\|C_0 x(t)\|_2^2 \,\mathrm{d}t + \int_{-h}^\infty k\|C_1 x(t)\|_2^2 \,\mathrm{d}t. \tag{5.41}$$

To make the dependency on $x(0)$ visible (similar to [93]), for each term the integration is restricted to a small time interval where $\|C_j x(t)\|_2$, $j \in \{1, 2\}$, deviates less than half from its value at $t = 0$, and thus $\|C_j x(t)\| \geq \frac{1}{2}\|C_j x(0)\|$. Lemma B.2.1 expresses a time bound $\delta(\alpha)$ that guarantees for $t \in [0, \delta(\alpha)]$ an arbitrarily small deviation $\|x(t) - x(0)\| \leq \alpha\|x_0\|_C$ relative to the initial function. Thus, $\|C_j x(t) - C_j x(0)\| \leq \alpha\|C_j\|\|x_0\|_C$. By the reverse triangle inequality[1]

$$\|C_j x(t)\| \geq \|C_j x(0)\| - \alpha\|C_j\|\|x_0\|_C \quad \text{if } t \in [0, \delta(\alpha)].$$

Hence, by considering only $t \in [0, \delta(\alpha_j)]$ with

$$\alpha_j = \frac{\frac{1}{2}\|C_j x(0)\|}{\|C_j\|\|x_0\|_C}, \tag{5.42}$$

it is achieved that $\|C_j x(t)\| \geq \frac{1}{2}\|C_j x(0)\|$, and (5.41) becomes

$$V(x_0) \geq \int_0^{\delta(\alpha_0)} \tfrac{k}{4}\|C_0 x(0)\|_2^2 \,\mathrm{d}t + \int_0^{\delta(\alpha_1)} \tfrac{k}{4}\|C_1 x(0)\|_2^2 \,\mathrm{d}t$$
$$= \delta(\alpha_0)\tfrac{k}{4}\|C_0 x(0)\|_2^2 + \delta(\alpha_1)\tfrac{k}{4}\|C_1 x(0)\|_2^2.$$

[1] respectively, $\|C_j x(0)\| = \|C_j x(t) - \big(C_j x(t) - C_j x(0)\big)\| \leq \|C_j x(t)\| + \|C_j x(t) - C_j x(0)\|$

According to Lemma B.2.1, δ can be chosen as a linear function, $\delta(\alpha) = m\alpha$, $m > 0$, yielding

$$V(x_0) \geq \frac{\frac{m}{2}\|C_0 x(0)\|}{\|C_0\|\|x_0\|_C} \frac{k}{4}\|C_0 x(0)\|_2^2 + \frac{\frac{m}{2}\|C_1 x(0)\|}{\|C_1\|\|x_0\|_C} \frac{k}{4}\|C_1 x(0)\|_2^2,$$

where $x(0) = x_0(0) = \phi(0)$ since the initial function x_0 represents the used argument ϕ in V. □

Two special cases should be emphasized.

First, as a consequence of the above theorem, if C_0 is chosen as a full-rank matrix or, more generally, if the combination $\left[\begin{smallmatrix} C_1 \\ C_0 \end{smallmatrix}\right] \in \mathbb{R}^{(p_0+p_1)\times n}$ has rank n, then $V(\phi)$ shares the same partial positive definiteness properties as the (not complete-type but related) LK functionals described in [93, 136].

Corollary 5.4.2. *Let* $\mathrm{rk}(\left[\begin{smallmatrix} C_0 \\ C_1 \end{smallmatrix}\right]) = n$. *If the conditions in Theorem 5.4.1 hold, then*
(a) **(Local cubic bound)** *for any* $r > 0$, *there exists a* $k_1 > 0$ *such that for all* $\phi \in C$ *with* $\|\phi\|_C < r$ *it holds* $k_1\|\phi(0)\|^3 \leq V(\phi)$;
(b) **(Global quadratic bound on a Razumikhin-like set)** *there exists a* $k_1 > 0$ *such that for all* $\phi \in C$ *with* $\|\phi\|_C = \|\phi(0)\|$ *it holds* $k_1\|\phi(0)\|^2 \leq V(\phi)$.

Proof. Theorem 5.4.1, using that (5.37) simplifies to $\exists k_1 > 0 : k_1 \frac{\|\phi(0)\|^3}{\|\phi\|_C} \leq V(\phi)$ if $\left[\begin{smallmatrix} C_0 \\ C_1 \end{smallmatrix}\right]$ has full rank. □

Second, if C_1 has full rank (even if $p_0 = 0$ in (5.8)), $V(\phi)$ even shares the same partial positive definiteness properties as the ones described in [111, Thm. 5] for LK functionals of complete type.

Theorem 5.4.3 (Global quadratic bound). *Let* $\mathrm{rk}(C_1) = n$. *If the conditions in Theorem 5.4.1 hold, then* $\exists k_1 > 0, \forall \phi \in C : k_1\|\phi(0)\|^2 \leq V(\phi)$.

Proof. The starting point is (5.40). If not only $\mathrm{rk}(C_1) = n$ but also $\mathrm{rk}(C_0) = n$ is satisfied, then the proof proceeds analogously to [110, Lem. 2.10]. That is,

$$V(x_0) \overset{(5.40)}{\geq} \int_0^\infty k\|\mathcal{C}x_t\|_2^2 \,\mathrm{d}t \tag{5.43}$$

$$\overset{(5.8)}{=} \int_0^\infty \begin{bmatrix} x(t) \\ x(t-h) \end{bmatrix}^\top \begin{bmatrix} C_0^\top C_0 & 0 \\ 0 & C_1^\top C_1 \end{bmatrix} \begin{bmatrix} x(t) \\ x(t-h) \end{bmatrix} \mathrm{d}t \tag{5.44}$$

is set in relation to

$$k_1\|x_0(0)\|_2^2 = k_1\|x(0)\|_2^2 = -\lim_{t_1\to\infty} k_1 \underbrace{x^\top(t_1)\,x(t_1)}_{\to 0 \quad \text{if ES}} + k_1 x^\top(0)\,x(0)$$

$$= -\lim_{t_1\to\infty} \int_0^{t_1} \underbrace{\frac{\mathrm{d}}{\mathrm{d}t}\big(k_1 x^\top(t)\,x(t)\big)}_{2k_1x^\top(t)\left(A_0x(t)+A_1x(t-h)\right)} \mathrm{d}t \tag{5.45}$$

by considering

$$V(x_0) - k_1\|x_0(0)\|_2^2 \tag{5.46}$$

$$\geq \int_0^\infty \begin{bmatrix} x(t) \\ x(t-h) \end{bmatrix}^\top \underbrace{\left(\begin{bmatrix} C_0^\top C_0 & 0 \\ 0 & C_1^\top C_1 \end{bmatrix} + k_1 \begin{bmatrix} 2\,\mathrm{sym}(A_0) & A_1 \\ A_1^\top & 0 \end{bmatrix}\right)}_{=:M} \begin{bmatrix} x(t) \\ x(t-h) \end{bmatrix} \mathrm{d}t.$$

If $C_0^\top C_0 \succ 0_{n\times n}$ and $C_1^\top C_1 \succ 0_{n\times n}$, then M is positive semidefinite for any sufficiently small k_1. Thus, it can be concluded that $V(x_0) - k_1\|x_0(0)\|_2^2 \geq 0$ for some $k_1 > 0$ if $\mathrm{rk}(C_0) = n$. Otherwise, if $\mathrm{rk}(C_0) < n$, note that

$$\int_{-h}^\infty k\|C_1x(t)\|_2^2 \,\mathrm{d}t \geq \int_0^\infty k\|C_1x(t)\|_2^2 \,\mathrm{d}t \tag{5.47}$$

is a lower bound on (5.41). Hence, a convex combination of both (5.40) and this lower bound on (5.40) can be used as starting point, and the same arguments apply. □

Concerning the classical LK theorem (Theorem 4.5.1), the following can be summarized: With a full-rank choice for C_1, an LK functional of robust type meets the partial positive definiteness requirement $\exists \kappa_1 \in \mathcal{K} : \kappa_1(\|\phi(0)\|) \leq V(\phi)$ globally (Theorem 5.4.3). With a full-rank choice for C_0, or more generally, if $\left[\begin{smallmatrix} C_1 \\ C_0 \end{smallmatrix}\right]$ has full column rank, it still meets this requirement on any arbitrarily large bounded set (Corollary 5.4.2).

Remark 5.4.4 (Observability). *Note that, in (5.40) and (5.41), observability Gramians [47, Def. 6.2.12] can be recognized. Therefore, if C_0, C_1 do not have full rank but, by chance, they render (5.36) observable in a certain sense (see [47, Cor. 6.2.15]), then the above discussed lower bound in terms of $\|\phi(0)\|$ still exists.*

5.5 Solvability of the Defining Equation

The objective of the present section is to characterize admissible perturbation restrictions for which the existence of a solution V of the defining equation (5.17) can be guaranteed. That is, the admissible range of the involved parameter that controls the resulting sector size, e.g., γ, ρ, or K_1 in Table 5.1, shall be determined. The objective will be achieved by resorting to a version of the Kalman–Yakubovich–Popov (KYP) lemma for C_0-semigroups on Hilbert spaces. To this end, the problem under consideration must first be brought in relation to an operator-valued algebraic Riccati equation which can be tackled by that lemma.

5.5.1 A Splitting Approach*

If C_1 is nonzero, then $(\mathcal{C}\phi)^\top \Pi_{\zeta\zeta}(\mathcal{C}\phi)$ on the right-hand side of the defining equation (5.17) explicitly depends on $\phi(-h)$, respectively $x_t(-h) = x(t-h)$. However, for the operator-theoretic treatment in the next two sections, as well as

* Chapter 5 is part of [S3].

the analysis of the numerical approach, a problem without such a dependency is more convenient. That is why the LK functional $V(\phi)$ will be split into, firstly, a part $V_0(\phi)$ that results from a defining equation without a quadratic delayed term, and, secondly, a remaining part $V_1(\phi)$. Due to the following transformation, the derivations do not have to cope with mixed term matrices $\Pi_{\zeta a}$, even if the original perturbation restriction belongs to row (II) or (III) in Table 5.1.

Lemma 5.5.1 (Transformation I). *$V(\phi)$ is an LK functional of robust type w.r.t. $\dot{x}(t) = f(x_t) = A_0x(t) + A_1x(t-h)$, $(B,\mathcal{C})$, and $\Pi = (\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$ if and only if $V(\phi) = V^{\mathrm{I}}(\phi)$ is an LK functional of robust type w.r.t. the transformed system*

$$\dot{x}(t) = A_0^{\mathrm{I}}x(t) + A_1^{\mathrm{I}}x(t-h) =: f^{\mathrm{I}}(x_t), \tag{5.48}$$

$$\textit{with} \quad A_0^{\mathrm{I}} = A_0 - B(-\Pi_{aa})^{-1}\Pi_{\zeta a}^{\top}\begin{bmatrix} 0_{p_1\times n} \\ C_0 \end{bmatrix}, \tag{5.49}$$

$$A_1^{\mathrm{I}} = A_1 - B(-\Pi_{aa})^{-1}\Pi_{\zeta a}^{\top}\begin{bmatrix} C_1 \\ 0_{p_0\times n} \end{bmatrix},$$

the original perturbation structure $(B,\mathcal{C})$, and the transformed perturbation restriction $\Pi^{\mathrm{I}} = (\Pi_{\zeta\zeta}^{\mathrm{I}}, \Pi_{\zeta a}^{\mathrm{I}}, \Pi_{aa}^{\mathrm{I}})$

$$\Pi_{\zeta\zeta}^{\mathrm{I}} = \Pi/\Pi_{aa} = \Pi_{\zeta\zeta} + \Pi_{\zeta a}(-\Pi_{aa})^{-1}\Pi_{\zeta a}^{\top}, \tag{5.50a}$$

$$\Pi_{\zeta a}^{\mathrm{I}} = 0, \quad \textit{and} \quad \Pi_{aa}^{\mathrm{I}} = \Pi_{aa}. \tag{5.50b}$$

Proof. Consider $g(x_t) = -B(-\Pi_{aa})^{-1}\Pi_{\zeta a}^{\top}\mathcal{C}x_t$. The defining equation (5.17) is not altered if $D_f^+V(\phi)$ and $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{\zeta\zeta})$ are replaced by $D_{(f+g)}^+V(\phi)$ from Lemma 5.3.1 and $(\Pi_{\zeta\zeta}^{\mathrm{I}}, \Pi_{\zeta a}^{\mathrm{I}}, \Pi_{\zeta\zeta}^{\mathrm{I}})$. □

Henceforth, it is assumed that (5.50a) has the structure

$$\Pi_{\zeta\zeta}^{\mathrm{I}} = \begin{bmatrix} \Pi_{\zeta\zeta}^{\mathrm{I},11} & 0_{p_1\times p_0} \\ 0_{p_0\times p_1} & \Pi_{\zeta\zeta}^{\mathrm{I},00} \end{bmatrix}. \tag{5.51}$$

With that block diagonal structure (5.51), the first term in the defining equation (5.17) for $V^{\mathrm{I}}(\phi) = V(\phi)$ becomes

$$(\mathcal{C}\phi)^\top \Pi^{\mathrm{I}}_{\zeta\zeta}(\mathcal{C}\phi) = \phi^\top(-h) \underbrace{C_1^\top \Pi^{\mathrm{I},11}_{\zeta\zeta} C_1}_{Q_1} \phi(-h) + \phi^\top(0) \underbrace{C_0^\top \Pi^{\mathrm{I},00}_{\zeta\zeta} C_0}_{Q_0} \phi(0). \tag{5.52}$$

For notational compactness, consider $e(\phi) \equiv 0$ in (5.17) (an extension to $e(\phi) \not\equiv 0$ is straightforward). The following splitting of the functional is along the lines of what has been pursued in Lemma 4.6.1 for complete-type LK functionals. However, for complete-type LK functionals, the defining equation is only a linear equation, for which the involved superposition from splitting the equation is obviously unproblematic. In contrast, LK functionals of robust type have a quadratic defining equation—but the required splitting still turns out to be possible.

Lemma 5.5.2 (Splitting). *Assume $\Pi^{\mathrm{I}}_{\zeta\zeta}$ in (5.50a) has the block diagonal structure (5.51) giving rise to Q_0, Q_1 from (5.52). Then any solution V of (5.17) can be split into*

$$V(\phi) = V_0(\phi) + V_1(\phi), \tag{5.53}$$

$$V_1(\phi) = \int_{-h}^{0} \phi^\top(\eta)\, Q_1 \phi(\eta)\, \mathrm{d}\eta, \tag{5.54}$$

where V_0 satisfies the modified defining equation

$$D^+_{f^{\mathrm{I}}} V_0(\phi) = -\phi^\top(0)(Q_0 + Q_1)\phi(0) - v_0^\top(\phi) B\, (-\Pi_{aa})^{-1} B^\top v_0(\phi) \tag{5.55}$$

without a term $\phi^\top(-h) Q_1 \phi(-h)$, $Q_1 \in \mathbb{R}^{n\times n}$.

Proof. According to Lemma 5.5.1, and with (5.52), the defining equation (5.17) for the overall functional V at $\phi = x_t$ is

$$\begin{aligned} D^+_{f^{\mathrm{I}}} V(x_t) = &-x^\top(t) Q_0 x(t) - v^\top(x_t) B(-\Pi_{aa})^{-1} B^\top v(x_t) \\ &- x^\top(t-h) Q_1 x(t-h). \end{aligned} \tag{5.56}$$

The latter shall be split into a sum $D^+_{f^{\mathrm{I}}} V(x_t) = D^+_{f^{\mathrm{I}}} V_0(x_t) + D^+_{f^{\mathrm{I}}} V_1(x_t)$. Note that (5.54) in terms of $\phi = x_t$ reads $V_1(x_t) = \int_{t-h}^{t} x^\top(\eta)\, Q_1 x(\eta)\, \mathrm{d}\eta$ with

$$D^+_{f^{\mathrm{I}}} V_1(x_t) = x^\top(t) Q_1 x(t) - x^\top(t-h) Q_1 x(t-h). \tag{5.57}$$

Thus, the remaining unknown V_0 must satisfy

$$\begin{aligned} D^+_{f^{\mathrm{I}}} V_0(x_t) = &-x^\top(t) Q_0 x(t) - x^\top(t) Q_1 x(t) - \big(v_0(x_t) \\ &+ v_1(x_t)\big)^\top B(-\Pi_{aa})^{-1} B^\top \big(v_0(x_t) + v_1(x_t)\big) \end{aligned} \tag{5.58}$$

where $v(x_t) = v_0(x_t) + v_1(x_t)$ are the corresponding subfunctionals according to (5.6). In V_1 from (5.54), the kernel functions in terms of (5.5) are $P_{\mathrm{xx}} = 0$, $P_{\mathrm{xz}}(\eta) \equiv 0$, $P_{\mathrm{zz}}(\xi, \eta) \equiv 0$, $P_{\mathrm{zz,diag}}(\eta) \equiv Q_1$, and thus (5.6) yields

$$v_1(x_t) \equiv 0. \tag{5.59}$$

Consequently, (5.58) becomes (5.55). □

5.5.2 Operator-Based Description*

As discussed in Section 4.6.2, LK functionals of complete type can be written as a quadratic form in $L_2 \times \mathbb{R}^n$ with an operator from an operator-valued Lyapunov

* Chapter 5 is part of [S3].

equation. Having the same structure, LK functionals of robust type can analogously be described. As will be shown below, only the Lyapunov equation that determines the involved operator is replaced by an algebraic Riccati equation.

All operators in this and the next section are still real, but, in the next section, complex arguments will occur. That is why, in contrast to Section 4.6.2, the Hilbert space $M_2 = L_2([-h,0],\mathbb{C}^n) \times \mathbb{C}^n$ over the field of complex numbers with the inner product[2]

$$\left\langle \left[\begin{smallmatrix}\phi_1\\ r_1\end{smallmatrix}\right], \left[\begin{smallmatrix}\phi_2\\ r_2\end{smallmatrix}\right] \right\rangle_{M_2} = \int_{-h}^{0} (\phi_2(\theta))^H \phi_1(\theta)\,\mathrm{d}\theta + r_2^H r_1, \tag{5.60}$$

$\phi_1, \phi_2 \in L_2$, $r_1, r_2 \in \mathbb{C}^n$, is henceforth considered in

$$\left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right] \in C([-h,0],\mathbb{R}^n) \times \mathbb{R}^n \subset L_2([-h,0],\mathbb{C}^n) \times \mathbb{C}^n. \tag{5.61}$$

Note that, in the case of complex arguments, $\langle \mathsf{z}, \mathsf{x}\rangle_{M_2}$ with $\mathsf{z}, \mathsf{x} \in M_2$ is no longer equivalent to $\langle \mathsf{x}, \mathsf{z}\rangle_{M_2} = \overline{\langle \mathsf{z}, \mathsf{x}\rangle_{M_2}}$. The focus of this section is on the functional V_0 that has been defined in Lemma 5.5.2. Compared to the overall functional (5.5), it does not show the term $P_{\mathrm{zz,diag}}$, which is only due to $V_1(\phi)$, cf. [115]. Based on (5.60), V_0 can be written as a quadratic form in M_2

$$V_0(\phi) = \left\langle \mathscr{P}_0 \left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right], \left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right] \right\rangle_{M_2} \tag{5.62}$$

[2] The convention to define inner products linear in the first argument—and thus, by the conjugate symmetry, conjugate linear in the second argument—is adhered to, e.g., $\langle r_1, r_2\rangle_{\mathbb{C}^n} = r_1^{\mathsf{T}} \overline{r}_2 = r_2^H r_1$.

with a real self-adjoint operator $\mathscr{P}_0 \colon M_2 \to M_2$ relying on suboperators $\mathscr{P}_{zz} \colon L_2 \to L_2$, and $\mathscr{P}_{zx} \colon \mathbb{C}^n \to L_2$,

$$\mathscr{P}_0 \begin{bmatrix} \phi \\ r \end{bmatrix} = \begin{bmatrix} \mathscr{P}_{zz}\phi + \mathscr{P}_{zx} r \\ \mathscr{P}_{zx}^* \phi + P_{\mathrm{xx}} r \end{bmatrix} = \begin{bmatrix} \tilde{\phi} \\ \tilde{r} \end{bmatrix}, \tag{5.63}$$

$$\text{with } \tilde{\phi}(\theta) = \int_{-h}^{0} P_{\mathrm{zz}}(\theta,\eta)\phi(\eta)\,\mathrm{d}\eta + (P_{\mathrm{xz}}(\theta))^H r,$$

$$\tilde{r} = \int_{-h}^{0} P_{\mathrm{xz}}(\eta)\phi(\eta)\,\mathrm{d}\eta + P_{\mathrm{xx}} r$$

that incorporate the kernel functions from (5.5). The quadratic form (5.62) shall be used in the defining equation (5.55). Consider $\dot{x}(t) = A_0^{\mathrm{I}} x(t) + A_1^{\mathrm{I}} x(t-h) = f^{\mathrm{I}}(x_t)$ from (5.48). According to (3.12), the evolution of $\left[\begin{smallmatrix} x_t \\ x_t(0) \end{smallmatrix}\right] \in M_2$ obeys the abstract ODE $\frac{\mathrm{d}}{\mathrm{d}t}\left[\begin{smallmatrix} x_t \\ x_t(0) \end{smallmatrix}\right] = \mathscr{A}^{\mathrm{I}}\left[\begin{smallmatrix} x_t \\ x_t(0) \end{smallmatrix}\right]$, with $\mathscr{A}^{\mathrm{I}} \colon D(\mathscr{A}^{\mathrm{I}}) \to M_2$,

$$\mathscr{A}^{\mathrm{I}} \begin{bmatrix} \phi \\ r \end{bmatrix} = \begin{bmatrix} \phi' \\ A_0^{\mathrm{I}} r + A_1^{\mathrm{I}} \phi(-h) \end{bmatrix}, \tag{5.64}$$

$$D(\mathscr{A}^{\mathrm{I}}) = \left\{ \left[\begin{smallmatrix} \phi \\ r \end{smallmatrix}\right] \in M_2 : r = \phi(0), \phi' \in L_2, \phi \in AC \right\}.$$

Using that abstract ODE, see [47], the left-hand side of (5.55) is

$$D_{f^{\mathrm{I}}}^{+} V_0(\phi) = \left\langle \mathscr{P}_0 \mathscr{A}^{\mathrm{I}} \left[\begin{smallmatrix} \phi \\ \phi(0) \end{smallmatrix}\right], \left[\begin{smallmatrix} \phi \\ \phi(0) \end{smallmatrix}\right] \right\rangle_{M_2} + \left\langle (\mathscr{A}^{\mathrm{I}})^* \mathscr{P}_0 \left[\begin{smallmatrix} \phi \\ \phi(0) \end{smallmatrix}\right], \left[\begin{smallmatrix} \phi \\ \phi(0) \end{smallmatrix}\right] \right\rangle_{M_2}.$$

The right-hand side of the defining equation (5.55) can also be expressed in terms of $\mathsf{x} = \left[\begin{smallmatrix} \phi \\ \phi(0) \end{smallmatrix}\right] \in M_2$. Altogether, if $V_0(\phi)$ solves (5.55), then $\mathscr{P}_0 = \mathscr{P}_0^*$ from (5.62) solves the operator-valued algebraic Riccati equation

$$\begin{aligned} &\langle \mathscr{P}_0 \mathscr{A}^{\mathrm{I}} \mathsf{x}, \mathsf{x} \rangle_{M_2} + \langle (\mathscr{A}^{\mathrm{I}})^* \mathscr{P}_0 \mathsf{x}, \mathsf{x} \rangle_{M_2} \\ &\quad = -\langle \mathscr{Q} \mathsf{x}, \mathsf{x} \rangle_{M_2} - \langle (-\Pi_{aa})^{-1} \mathscr{B}^* \mathscr{P}_0 \mathsf{x}, \mathscr{B}^* \mathscr{P}_0 \mathsf{x} \rangle_{\mathbb{C}^m} \end{aligned} \tag{5.65}$$

$\forall \mathsf{x} \in D(\mathscr{A}^{\mathrm{I}})$, where the involved operators $\mathscr{Q}\colon M_2 \to M_2$ relying on Q_0, Q_1 and $\mathscr{B}\colon \mathbb{C}^m \to M_2$ relying on $B \in \mathbb{R}^{n\times m}$, respectively its adjoint $\mathscr{B}^*\colon M_2 \to \mathbb{C}^m$, are given by

$$\mathscr{Q}\begin{bmatrix}\phi\\ r\end{bmatrix} = \begin{bmatrix}0_{L_2}\\ (Q_0+Q_1)r\end{bmatrix}, \qquad \mathscr{B}u = \begin{bmatrix}0_{L_2}\\ Bu\end{bmatrix}, \qquad \mathscr{B}^*\begin{bmatrix}\phi\\ r\end{bmatrix} = B^\top r \quad (5.66)$$

(only due to the splitting from Section 5.5.1, $\mathscr{Q}$ is a bounded operator). Conversely, the following lemma ensures that a solution $\mathscr{P}_0$ of (5.65) has the form given in (5.63). The result is well known for the stabilizing solution of algebraic Riccati equations from standard[3] LQR problems [70] and is analogously provable in the present case. As a consequence, $V(\phi) = V_0(\phi) + V_1(\phi)$ has the desired structure (5.5).

Lemma 5.5.3. *Let a bounded self-adjoint operator $\mathscr{P}_0$ be a solution of (5.65) and assume $\mathscr{A}^{\mathrm{I}}$ generates an exponentially stable[4] C_0-semigroup. Then $\mathscr{P}_0$ is described by (5.63), with $\mathscr{P}_{zz}\colon L_2 \to L_2$ being an integral operator.*

Proof. The right-hand side of (5.65) can be written as $-\langle \mathscr{Q}_{\mathrm{lyap}}\mathsf{x}, \mathsf{x}\rangle_{M_2} := -\langle \Gamma_1\mathsf{x}, \Gamma_1\mathsf{x}\rangle_{\mathbb{C}^n} - \langle \Gamma_2\mathscr{P}_0\mathsf{x}, \Gamma_2\mathscr{P}_0\mathsf{x}\rangle_{\mathbb{C}^m}$ where both $\Gamma_1\colon M_2 \to \mathbb{C}^n; \Gamma_1\left[\begin{smallmatrix}\phi\\ r\end{smallmatrix}\right] = (Q_0+Q_1)^{\frac{1}{2}}r$ and $\Gamma_2\colon M_2 \to \mathbb{C}^m; \Gamma_2\left[\begin{smallmatrix}\phi\\ r\end{smallmatrix}\right] = (-\Pi_{aa})^{-\frac{1}{2}}B^\top r$ are finite rank operators. Therefore, the arguments from [70, Thm. 5.2 and p. 102/103] apply. □

Altogether, the following is thus shown.

Theorem 5.5.4. *Assume $\dot{x}(t) = f^{\mathrm{I}}(x_t)$ defined in (5.48) has an exponentially stable equilibrium. An LK functional of robust type (Definition 5.2.4) exists if*

[3] In contrast to standard LQR problems with nonnegative costs, the present algebraic Riccati equation is associated to an indefinite LQR problem. See Section 7.1.

[4] If not $\mathscr{A}$ but only $\mathscr{A}^s := \mathscr{A} - \mathscr{B}\mathscr{K}$ with $\mathscr{K} = (-\Pi_{aa})^{-1}\mathscr{B}^*(-\mathscr{P}_0)$, cf. Remark 5.5.12, generates an exponentially stable C_0-semigroup, the statement still holds. In the proof, (5.65) is first rewritten with $\mathscr{A}^s$ on the left-hand side, yielding the right-hand side $-\langle \Gamma_1\mathsf{x}, \Gamma_1\mathsf{x}\rangle_{\mathbb{C}^n} + \langle \Gamma_2\mathscr{P}_0\mathsf{x}, \Gamma_2\mathscr{P}_0\mathsf{x}\rangle_{\mathbb{C}^m}$.

and only if a self-adjoint solution $\mathscr{P}_0$ of the operator-valued algebraic Riccati equation (5.65) exists.

Proof. In (5.53), the LK functional is split into $V = V_0 + V_1$, where V_0 from (5.55) is given by (5.62) according to the above derivations. Moreover, V_1 from (5.54) always exists. □

Finally, it should be noted that the boundedness condition on $V(\phi)$ in C that is imposed by the classical LK theorem in (4.28) is also ensured.

Lemma 5.5.5 (Upper bound in C). *If V_0 is described by (5.62) with a bounded operator $\mathscr{P}_0$ then $V = V_0 + V_1$ with V_1 from (5.54) satisfies $\exists k_2 > 0, \forall \phi \in C$: $V(\phi) \leq k_2 \|\phi\|_C^2$.*

Proof. By (5.62), $V_0(\phi) \leq \|\mathscr{P}_0\| \left\| \begin{bmatrix} \phi \\ \phi(0) \end{bmatrix} \right\|_{M_2}^2 = \|\mathscr{P}_0\| \big(\int_{-h}^{0} \|\phi(\theta)\|_2^2 \, \mathrm{d}\theta + \|\phi(0)\|_2^2 \big) \leq \|\mathscr{P}_0\|(h+1)\|\phi\|_{C,2}^2$, where $\|\phi\|_{C,2} = \max\limits_{\theta \in [-h,0]} \|\phi(\theta)\|_2$. Moreover, in (5.54), $V_1(\phi) \leq h\|Q_1\|\|\phi\|_{C,2}^2$. □

To sum up, the question of existence of an LK functional of robust type boils down to the question of solvability of the operator-valued algebraic Riccati equation (5.65).

5.5.3 Infinite-Dimensional Kalman–Yakubovich–Popov Lemma*

This section analyzes the solvability of the operator-valued algebraic Riccati equation (5.65) and thus the existence of an LK functional of robust type. To this end, consider the following Kalman–Yakubovich–Popov (KYP) lemma for C_0-semigroups on infinite-dimensional Hilbert spaces.

* Chapter 5 is part of [S3].

Lemma 5.5.6 (Infinite-dimensional KYP lemma [131, Thm. 3]). *Let X, U be complex Hilbert spaces, let $\mathscr{A}^{\mathrm{I}}\colon D(\mathscr{A}^{\mathrm{I}}) \to X$ be the infinitesimal generator of a C_0-semigroup on X, let $\mathscr{B}\colon U \to X$ be a bounded linear operator, and let*

$$\mathscr{F}(\mathsf{x}, u) = \langle F_{xx}\mathsf{x}, \mathsf{x}\rangle_X + 2\mathrm{Re}\langle F_{ux}\mathsf{x}, u\rangle_U + \langle F_{uu}u, u\rangle_U \tag{5.67}$$

be a continuous quadratic form on $X \times U$. Assuming that $\mathscr{A}^{\mathrm{I}}$ does not have a spectrum in the neighborhood of the imaginary axis, define

$$\alpha_3 = \inf_{\omega\in\mathbb{R}} \inf_{u\in U} \frac{1}{\|u\|_U^2} \mathscr{F}\big((\mathrm{i}\omega I_X - \mathscr{A}^{\mathrm{I}})^{-1}\mathscr{B}u, u\big). \tag{5.68}$$

Let $(\mathscr{A}^{\mathrm{I}}, \mathscr{B})$ be stabilizable, i.e., there exists a bounded linear operator $\mathscr{K}_s\colon X \to U$ such that $\mathscr{A}^{\mathrm{I}} - \mathscr{B}\mathscr{K}_s$ generates an exponentially stable C_0-semigroup. If $\alpha_3 > 0$, then there exist bounded linear operators $\mathscr{X}_0 = \mathscr{X}_0^*\colon X \to X$ and $\mathscr{K}\colon X \to U$ such that $\forall \mathsf{x} \in D(\mathscr{A}^{\mathrm{I}}), u \in U$:

$$2\mathrm{Re}\langle \mathscr{A}^{\mathrm{I}}\mathsf{x} + \mathscr{B}u, \mathscr{X}_0\mathsf{x}\rangle_X + \mathscr{F}(\mathsf{x}, u) = \|F_{uu}^{1/2}(\mathscr{K}\mathsf{x} + u)\|_U^2. \tag{5.69}$$

Moreover, if $\mathscr{A}^{\mathrm{I}}, \mathscr{B}, \mathscr{K}_s$ are real, then real operators $\mathscr{X}_0, \mathscr{K}$ exist. If $\alpha_3 < 0$, then no such operators exist.

The existence of an LK functional of robust type can be deduced from the given statement due to the following equivalence.

Lemma 5.5.7. *Let $\mathsf{x} \in X = M_2$, $u \in U = \mathbb{C}^m$, and*

$$\mathscr{F}(\mathsf{x}, u) = -\langle \mathscr{Q}\mathsf{x}, \mathsf{x}\rangle_{M_2} + \langle(-\Pi_{aa})u, u\rangle_{\mathbb{C}^m}. \tag{5.70}$$

Then the Lur'e equation (5.69), with $\mathscr{X}_0 = -\mathscr{P}_0$, is equivalent to the algebraic Riccati equation (5.65) and $\mathscr{K} = -(-\Pi_{aa})^{-1}\mathscr{B}^*\mathscr{P}_0$.

Proof. With $\mathscr{X}_0 = -\mathscr{P}_0$ and with $\mathscr{F}$ from (5.70), where $F_{uu} = F_{uu}^H = -\Pi_{aa}$, the Lur'e equation (5.69) becomes

$$\begin{aligned} -2\Big(\mathrm{Re}\langle\mathscr{P}_0\mathscr{A}^{\mathrm{I}}\mathsf{x},\mathsf{x}\rangle_{M_2}+\mathrm{Re}\langle u,\mathscr{B}^*\mathscr{P}_0\mathsf{x}\rangle_{\mathbb{C}^m}\Big) - \langle\mathscr{Q}\mathsf{x},\mathsf{x}\rangle_{M_2}+\langle(-\Pi_{aa})u,u\rangle_{\mathbb{C}^m} \\ = \langle\mathscr{K}\mathsf{x},(-\Pi_{aa})\mathscr{K}\mathsf{x}\rangle_{\mathbb{C}^m} + 2\mathrm{Re}\langle u,(-\Pi_{aa})\mathscr{K}\mathsf{x}\rangle_{\mathbb{C}^m} + \langle(-\Pi_{aa})u,u\rangle_{\mathbb{C}^m}. \end{aligned}$$

Comparing the mixed terms in u and x gives $(-\Pi_{aa})\mathscr{K} = -\mathscr{B}^*\mathscr{P}_0$. The terms quadratic in x result in (5.65). □

The decisive element in Lemma 5.5.6 is α_3 in (5.68). Since $\mathscr{A}^{\mathrm{I}}$ refers to (5.64), the first argument of $\mathscr{F}$ in (5.68) refers to

$$(sI_{M_2} - \mathscr{A}^{\mathrm{I}})^{-1}\mathscr{B} = \left[\begin{smallmatrix}\Phi\\ \Phi(0)\end{smallmatrix}\right], \tag{5.71}$$

$$\text{with } \Phi(\theta) = \mathrm{e}^{s\theta}H^{\mathrm{I}}(s) \tag{5.72}$$

$$\text{and } H^{\mathrm{I}}(s) = (sI - A_0^{\mathrm{I}} - \mathrm{e}^{-sh}A_1^{\mathrm{I}})^{-1}B, \tag{5.73}$$

see [47, Lem. 7.2.14]. Thus, combined with (5.70) and (5.66), the decisive α_3 from (5.68) in the KYP lemma relies on

$$\mathscr{F}\big((\mathrm{i}\omega I_{M_2} - \mathscr{A}^{\mathrm{I}})^{-1}\mathscr{B}u, u\big) = -u^H(H^{\mathrm{I}}(\mathrm{i}\omega))^H(Q_0+Q_1)H^{\mathrm{I}}(\mathrm{i}\omega)u - u^H\Pi_{aa}u. \tag{5.74}$$

5.5.4 A Result in Terms of the Original Transfer Function*

Rather than expressing the resulting existence condition $\alpha_3 > 0$ with α_3 from (5.68) in terms of $H^{\mathrm{I}}(s)$ from (5.73), the present section intends to state the existence criterion in terms of a transfer function $G(s)$ from the untransformed RFDE.

* Chapter 5 is part of [S3].

The first step is to incorporate how Q_0, Q_1 from (5.52) depend on C_0, C_1 to make

$$G^I(s) = \begin{bmatrix} C_1 \mathrm{e}^{-sh} \\ C_0 \end{bmatrix} (sI - A_0^I - \mathrm{e}^{-sh} A_1^I)^{-1} B \tag{5.75}$$

in (5.74) visible.

Lemma 5.5.8. *The following equivalence holds*

$$(H^{\mathrm{I}}(\mathrm{i}\omega))^H (Q_0 + Q_1) H^{\mathrm{I}}(\mathrm{i}\omega) = (G^{\mathrm{I}}(\mathrm{i}\omega))^H \Pi^{\mathrm{I}}_{\zeta\zeta} G^{\mathrm{I}}(\mathrm{i}\omega). \tag{5.76}$$

Proof. Based on $\Pi^{\mathrm{I}}_{\zeta\zeta}$ from (5.51) and H^{I} from (5.73), it holds

$$\begin{aligned}
&(G^{\mathrm{I}}(\mathrm{i}\omega))^H \Pi^{\mathrm{I}}_{\zeta\zeta} G^{\mathrm{I}}(\mathrm{i}\omega) \\
&\quad = (H^{\mathrm{I}}(\mathrm{i}\omega))^H \begin{bmatrix} C_1^H \mathrm{e}^{\mathrm{i}\omega h} & C_0^H \end{bmatrix} \begin{bmatrix} \Pi^{I,11}_{\zeta\zeta} & 0 \\ 0 & \Pi^{I,00}_{\zeta\zeta} \end{bmatrix} \begin{bmatrix} C_1 \mathrm{e}^{-\mathrm{i}\omega h} \\ C_0 \end{bmatrix} H^{\mathrm{I}}(\mathrm{i}\omega) \\
&\quad = (H^{\mathrm{I}}(\mathrm{i}\omega))^H (C_1^H \Pi^{\mathrm{I},11}_{\zeta\zeta} C_1 + C_0^H \Pi^{\mathrm{I},00}_{\zeta\zeta} C_0) H^{\mathrm{I}}(\mathrm{i}\omega),
\end{aligned}$$

involving Q_0 and Q_1 from (5.52). □

The second step consists in undoing the transformation from Lemma 5.5.1, to express (5.74) in terms of the original transfer function

$$G(s) = \underbrace{\begin{bmatrix} C_1 \mathrm{e}^{-sh} \\ C_0 \end{bmatrix}}_{=:C(s)} \underbrace{(sI - A_0 - \mathrm{e}^{-sh} A_1)^{-1}}_{=:\Delta^{-1}(s)} B. \tag{5.77}$$

Note that, due to Lemma 5.5.8, the negative of the overall right-hand side in (5.74) is described by the right-hand side in Lemma 5.5.9 below, where $\Pi^{\mathrm{I}} = \left[\begin{smallmatrix} \Pi^{\mathrm{I}}_{\zeta\zeta} & 0 \\ 0 & \Pi_{aa} \end{smallmatrix}\right]$, see (5.50). In contrast, $\Pi = \left[\begin{smallmatrix} \Pi_{\zeta\zeta} & \Pi_{\zeta a} \\ \Pi_{\zeta a}^\top & \Pi_{aa} \end{smallmatrix}\right]$.

Lemma 5.5.9. *Let* $Z = (-\Pi_{aa})^{-1}\Pi_{\zeta a}^{\top}$ *and assume* $\det(I_m + ZG(\mathrm{i}\omega)) \neq 0$ *for all* $\omega \in \mathbb{R}$. *Then,* $\forall v \in \mathbb{C}^m$,

$$v^H \begin{bmatrix} G(\mathrm{i}\omega) \\ -I \end{bmatrix}^H \Pi \begin{bmatrix} G(\mathrm{i}\omega) \\ -I \end{bmatrix} v = u^H \begin{bmatrix} G^I(\mathrm{i}\omega) \\ -I \end{bmatrix}^H \Pi^{\mathrm{I}} \begin{bmatrix} G^I(\mathrm{i}\omega) \\ -I \end{bmatrix} u,$$

where $u = (I + ZG(\mathrm{i}\omega))v$.

Proof. Consider the Aitken block diagonalization

$$\begin{aligned} & v^H \begin{bmatrix} G(\mathrm{i}\omega) \\ -I \end{bmatrix}^H \begin{bmatrix} \Pi_{\zeta\zeta} & \Pi_{\zeta a} \\ \Pi_{\zeta a}^{\top} & \Pi_{aa} \end{bmatrix} \begin{bmatrix} G(\mathrm{i}\omega) \\ -I \end{bmatrix} v \\ & = v^H \begin{bmatrix} G(\mathrm{i}\omega) \\ -I \end{bmatrix}^H T^H \underbrace{\begin{bmatrix} \Pi/\Pi_{aa} & 0 \\ 0 & \Pi_{aa} \end{bmatrix}}_{\Pi^{\mathrm{I}}} T \begin{bmatrix} G(\mathrm{i}\omega) \\ -I \end{bmatrix} v, \end{aligned} \tag{5.78}$$

where Π/Π_{aa} is the Schur complement (5.50a), relying on $T = \left[\begin{smallmatrix} I & 0 \\ -Z & I \end{smallmatrix}\right]$ with $Z = (-\Pi_{aa})^{-1}\Pi_{\zeta a}^{\top}$. Consider

$$T \begin{bmatrix} G(\mathrm{i}\omega) \\ -I \end{bmatrix} v = \begin{bmatrix} G(\mathrm{i}\omega)(I + ZG(\mathrm{i}\omega))^{-1} \\ -I \end{bmatrix} \underbrace{(I + ZG(\mathrm{i}\omega))v}_{u}.$$

Its upper term (cf. a closed loop transfer function with Z in the feedback path) simplifies to

$$\begin{aligned} G(\mathrm{i}\omega)\big(I + Z\,G(\mathrm{i}\omega)\,\big)^{-1} & \overset{(5.77)}{=} C(\mathrm{i}\omega)\Delta^{-1}(\mathrm{i}\omega)B\left(I + Z\,C(\mathrm{i}\omega)\Delta^{-1}(\mathrm{i}\omega)B\right)^{-1} \\ & = C(\mathrm{i}\omega)\Delta^{-1}(\mathrm{i}\omega)\left(I + B\,Z\,C(\mathrm{i}\omega)\Delta^{-1}(\mathrm{i}\omega)\,\right)^{-1}B \\ & = C(\mathrm{i}\omega)\left(\Delta(\mathrm{i}\omega) + B\,Z\,C(\mathrm{i}\omega)\,\right)^{-1}B \\ & \overset{(5.75),(5.48)}{=} G^{\mathrm{I}}(\mathrm{i}\omega) \end{aligned} \tag{5.79}$$

(using the push-through identity $B(I+QB)^{-1} = (I+BQ)^{-1}B$ in the second line), which completes the proof. □

Remark 5.5.10 (Stability assumption). *In the following theorem, it is assumed that the equilibrium of the transformed system $\dot{x}(t) = f^{\mathrm{I}}(x_t)$ from (5.48) is exponentially stable. (Note that, in the case of a linear norm bound, the latter coincides with the nominal system, $f^{\mathrm{I}} = f$.) In view of Lemma 5.5.6, weaker conditions (stabilizability, hyperbolicity) suffice, but, in view of Section 5.4, the simpler assumption is desirable anyway: Since $f^{\mathrm{I}}(x_t) = f(x_t) - BK\mathcal{C}x_t$ in (5.48), it implies that $K = (-\Pi_{aa})^{-1}\Pi_{\zeta a}^{\top}$ stabilizes the nominal system. With K fulfilling (5.35), this ensures that V satisfies the partial definiteness properties from Section 5.4, which would otherwise not have to be the case.*

Altogether, the following existence criterion for an LK functional of robust type w.r.t. the nominal system $\dot{x}(t) = f(x_t) = A_0x(t) + A_1x(t-h)$, the perturbation structure $(B, \mathcal{C})$, and the perturbation restriction $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$ is obtained.

Theorem 5.5.11 (Existence condition). *Assume that $\dot{x}(t) = f^{\mathrm{I}}(x_t)$ defined in (5.48) has an exponentially stable equilibrium, and $\dot{x}(t) = f(x_t)$ does not have characteristic roots on the imaginary axis. Moreover, let $\Pi_{\zeta\zeta}^{\mathrm{I}}$ from (5.50a) have the block diagonal structure (5.51). Based on the transfer function (5.77), consider*

$$W_G(\mathrm{i}\omega) = -\begin{bmatrix} G(\mathrm{i}\omega) \\ -I_m \end{bmatrix}^H \begin{bmatrix} \Pi_{\zeta\zeta} & \Pi_{\zeta a} \\ \Pi_{\zeta a}^{\top} & \Pi_{aa} \end{bmatrix} \begin{bmatrix} G(\mathrm{i}\omega) \\ -I_m \end{bmatrix}. \tag{5.80}$$

If $W_G(\mathrm{i}\omega) \succ 0_{m\times m}$, for all $\omega \in \mathbb{R}$, then an LK functional of robust type exists. If $W_G(\mathrm{i}\omega) \not\succeq 0_{m\times m}$ for some $\omega \in \mathbb{R}$, then no LK functional of robust type exists.

Proof. Because of Theorem 5.5.4, solvability of (5.65) must be shown. Thus, due to the equivalence from Lemma 5.5.7, the existence question is tackled by the KYP lemma 5.5.6. Concerning the characteristic roots $\{\lambda_k\}_k$, it need not be distinguished between $\forall k : |\mathrm{Re}(\lambda_k)| \neq 0$ required above and $\exists \epsilon > 0, \forall k : |\mathrm{Re}(\lambda_k)| > \varepsilon$ required in Lemma 5.5.6 since eigenvalue chains in RFDEs satisfy

$\mathrm{Re}(\lambda_k) \to -\infty$ if $|\lambda_k| \to \infty$, see [82, Lem. 1-4.1] and [14, Thm. 12.12]. By Lemma 5.5.8 and Lemma 5.5.9, (5.74) depends on (5.80) according to

$$\mathscr{F}((\mathrm{i}\omega I_{M_2} - \mathscr{A}^{\mathrm{I}})^{-1}\mathscr{B}u, u) = u^H(I + ZG(\mathrm{i}\omega))^{-H}W_G(\mathrm{i}\omega)(I + ZG(\mathrm{i}\omega))^{-1}u. \tag{5.81}$$

Thus, the existence statement of Lemma 5.5.6 relies on

$$\alpha_3 = \inf_{\omega} \lambda_{\min}\big((I + ZG(\mathrm{i}\omega))^{-H}W_G(\mathrm{i}\omega)(I + ZG(\mathrm{i}\omega))^{-1}\big).$$

By (5.79) and the assumptions, $\det((I+ZG(\mathrm{i}\omega))^{-1}) \neq 0$ holds for all $\omega \in \mathbb{R}$, and thus, Sylvester's law of inertia applies. Since G is strictly proper, and due to (5.15), $\lim_{|\omega|\to\infty} \lambda_{\min}(W_G(\mathrm{i}\omega)) = \lambda_{\min}(-\Pi_{aa}) > 0$. Hence, positive definiteness of $W_G(\mathrm{i}\omega)$ is equivalent to $\alpha_3 > 0$. □

Remark 5.5.12 (Uniqueness). *If uniqueness of V is desired, the considerations can be restricted to the unique so-called stabilizing solution $\mathscr{P}_0$ of the algebraic Riccati equation (5.65), see Section 7.2.1.*

5.5.5 Admissible Perturbation Restrictions*

The limiting factor on the admissible bounds in Table 5.1 becomes Theorem 5.5.11, which is decisive for the existence of an LK functional of robust type. Theorem 5.5.11 is expressed in terms of $\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa}$. The present section states the resulting explicit bounds on the involved parameters, e.g., γ, ρ, K_1, from Table 5.1. Subsequently derived bounds are summarized in Table 5.2. The first one addresses the linear norm bound γ from Table 5.1, row (I|a).

* Chapter 5 is part of [S3].

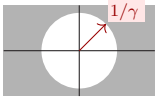
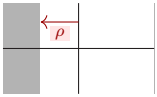
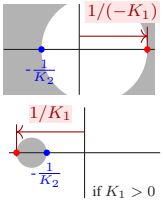
Row in Table 5.1	Bound on the respective parameter in $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$, based on the transfer function $G(s)$ of the nominal system from (5.77)	If $p = m = 1$: restriction on $G(\mathrm{i}\omega) \in \mathbb{C}$ in the Nyquist plot
(I|a)	$\gamma_{\max} = \dfrac{1}{\max_\omega \|G(\mathrm{i}\omega)\|_2} = \dfrac{1}{\|G\|_\infty}$ Reciprocal of the L_2-gain of the nominal system	
(II|a)	$\rho_{\min} = \max_\omega \mu_2(-G(\mathrm{i}\omega)) = -\nu(G)$ Negative of the input passivity index of the nominal system	
(III)	Assuming $K_2 = k_2 I_m$ is fixed and $K_1 = k_1 I_m$ variable: $k_{1,\min} = k_2 + \dfrac{1}{\nu(-G^{\mathrm{II}})}$ (see Corollary 5.5.17 for G^{II}), from a transformation to row (II)	

Table 5.2: Bounds on the admissible sectors in Table 5.1 according to Corollary 5.5.13, 5.5.14, 5.5.17, and Remark 5.5.15. If $p = m = 1$, then $\|G\|_\infty = \max_\omega |G(\mathrm{i}\omega)|$ and $\nu(G) = \min_\omega \mathrm{Re}(G(\mathrm{i}\omega))$, which corresponds to the gray-shaded Nyquist plot restrictions in the right column (depicted is the complex plane with the black lines being the real and imaginary axis). A less restricted Nyquist curve amounts to a less robust system, thus requiring a more restrictive perturbation sector. See Remark 5.5.10 for the role of a stability assumption on $\dot{x}(t) = f^{\mathrm{I}}(x_t)$ that is additionally imposed.

Corollary 5.5.13 (Maximum linear norm bound). *Assume the nominal system* $\dot{x}(t) = A_0x(t) + A_1x(t-h) =: f(x_t)$ *has an exponentially stable zero equilibrium. Let* $G(s)$ *be its transfer function (5.77) incorporating the perturbation structure* $(B, \mathcal{C})$*. If* $\gamma < \gamma_{\max}$

$$\gamma_{\max} := \frac{1}{\max_\omega \|G(\mathrm{i}\omega)\|_2} = \frac{1}{\|G\|_\infty}, \tag{5.82}$$

then a solution V *of (5.30) exists.*

Proof. Due to $f^{\mathrm{I}}(x_t) = f(x_t)$, the assumptions in Theorem 5.5.11 are satisfied. With $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$ from row (I|a) in Table 5.1, (5.80) becomes

$$W_G(\mathrm{i}\omega) = -\big(\gamma^2 (G(\mathrm{i}\omega))^H G(\mathrm{i}\omega) - I_m\big), \tag{5.83}$$

$$\lambda_{\min}\big(W_G(\mathrm{i}\omega)\big) = -\gamma^2 \lambda_{\max}\big((G(\mathrm{i}\omega))^H G(\mathrm{i}\omega)\big) + 1 > 0. \tag{5.84}$$

Moreover, the peak gain $\max_\omega \sqrt{\lambda_{\max}\big((G(\mathrm{i}\omega))^H G(\mathrm{i}\omega)\big)} = \max_\omega \|G(\mathrm{i}\omega)\|_2$ coincides with the H_∞-norm since $G \in H_\infty$ by the assumed exponential stability. □

Note that (5.82) coincides with the complex stability radius, cf. [159, example. 2.5]. Another way to read the above corollary is that the product of L_2-gains $\gamma\|G\|_\infty$ shall be smaller than one, mirroring the small-gain theorem cf. [53, Sec. III.2]. Similarly[5], for row (II|a) in Table 5.1, the following corollary mirrors a passivity theorem: the excess ρ of output passivity in the perturbation shall be larger than the shortage of input passivity in the nominal system. The latter is measured by the input passivity index $\nu(G) \leq 0$, cf. [203, eq. (7)].

[5] Both refer to interconnected dissipative elements [191] with $\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa}$ describing the QSR-dissipativity of $\zeta \mapsto a(\zeta)$. See also Section 6.3.2.

Corollary 5.5.14 (Minimum output passivity). *Let $G(s)$ be the transfer function (5.77) of the nominal system $\dot{x}(t) = A_0x(t) + A_1x(t-h)$ with $(B, \mathcal{C})$. Consider $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$ from row* (II|a) *in Table 5.1 with $\rho > \rho_{\min}$,*

$$\rho_{\min} := \max_{\omega} \mu_2(-G(\mathrm{i}\omega)) = -\nu(G), \tag{5.85}$$

where $\mu_2(M) = \lambda_{\max}(\frac{1}{2}(M^H + M))$ describes the logarithmic norm of a given matrix $M \in \mathbb{C}^{p\times p}$. Moreover, assume that $\dot{x}(t) = A_0x(t) + A_1x(t-h) - \frac{1}{2\rho}B\begin{bmatrix} C_1x(t-h) \\ C_0x(t)\end{bmatrix} =: f^{\mathrm{I}}(x_t)$ has an exponentially stable zero equilibrium and the nominal system has no characteristic roots on the imaginary axis. Then a solution V of (5.17) exists.

Proof. Consider (5.80) with Table 5.1, row (II|a),

$$W_G(\mathrm{i}\omega) = -\Big(-\tfrac{1}{2}\big((G(\mathrm{i}\omega))^H + G(\mathrm{i}\omega)\big) - \rho I\Big), \tag{5.86}$$

$$\lambda_{\min}\big(W_G(\mathrm{i}\omega)\big) = -\lambda_{\max}\big(\mathrm{He}(-G(\mathrm{i}\omega))\big) + \rho > 0. \tag{5.87}$$

□

Remark 5.5.15 (Circle criterion). *If $p = m = 1$, a Nyquist plot of $G(\mathrm{i}\omega)$ can be considered. In terms of the thus relevant real and imaginary parts of $G(\mathrm{i}\omega)$, (5.80) becomes*

$$W_G(\mathrm{i}\omega) = -\begin{bmatrix} \mathrm{Re}(G(\mathrm{i}\omega)) \\ -I_m \\ \mathrm{Im}(G(\mathrm{i}\omega))\end{bmatrix}^{\top}\begin{bmatrix} \Pi_{\zeta\zeta} & \Pi_{\zeta a} & 0_{p\times m} \\ \Pi_{\zeta a}^{\top} & \Pi_{aa} & 0_{m\times m} \\ 0_{m\times p} & 0_{p\times m} & \Pi_{\zeta\zeta}\end{bmatrix}\begin{bmatrix} \mathrm{Re}(G(\mathrm{i}\omega)) \\ -I_m \\ \mathrm{Im}(G(\mathrm{i}\omega))\end{bmatrix}. \tag{5.88}$$

The general sector perturbation restriction from row (III) in Table 5.1 results in the Nyquist plot restriction that is known from the circle criterion, cf. [108,

Sec. 7.1.1]. To see the circle in (5.88), note that given some radius $r > 0$ and some shift $x_\delta \in \mathbb{R}$,

$$\mp\big((x-x_\delta)^2+y^2-r^2\big) = \mp \begin{bmatrix} x \\ -1 \\ y \end{bmatrix}^\top \begin{bmatrix} 1 & x_\delta & 0 \\ x_\delta & x_\delta^2-r^2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ -1 \\ y \end{bmatrix} \geq 0 \tag{5.89}$$

describes a disc $(-)$ or the complement of a disc $(+)$ in the (x,y) plane. Analogously, if $p = m = 1$, an open disc or the interior of its complement in the $(\mathrm{Re}(G(\mathrm{i}\omega)), \mathrm{Im}(G(\mathrm{i}\omega)))$ plane is described by $W_G(\mathrm{i}\omega) > 0$ from (5.88) with Table 5.1, row (III|c). See the plots in Table 5.2.

Rather than plotting the Nyquist curve, transformations that eliminate either $\Pi_{\zeta a}$ (transformation I, Lemma 5.5.1) or $\Pi_{\zeta\zeta}$ (e.g., transformation II below) are preferable as these yield numerically traceable results in the manner of Corollary 5.5.13 or Corollary 5.5.14, not restricted to $p = m = 1$.

Lemma 5.5.16 (Transformation II). *$V(\phi)$ is an LK functional of robust type w.r.t. $\dot{x}(t) = A_0 x(t) + A_1 x(t)$, $(B,\mathcal{C})$, and $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$ from row (III) in Table 5.1 with $K_1 = k_1 I_m$ and $K_2 = k_2 I_m$, $k_2 > k_1 \in \mathbb{R}$, i.e.,*

$$\Pi_{\zeta\zeta} = -k_1 k_2 I_m, \quad 2\Pi_{\zeta a} = (k_1+k_2) I_m, \quad \Pi_{aa} = -I_m,$$

if and only if $V(\phi) = V^{\mathrm{II}}(\phi)$ is an LK functional of robust type w.r.t. $\dot{x}(t) = A_0^{\mathrm{II}} x(t) + A_1^{\mathrm{II}} x(t) =: f^{\mathrm{II}}(x_t)$,

$$A_0^{\mathrm{II}} = A_0 - k_2 B \begin{bmatrix} 0_{p_1 \times n} \\ C_0 \end{bmatrix}, \quad A_1^{\mathrm{II}} = A_1 - k_2 B \begin{bmatrix} C_1 \\ 0_{p_0 \times n} \end{bmatrix},$$

the original $(B,\mathcal{C})$, and the transformed

$$\Pi_{\zeta\zeta}^{\mathrm{II}} = 0_{m\times m}, \quad 2\Pi_{\zeta a}^{\mathrm{II}} = -(k_2-k_1) I_m, \quad \Pi_{aa}^{\mathrm{II}} = -I_m.$$

Proof. The defining equation (5.17) is not altered if instead of $D_f^+V(\phi)$ and $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$, the result for $D_{(f+g)}^+V(\phi)$ from Lemma 5.3.1 with $g(x_t) = -k_2B\mathcal{C}x_t$ and $(\Pi_{\zeta\zeta}^{\mathrm{II}}, \Pi_{\zeta a}^{\mathrm{II}}, \Pi_{aa}^{\mathrm{II}})$ are used. □

The final corollary is expedient for saturation nonlinearities, where the upper sector bound is usually fixed, and the best possible lower sector bound is of interest.

Corollary 5.5.17 (Minimum lower sector bound). *Consider Table 5.1* (III) *with* $K_2 = k_2I_m$. *Let* $K_1 = k_1I_m$ *and* $k_1 > k_{1,\min}$,

$$k_{1,\min} := k_2 - \frac{1}{\max_\omega \mu_2(G^{\mathrm{II}}(\mathrm{i}\omega))} = k_2 + \frac{1}{\nu(-G^{\mathrm{II}})},$$

where $G^{\mathrm{II}}(s) = \left[\begin{smallmatrix} C_1\mathrm{e}^{-sh} \\ C_0 \end{smallmatrix}\right](sI - A_0^{\mathrm{II}} - \mathrm{e}^{-sh}A_1^{\mathrm{II}})^{-1}B$ *and where* A_0^{II} *and* A_1^{II} *are defined in Lemma 5.5.16. Assume* $\dot{x}(t) = A_0x(t) + A_1x(t-h) - \frac{k_1+k_2}{2}B\left[\begin{smallmatrix} C_1x(t-h) \\ C_0x(t) \end{smallmatrix}\right] =: f^{\mathrm{I}}(x_t)$ *has an exponentially stable zero equilibrium and* $\dot{x}(t) = A_0^{\mathrm{II}}x(t) + A_1^{\mathrm{II}}x(t) =: f^{\mathrm{II}}(x_t)$ *has no characteristic roots on the imaginary axis. Then a solution* V *of (5.17) exists.*

Proof. With $(\Pi_{\zeta\zeta}^{\mathrm{II}}, \Pi_{\zeta a}^{\mathrm{II}}, \Pi_{aa}^{\mathrm{II}})$ from Lemma 5.5.16, (5.80) is

$$W_{G^{\mathrm{II}}}(\mathrm{i}\omega) = -\Big(\tfrac{1}{2}(k_2 - k_1)\big((G^{\mathrm{II}}(\mathrm{i}\omega))^H + G^{\mathrm{II}}(\mathrm{i}\omega)\big) - I_m\Big), \tag{5.90}$$

$$\lambda_{\min}\big(W_{G^{\mathrm{II}}}(\mathrm{i}\omega)\big) = -(k_2 - k_1)\lambda_{\max}(\mathrm{He}(G^{\mathrm{II}}(\mathrm{i}\omega))) + 1 > 0. \tag{5.91}$$

□

To sum up, the admissible perturbation restrictions for which an LK functional of robust type exists can be derived from the frequency response of the nominal system incorporating the input matrix and output operator that describe the perturbation structure. As these results rely on the Kalman–Yakubovich–Popov lemma, the resulting bounds, in fact, are familiar bounds from frequency domain

considerations (see also Section 6.3). For the special case of $p = m = 1$, the Nyquist curve could be considered, where, roughly speaking, a more restricted Nyquist curve amounts to a more robust system thus calling for a less restrictive sector bound on perturbations (as known from the circle criterion). However, the given formulas are not restricted to $p = m = 1$, which would be presumed in a graphical evaluation. For instance, for the linear norm bound γ, it is concluded that an LK functional of robust type exists whenever γ is smaller than the reciprocal of the H_∞-norm of the nominal system (which is the bound known from the complex stability radius or the small-gain theorem, see also Section 6.3).

5.6 Example*

The following example system (5.92) supplements in [111] the introduction of complete-type LK functionals.

Example 5.6.1. *With a vanishing perturbation $g(x_t) \equiv 0_n$, the system*

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix} x(t) + \begin{bmatrix} 0 & 0 \\ -1 & 1 \end{bmatrix} x(t-h) + g(x_t) \tag{5.92}$$

can be shown to have an exponentially stable equilibrium for any $h > 0$ (see Chapter 2). Henceforth, let $h = 1$.

(i) *Table 5.3 gives a linear norm bound on unstructured perturbations*

$$g(x_t) = -a\left(\begin{bmatrix} x(t-h) \\ x(t) \end{bmatrix}\right) \tag{5.93}$$

* Chapter 5 is part of [S3].

(i) Admissible linear norm bound on the perturbation
$\|g(x_t)\|_2 \le \tilde{\gamma} \left\| \begin{bmatrix} x(t) \\ x(t-h) \end{bmatrix} \right\|_2$ with $\tilde{\gamma} < \gamma_{\max}$

LK functional of complete type, (1.17)	$\gamma_{\max} = 0.0227$
LK functional of robust type, (5.82)	
(a) unstructured, $B = C_0 = C_1 = I$	$\gamma_{\max} = 0.1059$
(b) if $g(x_t) = \begin{bmatrix} 0 \\ g_2(x_t) \end{bmatrix}$: structured, $B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, C_0 = C_1 = I$	$\gamma_{\max}^{g_1 \equiv 0} = 0.2462$

(ii) Special case of bounds on uncertainties $\Delta_{0,1} \in \mathbb{R}^{n \times n}$
in $\dot{x}(t) = (A_0 + \Delta_0)x(t) + (A_1 + \Delta_1)x(t-h)$

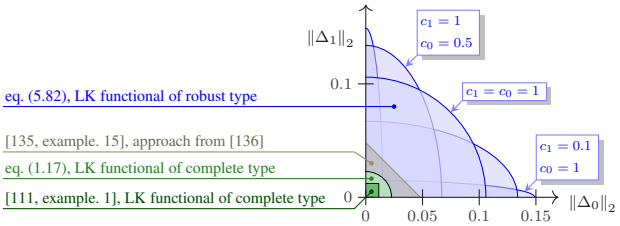


Table 5.3: Example 5.6.1. The bounds on admissible perturbations are significantly less restrictive than those from complete-type (choosing (1.14) as in [111, Example 1]) and related LK functionals.

and on perturbations

$$g(x_t) = \begin{bmatrix} 0 \\ g_2(x_t) \end{bmatrix} = - \begin{bmatrix} 0 \\ 1 \end{bmatrix} a \left(\begin{bmatrix} x(t-h) \\ x(t) \end{bmatrix} \right) \tag{5.94}$$

that only affect the second component, cf. (5.7). The latter are quite plausible if (5.92) represents the state space description of a second order system for x_1. According to (5.82), $\gamma_{\max} = \frac{1}{\|G\|_\infty}$. See, e.g., [8],[138] for a numerical implementation of

$$\|G\|_\infty = \max_{\omega \in \mathbb{R}} \left\| \begin{bmatrix} C_1 \mathrm{e}^{-\mathrm{i}\omega h} \\ C_0 \end{bmatrix} (\mathrm{i}\omega I_n - A_0 - A_1 \mathrm{e}^{-\mathrm{i}\omega h})^{-1} B \right\|_2, \tag{5.95}$$

and see also Section 7.3. Corollary 5.5.13 ensures that a solution $V(\phi)$ of (5.30) exists if $\gamma < \gamma_{\max}$. Due to the full-rank choice $C_0 = C_1 = I$, such an LK functional of robust type $V(\phi)$ satisfies the conditions of the classical LK theorem (Theorem 4.5.1):

- *monotonicity by Corollary 5.3.4 ($k_2 = \gamma^2 - \tilde{\gamma}^2$ from Table 5.3 with $\tilde{\gamma} < \gamma < \gamma_{\max}$, where $\tilde{\gamma}, \gamma$ are arbitrarily close to $\gamma_{\max}$),*
- *partial positive definiteness by Theorem 5.4.3 (with $K = 0_{n\times n}$ in (5.35), (5.36)), and boundedness by Lemma 5.5.5.*

Thus, for the family of accordingly perturbed systems, $V(\phi)$ is a common LK functional that proves global asymptotic stability of the zero equilibrium.

(ii) *Uncertainties $\Delta_0, \Delta_1 \in \mathbb{R}^{n\times n}$ in the coefficient matrices of (5.92) amount to $g(x_t) = \Delta_0 x(t) + \Delta_1 x(t-h)$ or*

$$g(x_t) = \begin{bmatrix} \frac{1}{c_1}\Delta_1 & \frac{1}{c_0}\Delta_0 \end{bmatrix} \begin{bmatrix} c_1 x(t-h) \\ c_0 x(t) \end{bmatrix} =: -a\left(\begin{bmatrix} c_1 x(t-h) \\ c_0 x(t) \end{bmatrix} \right), \tag{5.96}$$

for any $c_0, c_1 > 0$. Since

$$\|a(\zeta)\|_2 \le r(\Delta_0, \Delta_1)\|\zeta\|_2, \qquad r(\Delta_0, \Delta_1) := \sqrt{\tfrac{1}{c_1^2}\|\Delta_1\|_2^2 + \tfrac{1}{c_0^2}\|\Delta_0\|_2^2},$$

the linear norm bound is satisfied if $r(\Delta_0, \Delta_1) < \gamma_{\max}$ with $\gamma_{\max}$ from (5.82), choosing[6] $B = I, C_0 = c_0 I, C_1 = c_1 I$. See Table 5.3.

5.7 Revisiting the Main Points of the Chapter

- The chapter introduces the concept of LK functionals of robust type.

[6] equivalent to $B = C_0 = C_1 = I_n$ combined with Table 5.1, row (I|b), $L = \left[\begin{smallmatrix} c_1 I_n & 0 \\ 0 & c_0 I_n \end{smallmatrix}\right]$, $W = I_n$.

- Like LK functionals of complete type, LK functionals of robust type are also defined via their derivative along solutions of the linear nominal system. However, in contrast to complete-type and related LK functionals defined by (1.14), the prescribed derivative does not rely on three arbitrarily chosen matrices Q_0, Q_1, Q_2. Instead, the defining equation (5.17) relies on the following.

  1. The perturbation structure $(B, \mathcal{C})$: It gives rise to the Lur'e description $g(x_t) = -Ba(\mathcal{C}x_t)$ of the perturbation in (1.2), where

     - the input matrix B encodes which components of the RFDE are affected by the perturbation, and
     - the output operator $\mathcal{C}x_t = \left[\begin{smallmatrix} C_1\, x(t-h) \\ C_0\, x(t) \end{smallmatrix}\right]$ encodes what g relies upon (but C_0, C_1 can always be chosen as full-rank matrices by adding some εI if that is required for further properties), see Example 5.2.1.

  2. The perturbation restriction matrices $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$ that characterize the type of robustness bound that shall be obtained for the function $\zeta \mapsto \alpha = a(\zeta)$ in $g(x_t) = -Ba(\mathcal{C}x_t)$:

     - The three matrices $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$ define a sector in terms of the pairs (ζ, α) for which $w(\zeta, \alpha) = \left[\begin{smallmatrix} \zeta \\ \alpha \end{smallmatrix}\right]^\top \left[\begin{smallmatrix} \Pi_{\zeta\zeta} & \Pi_{\zeta a} \\ \Pi_{\zeta a}^\top & \Pi_{aa} \end{smallmatrix}\right] \left[\begin{smallmatrix} \zeta \\ \alpha \end{smallmatrix}\right] \geq 0$ is satisfied. Possible choices of $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$ and the corresponding sectors are summarized in Table 5.1.
     - The choice of $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$ can be tailored to the given nonlinear function $\zeta \mapsto \alpha = a(\zeta)$, see Example 5.2.2. If, in further steps, only local stability results shall be obtained, the graph $(\zeta, a(\zeta))$ need not even belong globally to the described sector. It suffices if the function resides locally (on a preferably large domain of ζ) within the sector.

  - Usually a parameter like γ (respectively ρ or k_1) is involved in $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$ that controls the size of the sector.
  - How large (respectively small) that parameter can be chosen in order not to hamper the existence of the LK functional of robust type only depends on the robustness of the nominal system for the given perturbation structure.
  - The corresponding admissible parameter bounds are summarized in Table 5.2. See Corollary 5.5.13 et seq. for the precise statements. Their derivations are based on a version of the Kalman–Yakubovich–Popov lemma for C_0-semigroups combined with a splitting approach.

3. The expression v defined in (5.6): It depends on terms of the functional itself, making the overall defining equation implicit. However, in the ODE-based numerical approach, which will be discussed in the next chapter, only the Lyapunov equation that has been encountered in Chapter 4 is replaced by an algebraic Riccati equation.

- The sense behind the proposed construction of the defining equation (5.17) is as follows.
  - It is achieved that the derivative of the LK functional $V(\phi)$ along solutions of the perturbed RFDE (1.2) is nonpositive wherever the argument ϕ is such that $\zeta = \mathcal{C}\phi$ with $\alpha = a(\mathcal{C}\phi)$ is compatible with the perturbation restriction. In other words, $w(\mathcal{C}\phi, a(\mathcal{C}\phi)) \geq 0$ implies $D^+_{(f+g)}V(\phi) \leq 0$.
  - Even more, any strengthening to $D^+_{(f+g)}V(\phi) \leq -\ell(\mathcal{C}\phi)$ with some desired offset function $\ell(\zeta)$ can be established by a corresponding strengthening $w(\mathcal{C}\phi, a(\mathcal{C}\phi)) \geq \ell(\mathcal{C}\phi)$, indicated by the turquoise shaded reduction of the sectors in the plots of Table 5.1 (Theorem 5.3.2).

  - In the special case of a linear-norm-bound perturbation restriction (where $\gamma < \gamma_{\max} = \frac{1}{\|G\|_\infty}$ ensures that a corresponding LK functional of robust type exists), it can be concluded that a more restrictive linear norm bound $\|a(\mathcal{C}\phi)\|_2 \leq \sqrt{\gamma^2 - k_3}\|\mathcal{C}\phi\|_2$ with some $k_3 \geq 0$ already implies $D^+_{(f+g)}V(\phi) \leq -k_3\|\mathcal{C}\phi\|_2^2$ (Corollary 5.3.4).

- Comparability with the known robustness statement from complete-type LK functionals, see (1.17) in the introduction, is established by choosing

  1. the perturbation structure $B = C_0 = C_1 = I_n$, in which case the nonlinearity $a(\cdot)$ covers the overall unstructured perturbation $g(x_t) = -Ba\big(\left[\begin{smallmatrix} C_1\, x(t-h) \\ C_0\, x(t) \end{smallmatrix}\right]\big) = -a\big(\left[\begin{smallmatrix} x(t-h) \\ x(t) \end{smallmatrix}\right]\big)$ and
  2. the perturbation restriction given by row (I|a) in Table 5.1, addressing a linear norm bound $\|a(\mathcal{C}x_t)\|_2 \leq \gamma\|\mathcal{C}x_t\|_2$. Note that, the defining equation (5.17) then simplifies to (5.18).

  For this choice of perturbation structure and perturbation restriction, a robustness bound having the same form as (1.17) is derived, and, even more, the following holds.

  - The concept shares what is commonly considered as the main advantage of complete-type LK functionals: A corresponding LK functional of robust type (which is guaranteed to exist whenever $\gamma < \gamma_{\max} = \frac{1}{\|G\|_\infty}$ is chosen) satisfies the conditions imposed by the classical LK theorem if and only if the equilibrium of the nominal system is exponentially stable. To be more precise, it then satisfies

    * the positive definiteness condition (4.28) by Theorem 5.4.3 and
    * the monotonicity condition (4.29) by Corollary 5.3.4.

    Thus, the approach is expedient even if the delay is arbitrarily close to a critical delay value at which the exponential stability is lost.

  - At the same time, the resulting linear norm bound γ on admissible perturbations, which according to Corollary 5.5.13 can be chosen as $\gamma < \gamma_{\max} = \frac{1}{\|G\|_\infty}$, is significantly less conservative compared to (1.18) from complete-type LK functionals, see Example 5.6.1.

- By offering additional freedom in both incorporating the structure of the perturbation and imposing a perturbation restriction in form of an arbitrary sector, the concept is more adaptable to the problem at hand, which is usually rewarded by an additional reduction of conservatism.

6 LK Functionals of Robust Type: Numerical Approach and ODE-Based Explanation

The numerical approach from Chapter 4 can also be applied to the proposed defining equation (5.17), thus providing a numerical approximation of LK functionals of robust type. Besides of the practical relevance, e.g., for the evaluation of the functional with a given argument or for bounds on the functional, the ODE-based approach also establishes a link to the simpler finite-dimensional setting.

The chapter is organized as follows. In **Section 6.1**, the defining equation of LK functionals of robust type is discretized based on the ODE approximation from the Legendre tau method. As a result, a finite-dimensional algebraic Riccati equation (ARE) takes the place of the finite-dimensional Lyapunov equation encountered in Chapter 4. **Section 6.2** explains the role of the ARE for the approximating ODE. Conditions on the solvability of the finite-dimensional ARE are provided in **Section 6.3** by means of the Kalman–Yakubovich–Popov (KYP) lemma. The connections to some related frequency-domain concepts are also explained. **Section 6.4** revisits the most important aspects of the chapter.

To shortly present the numerical approach that is proposed in Section 6.1 is a subject of [S4].

6.1 Numerical Approach via a Finite-Dimensional Algebraic Riccati Equation*

Let us briefly review the essence of the numerical approach that has already been applied to complete-type and related LK functionals in Section 4.1. The state function $x_t = \phi \in C([-h, 0], \mathbb{R}^n)$ is mapped by the LK functional V to a scalar value $V(x_t) \in \mathbb{R}$. In the numerical approach, instead, a polynomial approximation of the state function shall be mapped to the approximately same scalar value. The polynomial is uniquely represented by its coefficients $\tilde{x}$ (e.g. Legendre coordinates $\tilde{x} = c$, mixed coordinates $\tilde{x} = \chi$, or interpolation coordinates $\tilde{x} = y$). As a consequence, the described approximation of the quadratic LK functional can be established by a quadratic form

$$V(\phi) \approx \tilde{V}(\tilde{x}) = \tilde{x}^\top \tilde{P} \tilde{x} \tag{6.1}$$

in the polynomial coordinates $\tilde{x}$. The present chapter uses the placeholder

$$\tilde{x} \in \{c, \chi, y, \ldots\}. \tag{6.2}$$

to emphasize the independence of the chosen coordinates (in contrast to Chapter 4, the decomposition $y = [z^\top, \hat{x}^\top]^\top$ will not further be employed). Moreover, the introduced tilde for the corresponding matrices, e.g., $\tilde{P} = P_c$, $\tilde{P} = P_\chi$, or $\tilde{P} = P_y$, intends to contribute to a less cluttered notation in the matrix equations.

In the above described ansatz, only the matrix $\tilde{P}$ in (6.1) must be determined. Consider the defining equation of the LK functional (5.17). It relies on the derivative of $V(x_t)$ along the solution of the nominal linear time-delay system, i.e., the directional derivative in the direction in which the state x_t evolves. Therefore, in view of the approximation (6.1), it must be clarified how the coordinates $\tilde{x}(t)$ evolve with time t. Due to the convincing results from Chapter 4, the Legendre

* The contributions from Section 6.1 are considered in [S4].

tau method is used for this task. The result is an ODE $\dot{\tilde{x}} = \tilde{A}\tilde{x}$, see Chapter 3. Employing the latter, the left-hand side of the defining equation (5.17) becomes handleable in terms of matrices

$$D_f^+ V(\phi) \approx D^+_{(\dot{\tilde{x}}=\tilde{A}\tilde{x})} \tilde{x}^\top \tilde{P}\tilde{x} = \tilde{x}^\top \tilde{P}\tilde{A}\tilde{x} + \tilde{x}^\top \tilde{A}^\top \tilde{P}\tilde{x}.$$

In contrast to the defining equation of complete-type and related LK functionals from (1.14), which is tackled in Section 4.1, the right-hand side of the defining equation (5.17) involves a term that quadratically depends on the unknown functional. That is why, instead of the Lyapunov equation encountered in Section 4.1, an algebraic Riccati equation will arise in (6.18) for $\tilde{P}$.

6.1.1 Legendre-Tau-Based Approximation of a System with In- and Output

Relevant for the derivative $D_f^+ V$ on the left-hand side of the defining equation (5.17) is only the linear nominal system, which is still the autonomous RFDE $\dot{x}(t) = f(x_t) = A_0 x(t) + A_1 x(t-h)$. However, to cover the perturbation structure $(B, \mathcal{C})$ introduced in Section 5.2, it is of interest how appended in- and output terms in

$$\dot{x}(t) = A_0 x(t) + A_1 x(t-h) + Bu(t) \tag{6.3a}$$

$$\zeta(t) = \mathcal{C}x_t = \begin{bmatrix} C_1 x(t-h) \\ C_0 x(t) \end{bmatrix} \tag{6.3b}$$

are tackled by the Legendre tau method. Note that a combination with $u(t) = -a(\zeta(t))$ realizes the perturbed system

$$\dot{x}(t) = A_0 x(t) + A_1 x(t-h) - Ba(\left[\begin{smallmatrix} C_1 x(t-h) \\ C_0 x(t) \end{smallmatrix}\right]) \tag{6.4}$$

from (1.2) and (5.7).

The result of the Legendre tau method (cf. Chapter 3) for the approximated dynamics of the state x_t from (6.3a) becomes in terms of Legendre coordinates $\tilde{x} = c$

$$\dot{c}(t) = A_c c(t) + B_c u(t) \tag{6.5a}$$

$$\zeta(t) = C_c c(t), \tag{6.5b}$$

where the matrices $A_c \in \mathbb{R}^{n(N+1)\times n(N+1)}, B_c \in \mathbb{R}^{n(N+1)\times m}$, and $C_c \in \mathbb{R}^{(p_1+p_0)\times n(N+1)}$ are as follows.

A_c: Recap from (3.27) that

$$\begin{aligned} A_c = D_c \otimes I_n + &\begin{bmatrix} 0_{n\times n} & 0_{n\times n} & 0_{n\times n} & \cdots & 0_{n\times n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0_{n\times n} & 0_{n\times n} & 0_{n\times n} & \cdots & 0_{n\times n} \\ A_0{+}A_1 & A_0{-}A_1 & A_0{+}A_1 & \cdots & A_0{+}(-1)^N A_1 \end{bmatrix} \\ &+ \tfrac{2}{h} \begin{bmatrix} 0 & & & & & \cdots & 0 \\ \vdots & & & & & & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & -1 & -3 & -6 & -10 & \cdots & -\frac{N(N+1)}{2} \end{bmatrix} \otimes I_n \end{aligned} \tag{6.6}$$

$$\text{with } D_c = \frac{2}{h} \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & \dots & 0 \\ 0 & 0 & 3 & 0 & 3 & 0 & 3 & 0 & \dots & 3 \\ 0 & 0 & 0 & 5 & 0 & 5 & 0 & 5 & \dots & 0 \\ 0 & 0 & 0 & 0 & 7 & 0 & 7 & 0 & \dots & 7 \\ 0 & 0 & 0 & 0 & 0 & 9 & 0 & 9 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 11 & 0 & \dots & 11 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 13 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 15 \\ & & & & & & & & \ddots & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix} \tag{6.7}$$

(exemplarily, the differentiation matrix D_c is shown for N even, otherwise a last column $[1, 0, 5, 0, 9, \dots, 0, (2N-1), 0]^\top$ has to be appended).

B_c: The RFDE right-hand side only affects the last block row in (6.5a) when it is built via the Legendre tau method, cf. (3.38). Accordingly, only the lower row of B_c is nonzero

$$B_c = \begin{bmatrix} 0_{nN\times m} \\ B \end{bmatrix}. \tag{6.8}$$

C_c: Along the lines of Section 3.5.3, let $\mathcal{C}\phi \approx \mathcal{C}\phi^{[N]} =: C_c c$, where $\phi^{[N]}$ is defined in (3.46). Then, by $\phi^{[N]}(0) = \sum_{k=0}^{N} c^k$ and $\phi^{[N]}(-h) = \sum_{k=0}^{N} (-1)^k c^k$,

$$C_c = \begin{bmatrix} C_1 & -C_1 & C_1 & \cdots & (-1)^N C_1 \\ C_0 & C_0 & C_0 & \cdots & C_0 \end{bmatrix}. \tag{6.9}$$

6.1.2 Change of Basis

As mentioned above, (6.5) is henceforth replaced by

$$\dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t) + \tilde{B}u(t) \tag{6.10a}$$

$$\zeta(t) = \tilde{C}\tilde{x}(t), \tag{6.10b}$$

where $\tilde{x} \in \mathbb{R}^{\tilde{n}}$ stands for the chosen coordinates, and $\tilde{n} := n(N+1)$. All coordinate choices are related by a simple change of basis, see Section 3.6. Changing (A_c, B_c, C_c), which describes the dynamics of Legendre coordinates, to $(\tilde{A}, \tilde{B}, \tilde{C})$, which describes the dynamics of possibly different coordinates $\tilde{x}$, with $c = T_{c\tilde{x}}\tilde{x}$ and $T_{\tilde{x}c} = T_{c\tilde{x}}^{-1}$, amounts to

$$\tilde{A} = T_{\tilde{x}c} A_c T_{c\tilde{x}}, \qquad \tilde{B} = T_{\tilde{x}c} B_c, \qquad \tilde{C} = C_c T_{c\tilde{x}}. \tag{6.11}$$

A change of coordinates for the matrix $\tilde{P} \in \mathbb{R}^{n(N+1)\times n(N+1)}$ in (6.1) is accordingly mirrored by the matrix congruence

$$\tilde{P} = T_{c\tilde{x}}^{\top} P_c T_{c\tilde{x}}, \tag{6.12}$$

see (4.7).

The following aspects are not yet needed for the discretization of the defining equation, but will become relevant in later parts of the chapter.

First, it should be noted that the transfer function

$$\tilde{G}(s) = C_c(sI_{\tilde{n}} - A_c)^{-1}B_c = \tilde{C}(sI_{\tilde{n}} - \tilde{A})^{-1}\tilde{B} \tag{6.13}$$

remains unaltered no matter which coordinates are used (a change of basis only alters the state space realization (6.10) of $\tilde{G}(s)$). The tilde in the notation is still required to distinguish between the rational transfer function $\tilde{G}(s) \in (\mathbb{R}(s))^{p\times m}$ in (6.13) based on the ODE approximation and the original RFDE transfer function $G(s)$ from (5.77).

Second, it should be noted that the nonlinear or uncertain function $a\colon \mathbb{R}^p \to \mathbb{R}^m$ from (5.9) is already a mapping between the finite-dimensional output space $\mathbb{R}^p$ and the finite-dimensional input space $\mathbb{R}^m$, which remain unaltered. Combining $u(t) = -a(\zeta(t))$ with (6.10) amounts again to the approximation of the state dynamics from the perturbed system (6.4)

$$\dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t) - \tilde{B}a(\tilde{C}\tilde{x}(t)). \tag{6.14}$$

6.1.3 The Finite-Dimensional Approximation of the Defining Equation*

Altogether, with

$$D_f^+ V(\phi) \approx D^+_{(\dot{\tilde{x}}=\tilde{A}\tilde{x})}\tilde{V}(\tilde{x}) = \tilde{x}^\top(\tilde{P}\tilde{A} + \tilde{A}^\top\tilde{P})\tilde{x}, \tag{6.15}$$

$$v^\top(\phi)\, B \approx \tilde{x}^\top\tilde{P}\tilde{B}, \tag{6.16}$$

$$\mathcal{C}\phi \approx \tilde{C}\tilde{x}, \tag{6.17}$$

* The presented approach from Section 6.1.3 and the involved example are part of [S4].

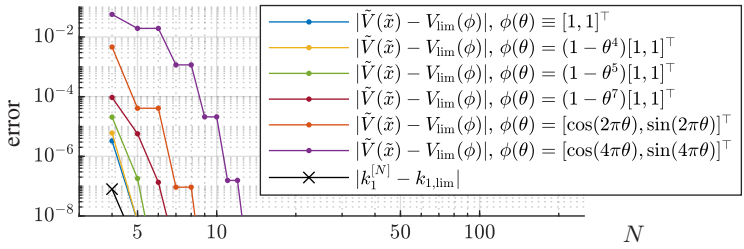


Figure 6.1: Convergence of the numerical approach in Example 6.1.1: The upper lines show the error between $\tilde{V}(\tilde{x}) = \tilde{x}^\top \tilde{P} \tilde{x}$ and its converged value $V_{\mathrm{lim}}(\phi)$ as N increases. Some exemplary polynomial and nonpolynomial arguments ϕ are considered. The lower line shows the convergence of the quadratic-lower-bound coefficient from (4.23).

the discretization of the defining equation (5.17) reads

$$\tilde{P}\tilde{A} + \tilde{A}^\top \tilde{P} = -\tilde{C}^\top \Pi_{\zeta\zeta} \tilde{C} - \Big[\tilde{P}\tilde{B} - \tilde{C}^\top \Pi_{\zeta a}\Big](-\Pi_{aa})^{-1}\Big[\tilde{B}^\top \tilde{P} - \Pi_{\zeta a}^\top \tilde{C}\Big] - \tilde{E}, \tag{6.18}$$

which is an algebraic Riccati equation for $\tilde{P}$ if $\tilde{E} = 0_{\tilde{n}\times\tilde{n}}$, or an algebraic Riccati inequality if the not further specified $\tilde{E} \succeq 0_{\tilde{n}\times\tilde{n}}$ allows for some arbitrary semidefinite discrepancy between the left- and right-hand side. In which sense (6.18) is the coordinate representation of an operator-valued equation is discussed in Appendix A.

The proposed numerical approach to LK functionals of robust type only requires to solve the matrix-valued algebraic Riccati equation (6.18) with $\tilde{E} = 0_{\tilde{n}\times\tilde{n}}$. To be more precise, $\tilde{V}(\tilde{x}) = \tilde{x}^\top \tilde{P} \tilde{x}$ is built from the unique so-called stabilizing (cf. Section 7.2.1) real symmetric matrix $\tilde{P}$ that solves (6.18) with $\tilde{E} = 0_{\tilde{n}\times\tilde{n}}$. This matrix $\tilde{P}$ is, e.g., numerically derivable via the standard routine `icare` for algebraic Riccati equations in Matlab. Note that, for a given $\phi \in C$, the vector $\tilde{x} = T_{\tilde{x}c} c$ is determined according to the discretization (3.47) on which the Legendre tau method relies. The resulting $\tilde{V}(\tilde{x}) = \tilde{x}^\top \tilde{P} \tilde{x} \approx V(\phi)$ is the searched approximation of the corresponding LK functional of robust type. A convergence statement for any $\phi \in C$ will be proven in Chapter 7.

Example 6.1.1. *Figure 6.1 shows for Example 5.6.1 the convergence of the above described approach with some exemplary arguments ϕ. To be more precise, Example 5.6.1 with the perturbation structure from Table 5.3 (i-b), i.e., $B = [0,1]^\top, C_0 = C_1 = I_2$, is considered. For the transfer function (6.13) from the Legendre tau method, $\tilde{\gamma}_{\max} = \frac{1}{\|\tilde{G}\|_\infty}$ converges rapidly to $\gamma_{\max} := \gamma_{\max}^{g_1 \equiv 0}$ from Table 5.3 as N increases, with $|\tilde{\gamma}_{\max} - \gamma_{\max}| < 10^{-8}$ for all tested $N \geq 4$. The chosen value is $\gamma = (1 - 10^{-5})\gamma_{\max}$. With the linear-norm-bound perturbation restriction from Table 5.1, the algebraic Riccati equation (6.18) simplifies to*

$$\tilde{P}\tilde{A} + \tilde{A}^\top \tilde{P} = -\gamma^2 \tilde{C}^\top \tilde{C} - \tilde{P}\tilde{B}\tilde{B}^\top \tilde{P}, \tag{6.19}$$

which, using mixed coordinates, is solved via `icare` *in Matlab. Note that the results in Figure 6.1 also account for the error in the underlying polynomial approximation of the argument ϕ, which would as well be relevant in a numerical integration of some known integral formula for $V(\phi)$.*

6.2 Explanation in Terms of the ODE

Besides of being the discretization of the defining equation (5.17), the algebraic Riccati equation (6.18), respectively the resulting $\tilde{V}(\tilde{x}) = \tilde{x}^\top \tilde{P}\tilde{x}$, is also an appropriate means for the analysis of the finite-dimensional system $\dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t) - \tilde{B}a(\tilde{C}\tilde{x}(t))$ from (6.14). In particular, the algebraic Riccati equation can directly be derived from considering this finite-dimensional approximation of the perturbed system dynamics. As a result, the link to various known concepts from finite-dimensional systems theory is established without abstractly elevating these concepts to the Hilbert-space setting. That is why, this and the next section focus on the role of the algebraic Riccati equation (6.18) in terms of the ODE (6.14).

The considerations proceed from a linear matrix inequality over the algebraic Riccati inequality to the desired algebraic Riccati equation. All these types of matrix equations respectively inequalities are widely discussed in the literature,

see, e.g., Willems' famous work [190]. In fact, they are related to Lyapunov-function-based approaches to the circle criterion addressing absolute stability, see Remark 5.2.5. However, the resulting $\tilde{V}(\tilde{x})$ from the above considered algebraic Riccati equation does not need to become a Lyapunov function in the classical sense. In view of Section 4.5, it is only expected to be a partial Lyapunov function. That is why, the present and the next section differ from such approaches in the following respects. First, the introduction of the offset function ℓ from Theorem 5.3.2 allows to impose tailored (partial) negative definiteness conditions on the derivative of $\tilde{V}$. It thus enables a uniform treatment of various stability concepts. Second, conditions that ensure the desired existence of $\tilde{P} = \tilde{P}^\top$ will not be mixed up with conditions that would be required to ensure that the resulting matrix $\tilde{P}$ is positive definite. The instead relevant partial positive definiteness could be characterized similarly to Remark 5.4.4, but, as it can simply be tested by means of Section 4.3, partial positive definiteness will not further be discussed.

6.2.1 The Linear Matrix Inequality (LMI)

Since the function $a(\zeta)$ is as in Section 5.2, the perturbation restriction in terms of $w(\zeta, a(\zeta))$ from Section 5.2 also remains unaltered. Consider the perturbation restriction with an offset ℓ (see Section 5.3, possibly $\ell(\zeta) \equiv 0$)

$$w(\tilde{C}\tilde{x}, a(\tilde{C}\tilde{x})) = \begin{bmatrix} \tilde{C}\tilde{x} \\ a(\tilde{C}\tilde{x}) \end{bmatrix}^\top \begin{bmatrix} \Pi_{\zeta\zeta} & \Pi_{\zeta a} \\ \Pi_{\zeta a}^\top & \Pi_{aa} \end{bmatrix} \begin{bmatrix} \tilde{C}\tilde{x} \\ a(\tilde{C}\tilde{x}) \end{bmatrix} \geq \ell(\tilde{C}\tilde{x}) \tag{6.20}$$

$\Pi_{\zeta\zeta} = \Pi_{\zeta\zeta}^\top, \Pi_{aa} = \Pi_{aa}^\top$. The inequality shall hold locally or globally (in terms of $\zeta = \tilde{C}\tilde{x}$, and thus also in terms of $\tilde{x}$) for the perturbation $a(\zeta)$ in (6.14). The fact that $w(\zeta, a(\zeta)) \geq 0$ is not satisfied by all arbitrary $\alpha = a(\zeta)$ is reflected by the involved matrix being indefinite, see Figure 6.2. Recap from Table 5.1 that relevant choices for the matrix in (6.20) are, e.g.,

$$\Pi_{(\mathrm{I}|\mathrm{a})} = \begin{bmatrix} \gamma^2 I & 0 \\ 0 & -I \end{bmatrix}, \quad \Pi_{(\mathrm{II}|\mathrm{a})} = \begin{bmatrix} 0 & \frac{1}{2}I \\ \frac{1}{2}I & -\rho I \end{bmatrix}, \quad \Pi_{(\mathrm{III}|\mathrm{a})} = \begin{bmatrix} -\mathrm{sym}(K_1^\top K_2) & \frac{1}{2}(K_1^\top + K_2^\top) \\ \frac{1}{2}(K_1 + K_2) & -I \end{bmatrix}.$$

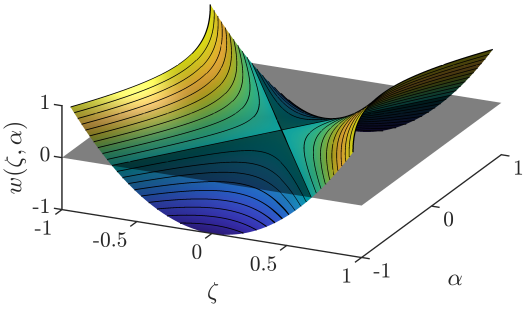


Figure 6.2: Let $p = m = 1$. The function $w(\zeta,\alpha) = \left[\begin{smallmatrix}\zeta\\ \alpha\end{smallmatrix}\right]^\top \left[\begin{smallmatrix}\Pi_{\zeta\zeta} & \Pi_{\zeta a}\\ \Pi_{\zeta a}^\top & \Pi_{aa}\end{smallmatrix}\right] \left[\begin{smallmatrix}\zeta\\ \alpha\end{smallmatrix}\right]$ must be a quadratic form from an indefinite matrix: If the matrix was positive semidefinite, the inequality $w(\zeta,\alpha) \geq 0$ would hold for all points—thus not yielding a perturbation restriction. If the matrix was negative semidefinite but nonzero, the inequality would hold only for a meagre set of points—amounting to a linear, exactly determined $\alpha = a(\zeta) = K\zeta$ that is no perturbation of interest. In contrast, if the matrix is indefinite, the zero-superlevel set $\{(\zeta,\alpha) : w(\zeta,\alpha) \geq 0\}$ gives a sector. This sector describes the admissible range for the graph of $\zeta \mapsto \alpha = a(\zeta)$, see Table 5.1 in Chapter 5. As a consequence, M in (6.24) is the sum of two indefinite matrices, and that sum might become negative semidefinite.

In view of the result on the derivative $D^+_{(\dot{x}(t)=f(x_t)-Ba(Cx_t))}V(x_t)$ that is established for LK functionals of robust type in (5.23), the objective shall be to find a function $\tilde{V}(\tilde{x})$ such that

$$D^+_{(\dot{\tilde{x}}=\tilde{A}\tilde{x}-\tilde{B}a(\tilde{C}\tilde{x}))}\tilde{V}(\tilde{x}) \leq -\ell(\tilde{C}\tilde{x}), \tag{6.21}$$

$\forall \tilde{x} \in G$ on some domain $G \subseteq \mathbb{R}^{\tilde{n}}$, $0_{\tilde{n}} \in G$. If a function $\tilde{V}(\tilde{x})$ can be found for which

$$D^+_{(\dot{\tilde{x}}=\tilde{A}\tilde{x}-\tilde{B}a(\tilde{C}\tilde{x}))}\tilde{V}(\tilde{x}) \leq -w(\tilde{C}\tilde{x}, a(\tilde{C}\tilde{x})), \tag{6.22}$$

$\forall \tilde{x} \in \mathbb{R}^n$, then this objective (6.21) is clearly achieved wherever

$$-w(\tilde{C}\tilde{x}, a(\tilde{C}\tilde{x})) \leq -\ell(\tilde{C}\tilde{x}),$$

i.e., wherever the perturbation restriction (6.20) holds.

Thus, is suffices to focus on (6.22). Being interested in a quadratic form $\tilde{V}(\tilde{x}) = \tilde{x}^\top \tilde{P}\tilde{x}$, the left-hand side of (6.22) can be made explicit,

$$2\tilde{x}^\top \tilde{P}\big(\tilde{A}\tilde{x} - \tilde{B}a(\tilde{C}\tilde{x})\big) \leq -w(\tilde{C}\tilde{x}, a(\tilde{C}\tilde{x})), \tag{6.23}$$

or in matrix notation[1], using (6.20),

$$\begin{aligned}
&2\tilde{x}^\top \tilde{P}\big(\tilde{A}\tilde{x} - \tilde{B}a(\tilde{C}\tilde{x})\big) + w(\tilde{C}\tilde{x}, a(\tilde{C}x)) \\
&= \begin{bmatrix} \tilde{x} \\ -a(\tilde{C}\tilde{x}) \end{bmatrix}^\top \underbrace{\left(\begin{bmatrix} 2\tilde{P}\tilde{A} & 2\tilde{P}\tilde{B} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \tilde{C}^\top \Pi_{\zeta\zeta} \tilde{C} & -\tilde{C}^\top \Pi_{\zeta a} \\ -\Pi_{\zeta a}^\top \tilde{C} & \Pi_{aa} \end{bmatrix} \right)}_{N} \begin{bmatrix} \tilde{x} \\ -a(\tilde{C}\tilde{x}) \end{bmatrix} \\
&= \begin{bmatrix} \tilde{x} \\ -a(\tilde{C}\tilde{x}) \end{bmatrix}^\top \underbrace{\left(\begin{bmatrix} \tilde{P}\tilde{A} + \tilde{A}^\top \tilde{P} & \tilde{P}\tilde{B} \\ \tilde{B}^\top \tilde{P} & 0 \end{bmatrix} + \begin{bmatrix} \tilde{C}^\top \Pi_{\zeta\zeta} \tilde{C} & -\tilde{C}^\top \Pi_{\zeta a} \\ -\Pi_{\zeta a}^\top \tilde{C} & \Pi_{aa} \end{bmatrix} \right)}_{M = \mathrm{sym}(N)} \begin{bmatrix} \tilde{x} \\ -a(\tilde{C}\tilde{x}) \end{bmatrix} \qquad \leq 0.
\end{aligned} \tag{6.24}$$

Hence, negative semidefiniteness of the involved real symmetric matrix M,

$$\exists \tilde{P} = \tilde{P}^\top \in \mathbb{R}^{\tilde{n}\times\tilde{n}} : \qquad M \preceq 0, \tag{6.25}$$

ensures that (6.22) holds for all $\tilde{x} \in \mathbb{R}^{\tilde{n}}$.

As a result, only the linear matrix inequality (6.25) must be solved for some symmetric matrix $\tilde{P}$. Provided a corresponding matrix $\tilde{P}$ is chosen in $\tilde{V}(\tilde{x}) = \tilde{x}^\top \tilde{P}\tilde{x}$, that function $\tilde{V}(\tilde{x})$ achieves that the objective (6.21) is accomplished for all $\tilde{x}$ for which the perturbation restriction (6.20) holds.

Remark 6.2.1 (Offset function vs. strict S-procedure). *The above introduced offset function ℓ enables a unified treatment of stability, asymptotic stability, partial asymptotic stability from Theorem 4.5.4, or other concepts which amount to a special choice of the right-hand side in (6.21). The associated strengthening only affects the perturbation restriction (6.20), whereas the requirement $M \preceq 0$ in (6.25), (and, in the end, the frequency domain condition from the KYP*

[1] For $N \in \mathbb{R}^{\nu\times\nu}$ and $z \in \mathbb{R}^\nu$: $z^\top \mathrm{skw}(N) z = 0$. Thus, $z^\top N z = z^\top \mathrm{sym}(N) z$.

lemma) remains unaltered. Although the underlying idea is the same, the usual S-procedure[2] approach chosen by Yakubovich for Lyapunov-function-based considerations of absolute stability [78, 200, 5, 197, 196], differs in this respect from the above result: The sector condition is commonly treated as remaining intact even if asymptotic stability is of interest, and, instead, negative definiteness of the Lyapunov function derivative is resolved by a "strict S-procedure"[3] resulting in a strict inequality $M \prec 0$. For the latter, the question of solvability then has to be tackled separately (amounting to a strict, also called regular or nondegenerate, variant of the KYP lemma [78, Thm. 2], in contrast to Lemma 6.3.1 below that addresses $M \preceq 0$).

In the following section, the known equivalence between the LMI (6.25) and the algebraic Riccati inequality, given by (6.18) with some arbitrary $\tilde{E} \succeq 0$, will be discussed. Moreover, it will be explained why nothing is lost by solving the algebraic Riccati equation (6.18) with $\tilde{E} = 0$ instead.

In fact, solving the algebraic Riccati equation is the way to go. First, solving an algebraic Riccati equation via standard routines is computationally less expensive than starting up a semidefinite programming solver that seeks for a matrix $\tilde{P}$ in the LMI (6.25). Second, to establish convergence to the searched LK functional of robust type, a unique solution must be agreed upon, which is why an inequality is inappropriate.

[2] Let $D = D^\top, W = W^\top \in \mathbb{R}^{n_z \times n_z}$. The relevant simplified form of the S-procedure is $\big(\forall z \in \{z \in \mathbb{R}^{n_z} : z^\top W z \geq 0\} : z^\top(-D)z \geq 0\big) \iff -M := -D - W \succeq 0$, cf. [25, Sec. 2.6.3]. See Lemma 6.2.3 (c3) for its application.

[3] $\big(\forall z \in \{z \in \mathbb{R}^{n_z} : z^\top W z \geq 0\} \setminus \{0\} : z^\top(-D)z > 0\big) \iff -M := -D - W \succ 0$

6.2.2 The Algebraic Riccati Inequality (ARI) and Comments on the LMI

Consider the LMI (6.25). The negative semidefiniteness of the involved block matrix from (6.24)

$$M = \begin{bmatrix} \tilde{P}\tilde{A} + \tilde{A}^\top \tilde{P} & \tilde{P}\tilde{B} \\ \tilde{B}^\top \tilde{P} & 0_{m\times m} \end{bmatrix} + \begin{bmatrix} \tilde{C}^\top \Pi_{\zeta\zeta} \tilde{C} & -\tilde{C}^\top \Pi_{\zeta a} \\ -\Pi_{\zeta a}^\top \tilde{C} & \Pi_{aa} \end{bmatrix}, \tag{6.26}$$

where $M_{22} = \Pi_{aa} \prec 0_{m\times m}$ according to (5.15), respectively positive semidefiniteness of $-M$, can be reformulated in terms of its Schur complement.

Lemma 6.2.2 (Semidefiniteness and Schur complement [91, Thm. 1.12]). *A matrix $M = M^H = \left[\begin{smallmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{smallmatrix}\right]$ with a positive definite submatrix M_{22} is positive semidefinite if and only if its Schur complement $M/M_{22} = M_{11} - M_{12}M_{22}^{-1}M_{21}$ is positive semidefinite.*

Therefore, $-M \succeq 0$ from (6.25) is equivalent to Lemma 6.2.3a below, which is the desired algebraic Riccati inequality. This equivalence is well known [190]. The following lemma also describes some further straightforward equivalences, respectively interpretations, that will become relevant in the sequel.

Lemma 6.2.3. *Let M be given by (6.26). The following statements are equivalent.*

(a) **(Algebraic Riccati inequality (ARI), Lemma 6.2.2 applied to (b))**
$M/M_{22} \preceq 0_{\tilde{n}\times\tilde{n}}$, *based on the Schur complement*

$$M/M_{22} = \underbrace{\tilde{P}\tilde{A} + \tilde{A}^\top \tilde{P} + \tilde{C}^\top \Pi_{\zeta\zeta}\tilde{C}}_{M_{11}} - \underbrace{\left[\tilde{P}\tilde{B} - \tilde{C}^\top \Pi_{\zeta a}\right]}_{M_{12}} \Pi_{aa}^{-1} \underbrace{\left[\tilde{B}^\top \tilde{P} - \Pi_{\zeta a}^\top \tilde{C}\right]}_{M_{21}}$$

(under the assumption $M_{22} = \Pi_{aa} \prec 0_{m\times m}$ from (5.15)).

(b) **(Linear matrix inequality (LMI))**
$M \preceq 0_{(\tilde{n}+m)\times(\tilde{n}+m)}$, *where M is given by (6.26).*

(c1) **(Negative semidefiniteness in $\mathbb{R}$)**
$\forall \tilde{x} \in \mathbb{R}^{\tilde{n}}, \forall u \in \mathbb{R}^m : \left[\begin{smallmatrix}\tilde{x}\\ u\end{smallmatrix}\right]^\top M \left[\begin{smallmatrix}\tilde{x}\\ u\end{smallmatrix}\right] \leq 0.$

(c2) **(Using u from (c1) instead of $-a(\tilde{C}\tilde{x})$ in (6.24))**
$\forall \tilde{x} \in \mathbb{R}^{\tilde{n}}, \forall u \in \mathbb{R}^m :$

$$2\tilde{x}^\top \tilde{P}\big(\tilde{A}\tilde{x} + \tilde{B}u\big) \leq -w(\tilde{C}\tilde{x}, -u),$$
$$\text{with } w(\tilde{C}\tilde{x}, -u) = \left[\begin{smallmatrix}\tilde{C}\tilde{x}\\ -u\end{smallmatrix}\right]^\top \left[\begin{smallmatrix}\Pi_{\zeta\zeta} & \Pi_{\zeta a}\\ \Pi_{\zeta a}^\top & \Pi_{aa}\end{smallmatrix}\right] \left[\begin{smallmatrix}\tilde{C}\tilde{x}\\ -u\end{smallmatrix}\right].$$

(c3) **(Interpreting (c2) as S-Procedure)**
$2\tilde{x}^\top \tilde{P}\big(\tilde{A}\tilde{x}+\tilde{B}u\big) \leq 0$ *for all* $\tilde{x} \in \mathbb{R}^{\tilde{n}}, u \in \mathbb{R}^m$ *that satisfy* $w(\tilde{C}\tilde{x}, -u) \geq 0$
(see, e.g., [25, Sec. 2.6.3] for the S-procedure; cf. Remark 6.2.1 above).

(c4) **(Interpreting (c2) as dissipativity inequality)**
For $V(\tilde{x}) = \tilde{x}^\top \tilde{P}\tilde{x}$ *it holds*

$$D^+_{(\dot{\tilde{x}}=A\tilde{x}+Bu)} V(\tilde{x}, u) \leq w^{\mathrm{sys}}(u, \tilde{C}\tilde{x}),$$

with

$$\begin{aligned} w^{\mathrm{sys}}(u, \tilde{C}\tilde{x}) &:= -w(\tilde{C}\tilde{x}, -u) \\ &= u^\top(-\Pi_{aa})u + 2\tilde{x}^\top \tilde{C}^\top \Pi_{\zeta a} u + \tilde{x}^\top \tilde{C}^\top(-\Pi_{\zeta\zeta})\tilde{C}\tilde{x}. \end{aligned}$$

Consequently, if additionally $V(\tilde{x}) \geq 0$ *is satisfied, then, following [191],* $V(\tilde{x})$ *qualifies as a storage function and the linear nominal system (6.10) is dissipative w.r.t. the supply rate* w^{sys}.

(d1) **(Negative semidefiniteness in $\mathbb{C}$)**
$\forall \tilde{x} \in \mathbb{C}^{\tilde{n}}, \forall u \in \mathbb{C}^m : \left[\begin{smallmatrix}\tilde{x}\\ u\end{smallmatrix}\right]^H M \left[\begin{smallmatrix}\tilde{x}\\ u\end{smallmatrix}\right] \leq 0.$

(d2) **(Using $\tilde{x}, u$ from (d1) in (6.24) where[4] $M = M^\top$)**
$\forall \tilde{x} \in \mathbb{C}^{\tilde{n}}, \forall u \in \mathbb{C}^m :$

$$2\mathrm{Re}\big(\tilde{x}^H \tilde{P}(\tilde{A}\tilde{x} + \tilde{B}u)\big) \leq -w_{\mathbb{C}}(\tilde{C}\tilde{x}, -u),$$
$$\text{with } w_{\mathbb{C}}(\tilde{C}\tilde{x}, -u) = \left[\begin{smallmatrix}\tilde{C}\tilde{x}\\ -u\end{smallmatrix}\right]^H \left[\begin{smallmatrix}\Pi_{\zeta\zeta} & \Pi_{\zeta a}\\ \Pi_{\zeta a}^\top & \Pi_{aa}\end{smallmatrix}\right] \left[\begin{smallmatrix}\tilde{C}\tilde{x}\\ -u\end{smallmatrix}\right].$$

[4] For $N \in \mathbb{R}^{\nu\times\nu}$ and $z \in \mathbb{C}^\nu$: $z^H \mathrm{skw}(N) z \in \mathrm{i}\mathbb{R}$ and $z^H \mathrm{sym}(N) z \in \mathbb{R}$. Thus, $\mathrm{Re}(z^H N z) = z^H \mathrm{sym}(N) z$.

Remark 6.2.4 (Absolute stability). *The designation absolute stability emphasizes that the whole family of ODEs $\dot{x}(t) = \tilde{A}\tilde{x}(t) + \tilde{B}u$ exhibiting a term $u = -a(\tilde{C}\tilde{x})$ that is globally compatible with the sector condition $w(\tilde{C}\tilde{x}, -u) \geq 0$ is simultaneously considered. Therefore, the term absolute stability refers to the global asymptotic stability of the zero equilibrium under all such perturbations [5]. Nevertheless, the underlying methods are also known to be useful for local considerations [108].*

Remark 6.2.5 (Necessary condition for a quadratic $\tilde{V}$, S-procedure). *Clearly, for any system in the above described family, the obtained $\tilde{V}(\tilde{x}) = \tilde{x}^\top \tilde{P}\tilde{x}$ has a nonpositive derivative $D^+_{\tilde{A}\tilde{x}+\tilde{B}(-a(\tilde{C}\tilde{x}))}\tilde{V}(\tilde{x})$ by (6.22). The significance of the S-Procedure in Lemma 6.2.3 (c3) rather lies in a converse statement. Due to the losslessness [25, 199] of the S-Procedure, the feasibility of the LMI or ARI in Lemma 6.2.3 is not only sufficient but even necessary for the existence of a common quadratic $\tilde{V}(\tilde{x})$ with a nonpositive derivative $D^+_{\tilde{A}\tilde{x}+\tilde{B}(-a(\tilde{C}\tilde{x}))}\tilde{V}(\tilde{x})$ for the entire family of ODEs. As a consequence, unless non-quadratic Lyapunov functions are taken into account or a non-sector-type perturbation family is used, the robustness bounds (the admissible parameter range describing the sector size, e.g., the maximum linear norm bound γ in Table 5.1 (I|a)) cannot further be improved.*

Remark 6.2.6 (Interconnection of dissipative elements). *On the one hand, concerning the nominal system, the LMI can be interpreted as a dissipativity requirement according to Lemma 6.2.3 (c4), provided additionally $\tilde{P} \succeq 0$ holds. On the other hand, concerning the perturbation, if $\ell \equiv 0$, then $\Pi_{aa}, \Pi_{\zeta a}^\top, \Pi_{\zeta\zeta}$ from the perturbation restriction can also be interpreted as Q, S, R matrices in a QSR-dissipativity requirement on the static perturbation map. See Figure 6.3. Therefore, provided $\ell \equiv 0$, and $\tilde{P} \succeq 0$, the overall system can be understood as a neutral interconnection of dissipative elements in the sense of Willems [191].*

The positive semidefiniteness of $\tilde{P}$, respectively $\tilde{V}(\tilde{x}) \geq 0$, that is required in Remark 6.2.6, is easily established. The conditions in the following lemma might

already be satisfied by a vanishing $a(\zeta) = K_C\zeta$ with $K_C = 0_{m\times p}$ (similarly to the discussions in Section 5.4 and Remark 5.5.10).

Lemma 6.2.7 (Positive semidefiniteness of the solution $\tilde{P}$). *Assume there exists some $K_C \in \mathbb{R}^{m\times p}$ such that $a(\zeta) = K_C\zeta$ belongs to the sector $w(\zeta, a(\zeta)) \geq 0$, i.e., $\left[\begin{smallmatrix} I_p \\ K_C \end{smallmatrix}\right]^\top \left[\begin{smallmatrix} \Pi_{\zeta\zeta} & \Pi_{\zeta a} \\ \Pi_{\zeta a}^\top & \Pi_{aa} \end{smallmatrix}\right] \left[\begin{smallmatrix} I_p \\ K_C \end{smallmatrix}\right] \succeq 0_{p\times p}$, and such that $u = -a(\tilde{C}\tilde{x})$ is a stabilizing input, i.e., $\tilde{A} - \tilde{B}K_C\tilde{C}$ is Hurwitz. If $\tilde{V}(\tilde{x}) = \tilde{x}^\top\tilde{P}\tilde{x}$ achieves (6.22), i.e., $D^+_{\tilde{A}\tilde{x}-\tilde{B}a(\tilde{C}\tilde{x})}\tilde{V}(\tilde{x}) \leq -w(\tilde{C}\tilde{x}, a(\tilde{C}\tilde{x}))$ (which, by construction, is the case for any $\tilde{P} = \tilde{P}^\top$ solving the ARI, ARE, LMI, etc. considered in this section), then $\tilde{P}$ is positive semidefinite.*

Proof. Any $\tilde{x}_0 \in \mathbb{R}^{\tilde{n}}$ can be used as an initial value $\tilde{x}_0$ in the stabilized problem $\dot{\tilde{x}} = (\tilde{A} - \tilde{B}K_C\tilde{C})\tilde{x}$. Then, analogously to the proof of Theorem 5.4.1,

$$\begin{aligned}\tilde{V}(\tilde{x}_0) &= -\int_0^\infty D^+_{(\dot{\tilde{x}}=\tilde{A}\tilde{x}-\tilde{B}K_C\tilde{C}\tilde{x})}\tilde{V}(\tilde{x}(t))\,\mathrm{d}t \geq \int_0^\infty w(\tilde{C}\tilde{x}(t), K_C\tilde{C}\tilde{x}(t))\,\mathrm{d}t \\ &\geq 0.\end{aligned}$$

□

Remark 6.2.8 (Complexification, KYP lemma). *Lemma 6.2.3 (d2) is a first step towards the KYP lemma, that provides conditions under which not all $\tilde{x} \in \mathbb{C}^{\tilde{n}}$ must be tested in the inequality of Lemma 6.2.3 (d2), but only*

$$\tilde{x}_f = (\mathrm{i}\omega I_{\tilde{n}} - \tilde{A})^{-1}\tilde{B}u, \tag{6.27}$$

$\forall\omega \in \mathbb{R}$. Note that, for $\tilde{x} = \tilde{x}_f$ from (6.27), the left-hand side of the inequality $2\mathrm{Re}\big(\tilde{x}^H\tilde{P}(\tilde{A}\tilde{x} + \tilde{B}u)\big) \leq -w_{\mathbb{C}}(\tilde{C}\tilde{x}, -u)$ in Lemma 6.2.3 (d2) vanishes by

$$\mathrm{Re}\big(\tilde{x}_f^H\tilde{P}(\tilde{A}\tilde{x}_f + \underbrace{\tilde{B}u}_{(\mathrm{i}\omega I_{\tilde{n}}-\tilde{A})\tilde{x}_f \text{ by (6.27)}})\big) = \mathrm{Re}(\mathrm{i}\omega \underbrace{\tilde{x}_f^H\tilde{P}\tilde{x}_f}_{\in\mathbb{R} \text{ if } \tilde{P}=\tilde{P}^H}) = 0 \tag{6.28}$$

(and with it the explicit occurrence of $\tilde{P}$). Section 6.3 will be devoted to the KYP lemma, which thus only relies on $0 \leq -w_{\mathbb{C}}(\tilde{C}\tilde{x}_f, -u)$.

6.2.3 The Algebraic Riccati Equation (ARE)

Statement (a) in Lemma 6.2.3 is the algebraic Riccati inequality encountered in (6.18) with $\tilde{E} \succeq 0_{\tilde{n}\times\tilde{n}}$. The present section considers the strengthening to an algebraic Riccati equation, thus arriving at (6.18) with $\tilde{E} = 0_{\tilde{n}\times\tilde{n}}$. Due to the equivalences in Lemma 6.2.3, the latter can not only be recognized as a strengthening of the ARI (a) but equivalently as a strengthening of the LMI from statement (b) in Lemma 6.2.3. This strengthening can be made explicit based on the following Gram-matrix-related statement on semidefiniteness.

Lemma 6.2.9 (Semidefiniteness and decomposability, cf. [121, Ch. 5.5, Cor. 2]). *A matrix $M = M^H \in \mathbb{C}^{n\times n}$ is positive semidefinite with $\mathrm{rk}(M) = r$ if and only if there exists a matrix $\Gamma \in \mathbb{C}^{r\times n}$ with full row rank $\mathrm{rk}(\Gamma) = r$ such that M can be decomposed into $M = \Gamma^H\Gamma$.*

In fact, replacing the ARI (a) from Lemma 6.2.3 by the ARE, which is the last point (A) in the following lemma, amounts to imposing an additional rank restriction to the LMI from (b), given by (B) below. Such a relation between the LMI and the ARE is already mentioned in [190, Rem. 10]. The following lemma describes this known equivalence between (B) and (A) through some straightforward intermediate steps (E1)-(E3), which are themselves relevant equivalent known statements, see Remark 6.2.12 and Remark 6.2.13 below.

Lemma 6.2.10. *Let M be given by (6.26). The following statements are equivalent.*

(B) **(LMI with rank restriction)**
$M = \left[\begin{smallmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{smallmatrix}\right] \preceq 0_{(\tilde{n}+m)\times(\tilde{n}+m)}$ *and* $\mathrm{rk}(M) = \underbrace{\dim(M_{22})}_{m}$.

(E1) **(Matrix equation** (also Lur'e eq.)**, decomposition from Lemma 6.2.9)**
$\exists \Gamma_x \in \mathbb{R}^{q\times\tilde{n}}, \Gamma_a \in \mathbb{R}^{q\times m}, q \in \mathbb{N}$:
$\left[\begin{smallmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{smallmatrix}\right] = -\left[\begin{smallmatrix} \Gamma_x^\top \\ \Gamma_a^\top \end{smallmatrix}\right] \left[\begin{smallmatrix} \Gamma_x & \Gamma_a \end{smallmatrix}\right], \qquad \mathrm{rk}(\left[\begin{smallmatrix} \Gamma_x & \Gamma_a \end{smallmatrix}\right]) = m.$

(E2) **(Lur'e equation, derived from using $-\Gamma_a^\top \Gamma_a = M_{22} = \Pi_{aa}$ in (E1))**
$\exists \tilde{K} \in \mathbb{R}^{m\times\tilde{n}}$:
$\left[\begin{smallmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{smallmatrix}\right] = -\left[\begin{smallmatrix} \tilde{K}^\top \\ I_m \end{smallmatrix}\right](-\Pi_{aa})\left[\begin{smallmatrix} \tilde{K} & I_m \end{smallmatrix}\right].$

(E3) **(Entrywise (E2))**
$\exists \tilde{K} \in \mathbb{R}^{m\times\tilde{n}}: \quad M_{11} = -\tilde{K}^\top(-\Pi_{aa})\tilde{K} \quad \textit{and} \quad M_{21} = -(-\Pi_{aa})\tilde{K}.$

(A) **(Algebraic Riccati equation (ARE), solving for $\tilde{K}$ in (E3))**
$$\underbrace{\tilde{P}\tilde{A} + \tilde{A}^\top\tilde{P} + \tilde{C}^\top\Pi_{\zeta\zeta}\tilde{C}}_{M_{11}} = \underbrace{\left[\tilde{P}\tilde{B} - \tilde{C}^\top\Pi_{\zeta a}\right]}_{-\tilde{K}^\top(-\Pi_{aa})}\underbrace{\Pi_{aa}^{-1}\left[\tilde{B}^\top\tilde{P} - \Pi_{\zeta a}^\top\tilde{C}\right]}_{\tilde{K}=-(-\Pi_{aa})^{-1}M_{21}}.$$

The supposed strengthening in replacing (a) by (A) turns out to be uncritical. The solvability in terms of the mere existence of a $\tilde{P} = \tilde{P}^\top \in \mathbb{R}^{\tilde{n}\times\tilde{n}}$ that not only solves the algebraic Riccati inequality (a) but that solves even the algebraic Riccati equation (A) remains unaltered if a stabilizability condition on $(\tilde{A}, \tilde{B})$ can be presumed. See [163] and references therein for the following result and more general ones[5].

Lemma 6.2.11 (ARI versus ARE). *Assuming stabilizability of $(\tilde{A}, \tilde{B})$, there exists a $\tilde{P} = \tilde{P}^\top \in \mathbb{R}^{\tilde{n}\times\tilde{n}}$ rendering the statements of Lemma 6.2.3 valid if and only if there exists a $\tilde{P} = \tilde{P}^\top \in \mathbb{R}^{\tilde{n}\times\tilde{n}}$ rendering the statements of Lemma 6.2.10 valid.*

Clearly, if $\tilde{A}$ is Hurwitz, the stabilizability is trivially satisfied. But even in more general cases, the assumed stabilizability is weaker than the conditions in Lemma 6.2.7 and therefore not at all restrictive. Some remarks on the various statements in Lemma 6.2.10 are in order.

Remark 6.2.12 (Matrix equation). *Lemma 6.2.10 (E1) gives a matrix equation (which can also be called Lur'e equation like the subsequent equation (E2)) in terms of two additional unknown matrices, see the last column of Table 6.1. In the*

[5] without loss of generality, concluding solvability of the transformed ARE in (7.24)

literature, such representations are frequently used for the positive-real lemma (cf. Remark 6.3.6) or related results, see, e.g., [108, Lem. 6.2].

Remark 6.2.13 (Relation to optimal control, Lur'e equation). *That algebraic Riccati equations are related to LQR problems is a standard result in control theory. The formulation as the Lur'e equation given in Lemma 6.2.10 (E2), which is, e.g, used in [198], is a rather convenient starting point for these considerations. See Section 7.1.2.*

Remark 6.2.14 (Hamiltonian matrix, ARE). *Solutions of the ARE in Lemma 6.2.10 (A) can be derived from $\tilde{n}$-dimensional invariant subspaces of the $2\tilde{n} \times 2\tilde{n}$ Hamiltonian matrix*

$$H := \begin{bmatrix} \tilde{A}^{\mathrm{I}} & \tilde{B}(-\Pi_{aa})^{-1}\tilde{B}^\top \\ -\tilde{C}^\top \Pi_{\zeta\zeta}^{\mathrm{I}} \tilde{C} & -(\tilde{A}^{\mathrm{I}})^\top \end{bmatrix}, \tag{6.29}$$

see, e.g., [168, 16, 66], where $\tilde{A}^{\mathrm{I}} = \tilde{A} - \tilde{B}(-\Pi_{aa})^{-1}\Pi_{\zeta a}^\top \tilde{C}$ and $\Pi_{\zeta\zeta}^{\mathrm{I}} = \Pi_{\zeta\zeta} + \Pi_{\zeta a}(-\Pi_{aa})^{-1}\Pi_{\zeta a}^\top$ (cf. Lemma 7.1.2). Being a Hamiltonian matrix, the eigenvalues are symmetric in the sense of $\sigma(H) = \{\lambda_1, \ldots, \lambda_{\tilde{n}}, -\lambda_1, \ldots, -\lambda_{\tilde{n}}\}$. Assume there are no eigenvalues on the imaginary axis[6], and let the stable invariant subspace of H be described as the column space of some matrix $\left[\begin{smallmatrix} U \\ L \end{smallmatrix}\right] \in \mathbb{C}^{2\tilde{n}\times\tilde{n}}$. For instance (although there are more appropriate approaches for numerical implementations, see, e.g., [9]),

$$\begin{bmatrix} U \\ L \end{bmatrix} = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_{\tilde{n}} \\ | & & | \\ | & & | \\ l_1 & \cdots & l_{\tilde{n}} \\ | & & | \end{bmatrix} \tag{6.30}$$

[6] Eigenvalues with $\mathrm{Re}(\lambda_k) = 0$ occur in the limit case of a nonstrict inequality in the KYP lemma, which can give rise to an almost stabilizing ARE solution.

can be built from eigenvectors $\left[\begin{smallmatrix} u_k \\ l_k \end{smallmatrix}\right] \in \mathbb{C}^{2\tilde{n}}$, $k \in \{1, \dots, \tilde{n}\}$, *(respectively generalized eigenvectors in the case of non-semisimple eigenvalues) of* H *corresponding to the eigenvalues* $\{\lambda_1, \dots, \lambda_{\tilde{n}}\}$ *with negative real part. Assume* U *is nonsingular*[7]*. Then*

$$\tilde{P} = LU^{-1} \tag{6.31}$$

solves the ARE and the eigenvalues of $\tilde{A} - B\tilde{K}$ *with* $\tilde{K}$ *defined in (7.11) coincide with* $\{\lambda_1, \dots, \lambda_{\tilde{n}}\}$*. Thus, if* $\tilde{P}$ *is real symmetric, then it is the searched stabilizing solution of the ARE. See [207] for the role of a Hamiltonian operator for operator-valued AREs like (5.65).*

To sum up, the numerical approach proposed in this chapter relies on the (unique stabilizing real symmetric) solution of the ARE that is encountered in Lemma 6.2.10 (A). The resulting $V(\tilde{x}) = \tilde{x}^\top \tilde{P}\tilde{x}$ achieves that the condition $D^+_{(\dot{\tilde{x}}=\tilde{A}\tilde{x}-\tilde{B}a(\tilde{C}\tilde{x}))}\tilde{V}(\tilde{x}) \leq -\ell(\tilde{C}\tilde{x})$ holds, see (6.21). This condition parallels what is achieved by LK functionals of robust type in (5.23). As such, it could have been proven similarly starting from the ARE. However, there is also a straightforward construction of an LMI to achieve that condition. This constructive derivation has been chosen in the present section with the intention to clarify the interrelations between various concepts that are known from the realm of robustness or absolute stability theory in ODEs. The LMI is equivalent to an ARI, and, under a stabilizability assumption, it is not restrictive to require that not only the ARI but the corresponding ARE (which is the ARE from the proposed numerical approach to LK functionals of robust type) shall be solved by some $\tilde{P} = \tilde{P}^\top$. The introduced offset function ℓ achieves that the LMI, ARI, or ARE are not affected by whether stability, asymptotic stability, or partial asymptotic stability are of interest (see Remark 6.2.1). Instead, only the perturbation restriction (6.20) is adapted.

[7] A singular U might occur if $(\tilde{A}, \tilde{B})$ is not stabilizable: Consider $\tilde{A} = a$, $\tilde{B} = 0$, $\Pi_{\zeta a} = 0$, and $\tilde{C}^\top \Pi_{\zeta\zeta} \tilde{C} = q$, i.e., $H = \left[\begin{smallmatrix} a & 0 \\ -q & -a \end{smallmatrix}\right]$, which amounts to a scalar Lyapunov equation $2Pa = -q$ as a special case of the ARE in Lemma 6.2.10 (A). Then $\lambda_1 = a$, $\left[\begin{smallmatrix} u_1 \\ l_1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} -2a \\ q \end{smallmatrix}\right]$ and $\lambda_2 = -a$, $\left[\begin{smallmatrix} u_2 \\ l_2 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$. Hence, if $a < 0$, then $P = LU^{-1} = q/(-2a)$, whereas for $a > 0$, $U = 0$.

6.3 Solvability of the Finite-Dimensional LMI, ARI, or ARE

The KYP lemma provides a frequency domain criterion by which the existence of a solution $\tilde{P} = \tilde{P}^\top \in \mathbb{R}^{\tilde{n}\times\tilde{n}}$ can be checked before trying to calculate such an explicit solution that satisfies the LMI or ARI in Lemma 6.2.3, respectively the matrix equation or ARE in Lemma 6.2.10.

6.3.1 The Kalman–Yakubovich–Popov Lemma

As already announced in Remark 6.2.8, the KYP lemma is based on $w_{\mathbb{C}}$ from statement (d2) in Lemma 6.2.3, which is the extension of w to complex arguments (in particular, with $\cdot^H$ instead of $\cdot^\top$). The following variant without a controllability assumption is proven in [45], incorporating [45, Rem. 1 and 2]. See, e.g., [78, Sec. 3.1] for alternative conditions.

Lemma 6.3.1 (Kalman–Yakubovich–Popov (KYP)). *Consider the statements of Lemma 6.2.3 with $\tilde{P}$ being unknown and $w_{\mathbb{C}}(\tilde{C}\tilde{x}, -u) = \left[\begin{smallmatrix}\tilde{C}\tilde{x}\\ -u\end{smallmatrix}\right]^H \left[\begin{smallmatrix}\Pi_{\zeta\zeta} & \Pi_{\zeta a}\\ \Pi_{\zeta a}^\top & \Pi_{aa}\end{smallmatrix}\right] \left[\begin{smallmatrix}\tilde{C}\tilde{x}\\ -u\end{smallmatrix}\right]$. Assume $\tilde{A}$ has no eigenvalues on the imaginary axis and $\Pi_{aa} \prec 0_{m\times m}$. If, based on $\tilde{x}_f = (\mathrm{i}\omega I_{\tilde{n}} - \tilde{A})^{-1}\tilde{B}u$,*

$$0 \leq -w_{\mathbb{C}}(\tilde{C}\tilde{x}_f, -u), \tag{6.32}$$

for all $\omega \in \mathbb{R}$ and all $u \in \mathbb{C}^m$, then there exists a $\tilde{P} = \tilde{P}^\top \in \mathbb{R}^{\tilde{n}\times\tilde{n}}$ rendering the statements of Lemma 6.2.3 valid.

In terms of the transfer function (6.13), note that $\left[\begin{smallmatrix}\tilde{C}\tilde{x}_f\\ -u\end{smallmatrix}\right] = \left[\begin{smallmatrix}\tilde{G}(\mathrm{i}\omega)u\\ -u\end{smallmatrix}\right]$. Consequently, (6.32) can be rewritten as $-w_{\mathbb{C}}(\tilde{C}\tilde{x}_f, -u) = u^H W_{\tilde{G}}\, u \geq 0$, $\forall u \in \mathbb{C}^m$, with the so-called Popov function or spectral density

$$W_{\tilde{G}}(\mathrm{i}\omega) = -\left[\begin{smallmatrix}\tilde{G}(\mathrm{i}\omega)\\ -I_m\end{smallmatrix}\right]^H \left[\begin{smallmatrix}\Pi_{\zeta\zeta} & \Pi_{\zeta a}\\ \Pi_{\zeta a}^\top & \Pi_{aa}\end{smallmatrix}\right] \left[\begin{smallmatrix}\tilde{G}(\mathrm{i}\omega)\\ -I_m\end{smallmatrix}\right]. \tag{6.33}$$

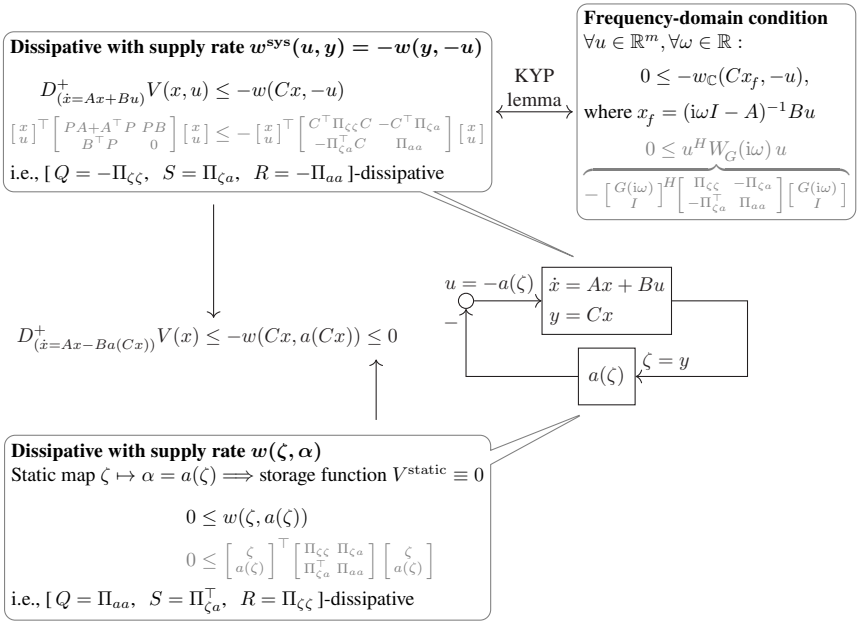


Figure 6.3: Interpretation as an interconnection of dissipative elements if $\ell \equiv 0$ in (6.20). The dissipativity interpretation additionally requires that $V(x) \ge 0$, see Lemma 6.2.7. For the sake of readability, a general LTI system (A, B, C) with state $x \in \mathbb{R}^n$ is depicted, instead of $(\tilde{A}, \tilde{B}, \tilde{C})$ and $\tilde{x} \in \mathbb{R}^{\tilde{n}}$ from (6.10), to which the results are applied in the present chapter. QSR-dissipativity amounts to a supply rate $w^{\mathrm{sys}}(u,y) = \left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right]^\top \left[\begin{smallmatrix} Q & S \\ S^\top & R \end{smallmatrix}\right] \left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right]$, given y describes the output and u the input of the respective element. The left upper box gives rise to the LMI in the first row of Table 6.1.

LMI in Lemma 6.2.3 (b)	**ARE** in Lemma 6.2.10 (A) (**ARI** in Lemma 6.2.3 (a))	**Matrix eq.** in Lemma 6.2.10 (E1) $\Gamma_x = L \in \mathbb{R}^{q\times n}, \Gamma_a = W \in \mathbb{R}^{q\times m}$
KYP lemma feasible if the parameters in $\Pi_{\zeta\zeta,\zeta a,aa}$ are such that $\lambda_{\min}\big(W_G(\mathrm{i}\omega)\big) = -\lambda_{\max}\Big((G(\mathrm{i}\omega))^H\Pi_{\zeta\zeta}G(\mathrm{i}\omega) - (G(\mathrm{i}\omega))^H\Pi_{\zeta a} - \Pi_{\zeta a}^\top G(\mathrm{i}\omega) + \Pi_{aa}\Big) \geq 0$		
$\begin{bmatrix} PA+A^\top P & PB \\ B^\top P & 0 \end{bmatrix} + \begin{bmatrix} C^\top\Pi_{\zeta\zeta}C & -C^\top\Pi_{\zeta a} \\ -\Pi_{\zeta a}^\top C & \Pi_{aa} \end{bmatrix} \preceq 0$	$PA+A^\top P + C^\top\Pi_{\zeta\zeta}C$ $+\Big[PB - C^\top\Pi_{\zeta a}\Big](-\Pi_{aa})^{-1}$ $\times\Big[B^\top P - \Pi_{\zeta a}^\top C\Big] = (\preceq)\, 0$	$PA+A^\top P + C^\top\Pi_{\zeta\zeta}C = -L^\top L$ $PB - C^\top\Pi_{\zeta a} = -L^\top W$ $\Pi_{aa} = -W^\top W$
Result for $\begin{bmatrix}\Pi_{\zeta\zeta} & \Pi_{\zeta a}\\ \Pi_{\zeta a}^\top & \Pi_{aa}\end{bmatrix} = \begin{bmatrix}\gamma^2 I & 0\\ 0 & -I\end{bmatrix}$ **Cor. 6.3.2 (a)** feasible if $\gamma \leq \gamma_{\max} := \frac{1}{\sup_\omega \|G(\mathrm{i}\omega)\|_2}$		
$\begin{bmatrix} PA+A^\top P & PB \\ B^\top P & 0 \end{bmatrix} + \begin{bmatrix} \gamma^2 C^\top C & 0 \\ 0 & -I \end{bmatrix} \preceq 0$	$PA+A^\top P + \gamma^2 C^\top C$ $+PBB^\top P = (\preceq)\, 0$	$PA+A^\top P + \gamma^2 C^\top C = -L^\top L$ $PB = -L^\top W$ $-I = -W^\top W$
Generalized bounded-real lemma feasible if $\gamma^{sys} \geq \gamma_{\min}^{sys} := \sup_\omega \|G(\mathrm{i}\omega)\|_2$		
$\begin{bmatrix} PA+A^\top P & PB \\ B^\top P & 0 \end{bmatrix} + \begin{bmatrix} C^\top C & 0 \\ 0 & -(\gamma^{sys})^2 I \end{bmatrix} \preceq 0$	$PA+A^\top P + C^\top C$ $+(\gamma^{sys})^2 PBB^\top P = (\preceq)\, 0$	$PA+A^\top P + C^\top C = -L^\top L$ $PB = -L^\top W$ $-(\gamma^{sys})^2 I = -W^\top W$
Result for $\begin{bmatrix}\Pi_{\zeta\zeta} & \Pi_{\zeta a}\\ \Pi_{\zeta a}^\top & \Pi_{aa}\end{bmatrix} = \begin{bmatrix}0 & \frac{1}{2}I\\ \frac{1}{2}I & -\rho I\end{bmatrix}$ **Cor. 6.3.2 (b)** feasible if $\rho \geq \rho_{\min} := \sup_\omega \mu_2(-G(\mathrm{i}\omega))$		
$\begin{bmatrix} PA+A^\top P & PB \\ B^\top P & 0 \end{bmatrix} + \begin{bmatrix} 0 & -\frac{1}{2}C^\top \\ -\frac{1}{2}C & -\rho I \end{bmatrix} \preceq 0$	$PA+A^\top P$ $+\frac{1}{\rho}[PB - \frac{1}{2}C^\top][B^\top P - \frac{1}{2}C]$ $= (\preceq)\, 0$	$PA+A^\top P = -L^\top L$ $PB - \frac{1}{2}C^\top = -L^\top W$ $-\rho I = -W^\top W$
Generalized positive-real lemma feasible if $\nu^{sys} \leq \nu_{\max}^{sys} := -\sup_\omega \mu_2(-G(\mathrm{i}\omega))$		
$\begin{bmatrix} PA+A^\top P & PB \\ B^\top P & 0 \end{bmatrix} + \begin{bmatrix} 0 & -\frac{1}{2}C^\top \\ -\frac{1}{2}C & \nu^{sys} I \end{bmatrix} \preceq 0$	$PA+A^\top P$ $-\frac{1}{\nu^{sys}}[PB - \frac{1}{2}C^\top][B^\top P - \frac{1}{2}C]$ $= (\preceq)\, 0$ if $\nu^{sys} < 0$	$PA+A^\top P = -L^\top L$ $PB - \frac{1}{2}C^\top = -L^\top W$ $\nu^{sys} I = -W^\top W$

Table 6.1: Explicit conditions (as in Figure 6.3, written without tildes for readability) and comparison with common forms of the (generalized—since not being restricted to $\gamma^{sys} = 1$) bounded-real lemma and a generalization (not restricted to $\nu^{sys} = 0$) of the positive-real lemma, see Remark 6.3.3 et seq.. Feasibility amounts to the existence of a matrix $P = P^\top$ that renders the equalities, respectively inequalities, valid. It is assumed that A has no eigenvalues on the imaginary axis (Lemma 6.3.1), and, concerning the ARE and the matrix equation, also that (A, B) is stabilizable (Lemma 6.2.11).

Thus, (6.32) requires

$$W_{\tilde{G}}(\mathrm{i}\omega) \succeq 0_{m\times m} \tag{6.34}$$

for all $\omega \in \mathbb{R}$. In contrast to the infinite-dimensional version in Lemma 5.5.6, which does not make any statement in the limit case $\alpha_3 = 0$ (associated to semidefiniteness without definiteness), the above finite-dimensional KYP lemma states that positive semidefiniteness of $W_{\tilde{G}}(\mathrm{i}\omega)$ is already sufficient for the existence of a symmetric $\tilde{P}$. Hence, the following corollaries, which are derived analogously to Corollary 5.5.13, 5.5.14, and 5.5.17, encounter nonstrict inequalities. The chosen terminology "complementary generalized bounded- (respectively, positive-) real lemma" will be explained below in Section 6.3.2, but might also already become clear from the juxtaposition in Table 6.1.

Corollary 6.3.2. *Assume $\tilde{A}$ has no eigenvalues on the imaginary axis, and let $\Pi = \left[\begin{smallmatrix} \Pi_{\zeta\zeta} & \Pi_{\zeta a} \\ \Pi_{\zeta a}^{\mathsf{T}} & \Pi_{aa} \end{smallmatrix}\right]$ be chosen according to one of the following three cases in Table 5.1: (a) row* (I|a)*; (b) row* (II|a)*; (c) row* (III) *with $K_2 = k_2 I_m$. It is achieved that*

- *the LMI (6.25) is feasible, and, equivalently,*
- *there exists a real symmetric matrix $\tilde{P}$ that satisfies the ARI from Lemma 6.2.3 (a), respectively (6.18) with some $\tilde{E} \succeq 0_{\tilde{n}\times n}$, and*
- *provided $(\tilde{A}, \tilde{B})$ is stabilizable, there exists a real symmetric solution $\tilde{P}$ of the ARE given by (6.18) with $\tilde{E} = 0_{\tilde{n}\times n}$*

if, relying on $\tilde{G}(\mathrm{i}\omega)$ from (6.13), the following holds for the respective case:

(a) Complementary generalized bounded-real lemma

$$\gamma \leq \gamma_{\max} := \frac{1}{\sup\limits_{\omega\in\mathbb{R}} \|\tilde{G}(\mathrm{i}\omega)\|_2} \tag{6.35}$$

in $\Pi_{(I|a)} = \begin{bmatrix} \gamma^2 I & 0 \\ 0 & -I \end{bmatrix}$ *from Table 5.1, row* (I|a).

(b) Complementary generalized positive-real lemma

$$\rho \geq \rho_{\min} := \sup_{\omega \in \mathbb{R}} \mu_2(-\tilde{G}(\mathrm{i}\omega)) \tag{6.36}$$

in $\Pi_{(II|a)} = \begin{bmatrix} 0 & \frac{1}{2} I \\ \frac{1}{2} I & -\rho I \end{bmatrix}$ *from Table 5.1, row* (II|a).

(c) Lower sector bound

Based on $\tilde{A}^{\mathrm{II}} = \tilde{A} - k_2 \tilde{B}\tilde{C}$ *and* $\tilde{G}^{\mathrm{II}} = \tilde{C}(\mathrm{i}\omega I_{\tilde{n}} - \tilde{A}^{\mathrm{II}})^{-1}\tilde{B}$,

$$K_1 = k_1 I_m \quad \text{with } k_1 \geq k_{1,\min} := k_2 - \frac{1}{\max_\omega \mu_2(\tilde{G}^{\mathrm{II}}(\mathrm{i}\omega))} \tag{6.37}$$

in $\Pi_{(III|a)} = \begin{bmatrix} -\mathrm{sym}(K_1^\top K_2) & \frac{1}{2}(K_1^\top + K_2^\top) \\ \frac{1}{2}(K_1 + K_2) & -I \end{bmatrix}$ *from Table 5.1, row* (III)*, with* $K_2 = k_2 I_m$.

Proof. Lemma 6.3.1 and Lemma 6.2.11. See Corollary 5.5.13, 5.5.14, and 5.5.17 on how to derive $\gamma_{\max}$, $\rho_{\min}$, and $k_{1,\min}$ from (6.34). □

The corresponding Nyquist plot restrictions in the SISO case $m = p = 1$ are already provided in Table 5.2 (the only difference is the nonstrict inequality in the above results).

6.3.2 Relation to Some Frequency-Domain and Dissipativity-Based Results for the ODE

See Remark 5.5.15 for the relation between (6.34) and the circle criterion. The above statements are not yet equivalent to the circle criterion since the KYP lemma only ensures the existence of a real symmetric solution $\tilde{P}$. In order to conclude (partial) stability of the ODE equilibrium via the (partial) Lyapunov function $\tilde{V}(\tilde{x}) = \tilde{x}^\top \tilde{P} \tilde{x}$, (partial) definiteness properties of $\tilde{P}$ are additionally

required. Such properties could be guaranteed a priori (similarly to Theorem 5.4.1 or Remark 5.4.4 involving observability properties), but once $\tilde{P}$ is computed from the ARE, (partial) positive definiteness can also simply be tested for that individual result. The same holds for the relation to the small-gain theorem or passivity theorems.

Remark 6.3.3 (Relation to the small-gain theorem). *As already noted below Corollary 5.5.13, the result for the maximum admissible linear norm bound $\gamma_{\max}$ in Corollary 6.3.2a mirrors the small-gain theorem, stating that the product of the perturbation gain γ and the L_2-gain of the nominal system (which is well known to be addressed by $\|G\|_\infty$) shall be smaller than one. However, this reciprocal relation between both gains does not even rely on the KYP lemma but can already be deduced from dissipativity considerations without involving the frequency domain. Compare the following:*

1) *Concerning the perturbation gain: The parameter γ from the perturbation restriction $\|a(\zeta)\|_2 \leq \gamma\|\zeta\|_2$ in row* (I|a) *from Table 5.1 is an upper bound on the gain of the static map*

$$\zeta \mapsto \alpha = a(\zeta). \tag{6.38}$$

Moreover, with the storage function $V^{\mathrm{static}} \equiv 0$ (see the lower box in Figure 6.3), this perturbation restriction, when written out as in Table 5.1 (I|a)

$$0 \leq \gamma^2 \zeta^\top \zeta - a^\top(\zeta) a(\zeta) = w(\zeta, a(\zeta)), \tag{6.39}$$

can be interpreted as a dissipativity inequality $D^+ V^{\mathrm{static}} \leq w(\zeta, \alpha)$ with supply rate $w(\zeta, \alpha)$ in terms of the perturbation input ζ and perturbation output α.

2a) Concerning the L_2-gain of the nominal system: An analogous supply rate (but in terms of the respective input u and output $\tilde{C}\tilde{x}$ of this system) would become relevant for an upper bound on the L_2-gain of

$$u \mapsto \tilde{C}\tilde{x}, \tag{6.40}$$

$u \in L_{2,\mathrm{loc}}([0,\infty),\mathbb{R}^m), \tilde{x} \in L_{2,\mathrm{loc}}([0,\infty),\mathbb{R}^{\tilde{n}})$, from the dynamical system $(\tilde{A},\tilde{B},\tilde{C})$. To be more precise, in that case, the dissipativity inequality

$$D^+_{(\dot{\tilde{x}}=\tilde{A}\tilde{x}+\tilde{B}u)}\tilde{V}_s(\tilde{x},u) \le (\gamma^{sys})^2 u^\top u - (\tilde{C}\tilde{x})^\top(\tilde{C}\tilde{x}) \tag{6.41}$$

would be feasible with a nonnegative storage function $\tilde{V}_s$ [108, eq. (5.29)]. In fact, the smallest possible value $\gamma^{sys} = \gamma^{sys}_{\min}$ in (6.41) is exactly the L_2-gain (provided A is Hurwitz, which ensures that $\tilde{V}_s \ge 0$, cf. Lemma 6.2.7 with $K_C = 0$).

2b) Concerning the derived admissible perturbation gain compatible with the nominal system: In contrast to (6.41), Lemma 6.2.3 (c4) with Table 5.1 (I|a) requires

$$D^+_{(\dot{\tilde{x}}=\tilde{A}\tilde{x}+\tilde{B}u)}\tilde{V}(\tilde{x},u) \le -w(\tilde{C}\tilde{x},-u) = -\gamma^2(\tilde{C}\tilde{x})^\top(\tilde{C}\tilde{x}) + u^\top u \tag{6.42}$$

(see the left upper box in Figure 6.3).

Comparing (6.41) and (6.42) shows that if 2a) is feasible for some given γ^{sys}, then 2b) is feasible for

$$\gamma = \frac{1}{\gamma^{sys}} \tag{6.43}$$

(with $\tilde{V} = \frac{1}{(\gamma^{sys})^2}\tilde{V}_s$). Consequently, the best possible value in 2b) becomes $\gamma_{\max} = \frac{1}{\gamma^{sys}_{\min}}$, the reciprocal of the L_2-gain.

Remark 6.3.4 (Bounded-real lemma). *The KYP lemma establishes the link to the frequency domain for both 2a) and 2b) in Remark 6.3.3. For simplicity assume that $\tilde{A}$ is Hurwitz, and thus $\max_\omega \|\tilde{G}(\mathrm{i}\omega)\|_2 = \|\tilde{G}\|_\infty$ is the H_∞-norm.*

2a) The corollary of the KYP lemma for the supply rate from 2a) (addressing the L_2-gain of the nominal system) states that $\gamma^{sys} \geq \gamma^{sys}_{\min} = \|\tilde{G}\|_\infty$ renders (6.42), and thus the corresponding LMI, ARI, etc. feasible. See Table 6.1. This result (or this result with $\gamma^{sys} = 1$) is commonly referred to as bounded-real lemma [25, 120, 33]. Strictly speaking, the bounded-real property of $\tilde{G}$ only refers to $\|\tilde{G}\|_\infty \leq 1$, i.e., $\gamma^{sys} = 1$ is fixed, see [25]. In order to emphasize that γ^{sys} above is some parameter, the term "generalized bounded-real lemma" is used in this thesis. Note that the bounded-real lemma usually requires that $\tilde{G} \in H_\infty$ since bounded realness is defined by $(\tilde{G}(s))^* \tilde{G}(s) \preceq I$ for $\mathrm{Re}(s) > 0$, see [25], which is equivalent to $\sup_{s\in\mathbb{C}^+} \|\tilde{G}(s)\|_2 \leq 1$. Moreover, $(\tilde{A}, \tilde{B}, \tilde{C})$ is usually assumed to be a minimal realization. In view of the mere KYP lemma, both is not necessary for the existence of $\tilde{P} = \tilde{P}^\top$. Nevertheless, under these additional conditions, the system matrix $\tilde{A}$ is necessarily Hurwitz, which by Lemma 6.2.7 with $K_C = 0$ (similarly to Remark 5.5.10) ensures positive semidefiniteness of $\tilde{P}$. The latter, in turn, is required anyway for γ^{sys} representing a bound on the L_2-norm, see [108, Thm. 5.5], respectively to allow for a storage function interpretation of $\tilde{V}$ and thus the dissipativity-based characterization of the L_2-gain in (6.41).

2b) Corollary 6.3.2a independently concludes that $\gamma \leq \gamma_{\max} = \frac{1}{\|\tilde{G}\|_\infty}$ renders (6.42), and thus the corresponding LMI, ARI, etc. in Table 6.1, feasible. That is why the term "complementary generalized bounded-real lemma" is used in Corollary 6.3.2a.

Remark 6.3.5 (Relation to a passivity theorem). *Analogously, Corollary 6.3.2b mirrors a passivity theorem if $\tilde{A}$ is assumed to be Hurwitz: The shortage of input passivity in the nominal dynamical system yields the minimum required excess of output passivity $\rho_{\min}$ in the perturbation; see, e.g., [203, 108, 165] for discussions on passivity properties. Again, to recognize this relation, the KYP lemma is not*

needed. The statement already becomes visible from comparing the respective dissipativity-based characterizations of the involved properties.

1) The parameter ρ from the perturbation restriction $a^\top(\zeta)\,\zeta \geq \rho\|a(\zeta)\|_2^2$ from row (II|a) *in Table 5.1 is a lower bound on the output passivity index of the static map $\zeta \mapsto a(\zeta)$ (defined as being the largest possible coefficient ρ in that inequality), and*

$$0 \leq a^\top(\zeta)\zeta - \rho a^\top(\zeta)a(\zeta) = w(\zeta, a(\zeta)) \tag{6.44}$$

can be interpreted as a dissipativity inequality with supply rate $w(\zeta, \alpha)$.

2a) If an analogous supply rate in terms of $u \mapsto \tilde{C}\tilde{x}$,

$$D^+_{(\dot{\tilde{x}}=\tilde{A}\tilde{x}+\tilde{B}u)}\tilde{V}(\tilde{x}, u) \leq (\tilde{C}\tilde{x})^\top u - \rho^{sys}(\tilde{C}\tilde{x})^\top(\tilde{C}\tilde{x}), \tag{6.45}$$

is feasible, and if $\tilde{V}(\tilde{x}) \geq 0$, then ρ^{sys} is a lower bound on the output passivity index of the dynamical system, whereas, in

$$D^+_{(\dot{\tilde{x}}=\tilde{A}\tilde{x}+\tilde{B}u)}\tilde{V}(\tilde{x}, u) \leq (\tilde{C}\tilde{x})^\top u - \nu^{sys}u^\top u, \tag{6.46}$$

ν^{sys} refers to a lower bound on its input passivity index. In both cases, $\tilde{V}(\tilde{x}) \geq 0$ must be presupposed in order to allow for the storage function interpretation, which, however, holds if $\tilde{A}$ is Hurwitz (cf. Lemma 6.2.7 with $K_C = 0$).

2b) In contrast, (see the left upper box in Figure 6.3)

$$D^+_{(\dot{\tilde{x}}=\tilde{A}\tilde{x}+\tilde{B}u)}\tilde{V}(\tilde{x}, u) \leq -w(\tilde{C}\tilde{x}, -u) = u^\top(\tilde{C}\tilde{x}) + \rho u^\top u \tag{6.47}$$

is required in Lemma 6.2.3 (c4) with Table 5.1 (II|a).

Since (6.47) is equivalent to (6.46) with $\rho = -\nu^{sys}$, it can be concluded that the minimum required $\rho_{\min} = -\nu^{sys}_{\max}$ is the negative of the input passivity index of the nominal dynamical system.

Remark 6.3.6 (Positive-real lemma). *See Table 6.1 for the frequency domain conditions from the KYP lemma for the supply rates from both 2a) and 2b) in Remark 6.3.5. In the limit case $\nu^{sys} \to 0$ in 2a), the condition simplifies to the well-known positive-real*[8] *lemma [108, 25, 120, 33]. That is why, Corollary 6.3.2 (b), which addresses the above case 2b) is termed "complementary generalized positive-real lemma" in the present thesis. Finally, note that the input passivity index $\nu^{sys}_{\max}$ is always nonpositive since $(\tilde{A}, \tilde{B}, \tilde{C})$ is necessarily strictly proper, i.e., $\|\tilde{G}(s)\| \to 0$, as $|s| \to \infty$. If the input passivity index $\nu^{sys}_{\max}$ is negative, it describes a shortage of passivity. In the case of $p = m = 1$, the latter is obtained by measuring how far the Nyquist plot extends into the left complex half-plane, see Table 5.2 for a respective plot.*

To sum up, the solvability of the LMI, ARI, ARE or matrix equation for the matrix $\tilde{P}$ can be checked beforehand by the KYP lemma. The resulting frequency-domain conditions for this solvability coincide with those known from the circle criterion (though the conditions are weaker since only the existence of $\tilde{P} = \tilde{P}^\top$ is addressed but no positive-definiteness requirements are imposed on $\tilde{P}$). These solvability conditions are closely related to famous corollaries of the KYP lemma like the bounded-real lemma and the positive-real lemma. In fact, they tackle closely related LMIs, respectively ARIs, AREs, or matrix equations, see Table 6.1. Whilst the outcome of these corollaries are system properties like the L_2-gain or the input passivity index, the outcome of interest in the present context is the admissible sector bound on a perturbation of the nominal system. The latter is instead given by the reciprocal of the L_2-gain or the negative of the input passivity index

[8] Similarly to the discussion on bounded-realness property in Remark 6.3.4, the positive-realness property does not only amount to $-\sup_{\omega\in\mathbb{R}} \mu_2(-G(\mathrm{i}\omega)) \geq \nu^{sys} := 0$ stated in Table 6.1. Being defined by $(G(s))^* + G(s) \succeq 0$ for all $\mathrm{Re}(s) > 0$, cf. [25], it actually amounts to the property $-\sup_{s\in\mathbb{C}+} \mu_2(-G(s)) \geq 0$ referring to the whole half plane, derived from $\inf_{s\in\mathbb{C}+} \lambda_{\min}((G(s))^* + G(s)) = -\sup_{s\in\mathbb{C}+} \lambda_{\max}(2\mathrm{He}(-G(s))) = -2\sup_{s\in\mathbb{C}+} \mu_2(-G(s))$. For the latter to hold, $G(s)$ cannot have any poles s_k with $\mathrm{Re}(s_k) > 0$, see [77, Prop. 3.3]. Nonexistence of such unstable poles is actually not required for the mere solvability from the KYP lemma, but, similarly to Remark 6.3.4, desirable anyway.

of the nominal system (yielding the admissible gain of the static perturbation, respectively the admissible output passivity index of the static perturbation).

6.4 Revisiting the Main Points of the Chapter

- The Legendre-tau based approach is applied to the defining equation of LK functionals of robust type.
  - Instead of the finite-dimensional Lyapunov equation that occurs in the numerical approach for complete-type and related LK functionals in Section 4.1, a finite-dimensional algebraic Riccati equation (ARE), given by (6.18), has to be solved.
  - To be more precise, the considered matrix $\tilde{P} \in \mathbb{R}^{n(N+1)\times n(N+1)}$ is the so-called stabilizing solution of the ARE, which can be obtained via standard routines like `icare` in Matlab.
- Section 6.2 intends to clarify the role of the finite-dimensional ARE for the finite-dimensional ODE from Chapter 3.
  - Assume the objective is to construct a Lyapunov function $\tilde{V}(\tilde{x}) = \tilde{x}^\top \tilde{P}\tilde{x}$ that has a negative semidefinite Lyapunov-function derivative along solutions of the perturbed ODE (6.14) wherever the perturbation resides in the given sector (see also Example 5.2.2 for the choice of that perturbation restriction). This objective is straightforwardly achieved by solving the linear matrix inequality (LMI) (6.25) for a suitable matrix $\tilde{P}$.
  - As already proposed in Section 5.3 of the last chapter, an offset function ℓ is introduced in the description of the perturbation restriction. As a result, a strengthening of the negative semidefiniteness of the Lyapunov function derivative to a partial negative definiteness does

not amount to a strengthening of the LMI but is covered by this strengthening of the perturbation restriction (Remark 6.2.1).

  - The LMI is known to be equivalent to an algebraic Riccati inequality (ARI) in $\tilde{P}$ (Lemma 6.2.3). The corresponding ARE (Section 6.2.3), in turn, is the one that has to be solved in the proposed numerical approach for LK functionals of robust type.

  - Although solvability of an equation seems to be a stronger requirement than solvability of an inequality, it is a known result (Lemma 6.2.11) that, under a stabilizability assumption, solvability of the ARI is equivalent to solvability of the corresponding ARE.

- The Kalman–Yakubovich–Popov (KYP) lemma (Section 6.3.1) provides a frequency-domain condition for the solvability of the LMI, ARI and (as stabilizability is assumed) also of the ARE of interest.

  - Corollary 6.3.2 summarizes the most important corollaries, which, in fact, are the sector bounds already encountered in Table 5.2 from the infinite-dimensional problem.

- Relations to dissipativity (Figure 6.3), the small-gain theorem (Remark 6.3.3), passivity theorems (Remark 6.3.5), and the relation to common corollaries of the KYP lemma in Table 6.1, like the bounded-real lemma (Remark 6.3.4) or the positive-real lemma (Remark 6.3.6) are discussed.

7 LK Functionals of Robust Type: Relation to LQR Problems and Convergence of the Numerical Approach

It remains to derive a convergence statement for the thus obtained numerical approach to LK functionals of robust type. As the literature already provides an in-depth convergence analysis for the Legendre-tau-based solution of optimal control problems in time-delay systems, a comparison between the operator-valued algebraic Riccati equations in the present context and those from standard optimal control problems is pursued as a first step in this chapter.

The chapter is organized as follows. The closer look on standard optimal control problems in **Section 7.1** shows the differences to the problem of interest, which thus are taken into account in the convergence analysis that is tackled in **Section 7.2**. One of these issues is the transition from the solvability condition of the finite-dimensional ARE to the solvability condition of the operator-valued ARE, which is resolved in **Section 7.3** by considering the Legendre-tau-based transfer function. Based on the latter, the convergence proof for the numerical approach is finally completed. **Section 7.4** concludes the chapter by revisiting the main points.

7.1 Relation to LQR Problems

Algebraic Riccati equations (AREs) are well known for their role in optimal control. In fact, the chosen numerical approach, i.e., to consider the Legendre-tau-based finite-dimensional ARE (6.18) when actually being interested in an operator-valued ARE, is along the lines of the Legendre-tau-based early lumping approach proposed by Ito and Teglas in [99] for time-delay linear-quadratic regulator (LQR) problems. For the latter, in-depth convergence considerations are available [99], see also [70]. These existing results motivate a closer look at the similarities and differences between the problem under consideration and such LQR problems.

7.1.1 Standard LQR Problems

For time-delay systems, standard LQR problems have the form[1]

$$-V^{\mathrm{LQR}}(\phi) = \min_{u\in L_2} \int_0^\infty (x^\top(t)Q^{\mathrm{LQR}}x(t) + u^\top(t)R^{\mathrm{LQR}}u(t))\,\mathrm{d}t \qquad (7.1)$$
$$\text{subject to } \begin{cases} \dot{x}(t)=A_0x(t)+A_1x(t-h)+Bu(t), \\ x_0=\phi, \end{cases}$$

with $Q^{\mathrm{LQR}} \succeq 0_{n\times n}$, $R^{\mathrm{LQR}} \succ 0_{m\times m}$ (assuming the system is stabilizable).

[1] Clearly, the minimum is nonnegative. Still, $V^{\mathrm{LQR}}(\phi) \leq 0$ is henceforth defined as the negative of the value function in (7.1) to be closer to the algebraic Riccati equations in the present chapter, where Q^{LQR} will be replaced by a negative semidefinite matrix.

Instead of considering only $\phi \in C([-h,0],\mathbb{R}^n)$, it is a common practice to use the larger state space M_2 from Section 3.2.2, cf. [70], i.e.,

$$
\begin{aligned}
&-\left\langle \mathscr{P}^{\mathrm{LQR}} \left[\begin{smallmatrix} \phi \\ \phi(0) \end{smallmatrix}\right], \left[\begin{smallmatrix} \phi \\ \phi(0) \end{smallmatrix}\right] \right\rangle_{M_2} \\
&\quad = \min_{u\in L_2} \int_0^\infty \left(\left\langle \mathscr{Q}^{\mathrm{LQR}} \left[\begin{smallmatrix} x_t \\ x(t) \end{smallmatrix}\right], \left[\begin{smallmatrix} x_t \\ x(t) \end{smallmatrix}\right] \right\rangle_{M_2} + \left\langle R^{\mathrm{LQR}} u(t), u(t) \right\rangle_{\mathbb{R}^m} \right) \mathrm{d}t
\end{aligned}
\tag{7.2}
$$

$$
\text{subject to } \begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} \left[\begin{smallmatrix} x_t \\ x(t) \end{smallmatrix}\right] = \mathscr{A} \left[\begin{smallmatrix} x_t \\ x(t) \end{smallmatrix}\right] + \mathscr{B} u(t), \\ \left[\begin{smallmatrix} x_0 \\ x(0) \end{smallmatrix}\right] = \left[\begin{smallmatrix} \phi \\ \phi(0) \end{smallmatrix}\right] \in M_2, \end{cases}
$$

where $\mathscr{A}$ is defined in (3.12) and $\mathscr{Q}^{\mathrm{LQR}} \left[\begin{smallmatrix} \phi \\ \phi(0) \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0_{n_{[-h,0]}} \\ \tilde{Q}^{\mathrm{LQR}} \phi(0) \end{smallmatrix}\right]$. Note that, compared with (7.1), the optimization space $L_2([0,\infty),\mathbb{R}^m) \ni u$ is not altered. The extension only refers to the domain of admissible arguments ϕ in the functional $V^{\mathrm{LQR}}(\phi)$ from (7.1), amounting to more general initial functions ϕ in the dynamic constraint. Therefore, on the restriction to $\phi \in C$, both results coincide.

7.1.1.1 The Operator-Valued ARE from Standard LQR Problems

Since the minimum (7.2) is nonnegative, considering

$$
\mathscr{X}^{\mathrm{LQR}} := -\mathscr{P}^{\mathrm{LQR}} \tag{7.3}
$$

is more convenient from an optimal control point of view, whereas the above introduced operator $\mathscr{P}^{\mathrm{LQR}}$ is motivated by the context of the present thesis. The minimum (7.2) is attained by the control law

$$
u(t) = -\mathscr{K}^{\mathrm{LQR}} \left[\begin{smallmatrix} x_t \\ x(t) \end{smallmatrix}\right], \quad \text{with} \quad \mathscr{K}^{\mathrm{LQR}} = (R^{\mathrm{LQR}})^{-1} \mathscr{B}^* \mathscr{X}^{\mathrm{LQR}} \tag{7.4}
$$

$$
= -(R^{\mathrm{LQR}})^{-1} \mathscr{B}^* \mathscr{P}^{\mathrm{LQR}}, \tag{7.5}
$$

where $\mathscr{P}^{\mathrm{LQR}}$ is the unique stabilizing (i.e., such that the given control law renders the equilibrium in the closed loop system exponentially stable) self-adjoint solution of the associated operator-valued ARE[2]

$$\begin{aligned}&\langle \mathscr{P}^{\mathrm{LQR}}\mathscr{A}\mathsf{x},\mathsf{x}\rangle_{M_2} + \langle \mathscr{A}^*\mathscr{P}^{\mathrm{LQR}}\mathsf{x},\mathsf{x}\rangle_{M_2}\\ &\quad - \langle \mathscr{Q}^{\mathrm{LQR}}\mathsf{x},\mathsf{x}\rangle_{M_2} + \langle (R^{\mathrm{LQR}})^{-1}\mathscr{B}^*\mathscr{P}^{\mathrm{LQR}}\mathsf{x}, \mathscr{B}^*\mathscr{P}^{\mathrm{LQR}}\mathsf{x}\rangle_{\mathbb{R}^m} = 0 \end{aligned} \tag{7.6}$$

$\forall \mathsf{x} \in D(\mathscr{A})$, see [70].

7.1.1.2 The Finite-Dimensional ARE in an Early-Lumping Approach

To discretize the problem, and thus to address instead a finite-dimensional LQR problem, see, e.g., [70],

$$-\tilde{V}^{\mathrm{LQR}}(\tilde{x}_0) = \min_{u\in L_2} \int_0^\infty (\tilde{x}^\top(t)\tilde{Q}^{\mathrm{LQR}}\tilde{x}(t) + u^\top(t)R^{\mathrm{LQR}}u(t))\,\mathrm{d}t \tag{7.7}$$
$$\text{subject to } \begin{cases}\dot{\tilde{x}}(t)=\tilde{A}\tilde{x}(t)+\tilde{B}u(t),\\ \tilde{x}(0)=\tilde{x}_0\in\mathbb{R}^{\tilde{n}},\end{cases}$$

with $\tilde{V}^{\mathrm{LQR}}(\tilde{x}) = \tilde{x}^\top \tilde{P}^{\mathrm{LQR}}\tilde{x}$, by solving the finite-dimensional ARE[3]

$$\tilde{P}^{\mathrm{LQR}}\tilde{A} + \tilde{A}^\top\tilde{P}^{\mathrm{LQR}} - \tilde{Q}^{\mathrm{LQR}} + \tilde{P}^{\mathrm{LQR}}\tilde{B}(R^{\mathrm{LQR}})^{-1}\tilde{B}^\top\tilde{P}^{\mathrm{LQR}} = 0_{\tilde{n}\times\tilde{n}}, \tag{7.8}$$

is tantamount to the (computational tempting but not risk-free) paradigm "discretize-then-optimize" in contrast to “optimize-then-discretize”. In the terminology of the control synthesis for PDEs [144, 55], such an approach follows an “early lumping” strategy—in contrast to “late lumping”, where the discretization is postponed to the realization of the distributed control law (if it was known analytically).

[2] usually expressed in terms of $\mathscr{X}^{\mathrm{LQR}} = -\mathscr{P}^{\mathrm{LQR}}$
[3] usually expressed in terms of $\tilde{X}^{\mathrm{LQR}} = -\tilde{P}^{\mathrm{LQR}}$

7.1.2 An LQR Problem Associated with the ARE

The following section considers the finite-dimensional ARE from (6.18) (with $\tilde{E} = 0$), which, in fact, is the core of the numerical approach from Section 6.1. In view of the role of AREs in optimal control problems, it is not surprising that under certain conditions, the ARE solution $\tilde{P}$ is also related to the optimal value function of a certain optimal control problem. This observation is, e.g., employed by Yakubovich in his works on the KYP lemma [197, 198, 131], or by Willems studying dissipativity [190].

7.1.2.1 The Finite-Dimensional Indefinite LQR Problem Associated with the ARE in the Numerical Approach

Recap the origin of the matrix M on which the LMI discussed in Section 6.2.1 relies. It stems from (6.24) and (6.22), which, when substituting $u = -a(\tilde{C}\tilde{x})$, requires

$$\underbrace{D^+_{(\dot{\tilde{x}}=\tilde{A}\tilde{x}+\tilde{B}u)}\tilde{V}(\tilde{x},u) + w(\tilde{C}\tilde{x},-u)}_{\left[\begin{smallmatrix}\tilde{x}\\u\end{smallmatrix}\right]^\top M \left[\begin{smallmatrix}\tilde{x}\\u\end{smallmatrix}\right]} \leq 0. \tag{7.9}$$

Thus, the strengthening from the LMI $M \preceq 0$ in Lemma 6.2.3 to the Lur'e equation $M = -\begin{bmatrix}\tilde{K}^\top \\ I_m\end{bmatrix}(-\Pi_{aa})\begin{bmatrix}\tilde{K} & I_m\end{bmatrix}$ in Lemma 6.2.10 (E2) corresponds to

$$\underbrace{D^+_{(\dot{\tilde{x}}=\tilde{A}\tilde{x}+\tilde{B}u)}\tilde{V}(\tilde{x},u) + w(\tilde{C}\tilde{x},-u)}_{\left[\begin{smallmatrix}\tilde{x}\\u\end{smallmatrix}\right]^\top M \left[\begin{smallmatrix}\tilde{x}\\u\end{smallmatrix}\right]} = \underbrace{-(\tilde{K}\tilde{x}+u)^\top(-\Pi_{aa})(\tilde{K}\tilde{x}+u)}_{\left[\begin{smallmatrix}\tilde{x}\\u\end{smallmatrix}\right]^\top\left(-\left[\begin{smallmatrix}\tilde{K}^\top\\I_m\end{smallmatrix}\right](-\Pi_{aa})\begin{bmatrix}\tilde{K} & I_m\end{bmatrix}\right)\left[\begin{smallmatrix}\tilde{x}\\u\end{smallmatrix}\right]} \tag{7.10}$$

(equivalently, (7.15) below). The additionally introduced matrix $\tilde{K}$ becomes, as described in item (E3) of Lemma 6.2.10,

$$\tilde{K} = -(-\Pi_{aa})^{-1}M_{21} = -(-\Pi_{aa})^{-1}\big(\tilde{B}^\top\tilde{P} - \Pi_{\zeta a}^\top\tilde{C}\big). \tag{7.11}$$

The significance of the Lur'e equation, compared to the equivalent ARE in Lemma 6.2.10 (A), lies in the simplicity of recognizing from (7.10) the relation to the following optimal control problem, cf. [198, eq. (1.18)].

Lemma 7.1.1 (The LQR problem associated with the ARE). *Let $\tilde{P}$ be a solution of (6.18) with $\tilde{E} = 0$, i.e., of the ARE*

$$\tilde{P}\tilde{A} + \tilde{A}^\top \tilde{P} + \tilde{C}^\top \Pi_{\zeta\zeta} \tilde{C} + \Big[\tilde{P}\tilde{B} - \tilde{C}^\top \Pi_{\zeta a}\Big](-\Pi_{aa})^{-1}\Big[\tilde{B}^\top \tilde{P} - \Pi_{\zeta a}^\top \tilde{C}\Big] = 0_{\tilde{n}\times\tilde{n}}. \tag{7.12}$$

If $\tilde{P}$ is positive semidefinite and if $\tilde{K}$ from (7.11) is stabilizing, i.e., $u = -\tilde{K}\tilde{x}$ renders the zero equilibrium of $\dot{\tilde{x}} = \tilde{A}\tilde{x} + \tilde{B}u$ exponentially[4] stable, then

$$\tilde{V}(\tilde{x}_0) = \tilde{x}_0^\top \tilde{P} \tilde{x}_0, \tag{7.13}$$

for any $\tilde{x}_0 \in \mathbb{R}^{\tilde{n}}$, satisfies

$$-\tilde{V}(\tilde{x}_0) = \inf_{u\in L_2} \int_0^\infty -w(\tilde{C}\tilde{x}(t), -u(t))\,\mathrm{d}t \tag{7.14}$$
$$\text{subject to } \begin{cases} \dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t) + \tilde{B}u(t), \\ \tilde{x}(0) = \tilde{x}_0 \in \mathbb{R}^{\tilde{n}}, \end{cases}$$

in terms of $w(\zeta, \alpha) = \zeta^\top \Pi_{\zeta\zeta}\zeta + 2\zeta^\top \Pi_{\zeta a}\alpha + \alpha^\top \Pi_{aa}\alpha$. The minimum is attained[5] by choosing $u(t) = -\tilde{K}\tilde{x}(t)$, with $\tilde{K}$ defined in (7.11).

[4] See, e.g., [32, Thm. 25-2] for a theorem that includes the almost stabilizing limit case.

[5] If $\tilde{K}$ is only almost stabilizing, then the infimum is not attained by $u = -\tilde{K}\tilde{x}$. Consider the scalar example $\inf_u J(u)$ with $J(u) = \int_0^\infty (-\gamma^2 x^2(t) + u^2(t))\,\mathrm{d}t$, s.t. $\dot{x} = -x + u$. The ARE is $-2P = -\gamma^2 - P^2$, i.e., $P = 1 \pm \sqrt{1-\gamma^2}$. If $\gamma < \gamma_{\max} = 1/\|G\|_\infty = 1$, then a stabilizing $K = -B^\top P = -P$ exists, whereas if $\gamma = 1$, both solutions are only almost stabilizing with $\dot{x} = -x + u = -x - Kx = 0$, i.e., $x(t) \equiv x_0$. However, choosing (instead of $K = -B^\top P = -1$) some $K_\varepsilon = -1 + \varepsilon$, then $\dot{x} = -x - K_\varepsilon x = -\varepsilon x$ gives an exponentially decaying solution for any $\varepsilon > 0$. As a result, with $\gamma = \gamma_{\max}$, $J(Kx) = \int_0^\infty (-\gamma^2 x^2(t) + (Kx(t))^2)\,\mathrm{d}t = 0$ but, considering (7.16), $J(K_\varepsilon x) = \lim_{t\to\infty} x^\top(t)Px(t) - x_0^\top P x_0 + \int_0^\infty \|(Kx(t) - K_\varepsilon x(t))\|_2^2\,\mathrm{d}t \to -x_0^\top P x_0 = -x_0^2$, as $\varepsilon \to 0^+$.

Proof. From (7.10), i.e., from

$$D^+_{(\dot{\tilde{x}}=\tilde{A}\tilde{x}+\tilde{B}u)}\tilde{V}(\tilde{x},u)+\|(-\Pi_{aa})^{\frac{1}{2}}(\tilde{K}\tilde{x}+u)\|_2^2=-w(\tilde{C}\tilde{x},-u), \tag{7.15}$$

it follows that

$$\begin{aligned}&\inf_{u\in L_2}\int_0^\infty -w(\tilde{C}\tilde{x}(t),-u(t))\,\mathrm{d}t \qquad (7.16)\\ &=\inf_{u\in L_2}\Big(\underbrace{\lim_{t\to\infty}\tilde{V}(\tilde{x}(t))}_{\geq 0 \ \text{ if } \tilde{P}\succeq 0}-\tilde{V}(\tilde{x}(0))+\underbrace{\int_0^\infty\|(-\Pi_{aa})^{\frac{1}{2}}(\tilde{K}\tilde{x}(t)+u(t))\|_2^2\,\mathrm{d}t}_{=0 \ \text{ if } u=-\tilde{K}\tilde{x}}\Big),\end{aligned}$$

where $\lim_{t\to\infty}\tilde{V}(\tilde{x}(t))=0$ if a stabilizing input u is chosen. Thus, the infimum is $-\tilde{V}(\tilde{x}(0))$, which is attained by the (by assumption) stabilizing input $u(t)=-\tilde{K}\tilde{x}(t)$. Moreover, it is known that the quadratic ansatz $\tilde{x}^\top\tilde{P}\tilde{x}$ is not restrictive in view of a quadratic cost functional [198]. □

Note that the integrand in (7.14) is the dissipativity supply rate from Lemma 6.2.3 (c4), which is

$$\begin{aligned}-w(\tilde{C}\tilde{x},-u)&=-\begin{bmatrix}\tilde{C}\tilde{x}\\-u\end{bmatrix}^\top\begin{bmatrix}\Pi_{\zeta\zeta}&\Pi_{\zeta a}\\\Pi_{\zeta a}^\top&\Pi_{aa}\end{bmatrix}\begin{bmatrix}\tilde{C}\tilde{x}\\-u\end{bmatrix}\\&=-\tilde{x}^\top\tilde{C}^\top\Pi_{\zeta\zeta}\tilde{C}\tilde{x}+2\tilde{x}^\top\tilde{C}^\top\Pi_{\zeta a}u+u^\top(-\Pi_{aa})u, \qquad (7.17)\end{aligned}$$

cf. Figure 6.3. In contrast to the costs in standard LQR problems (7.7), which are always nonnegative, (7.17) is an indefinite quadratic form, see Figure 6.2. Therefore, (7.14) is henceforth referred to as an indefinite LQR problem.

In fact, the most important difference to the ARE (7.8) from standard LQR problems is that the ARE (7.12) that occurs in the numerical approach to LK functionals of robust type is associated to an indefinite LQR problem.

7.1.2.2 The Infinite-Dimensional Counterpart

The RFDE-based counterpart of (7.14) would be

$$-V(\phi) = \inf_{u \in L_2} \int_0^\infty -w(\mathcal{C}x_t, -u(t))\,\mathrm{d}t \qquad (7.18)$$
$$\text{subject to } \begin{cases} \dot{x}(t) = A_0 x(t) + A_1 x(t-h) + Bu(t), \\ x_0 = \phi. \end{cases}$$

However, the integrand

$$-w(\mathcal{C}x_t, -u(t)) = -\begin{bmatrix} C_1 x(t-h) \\ C_0 x(t) \end{bmatrix}^\top \Pi_{\zeta\zeta} \begin{bmatrix} C_1 x(t-h) \\ C_0 x(t) \end{bmatrix} + 2\begin{bmatrix} C_1 x(t-h) \\ C_0 x(t) \end{bmatrix}^\top \Pi_{\zeta a} u(t) + u^\top(t)(-\Pi_{aa})u(t) \qquad (7.19)$$

involves the point evaluation $x(t-h) = x_t(-h)$ of x_t. As a consequence, when the first term of (7.19) is expressed as a quadratic form in $\left[\begin{smallmatrix} x_t \\ x(t) \end{smallmatrix}\right] \in M_2$, an unbounded operator arises. In contrast, in the standard LQR problem from (7.2), only a bounded operator $\mathscr{Q}^{\mathrm{LQR}}$ occurs since the costs from (7.1) only rely on $x(t)$ but not on $x(t-h)$.

In fact, an operator-valued ARE that is associated with (7.18) is not considered in the present thesis. Note that the operator-valued ARE introduced in Section 5.5.2 is an equation for $\mathscr{P}_0$, not addressing the LK functional $V(\phi)$ but only $V_0(\phi)$ from the splitting approach. This splitting approach has been applied as an intermediate step in Section 5.5.1 exactly to prevent the occurrence of a point evaluation in the operator-valued ARE.

To sum up, a second major difference to standard LQR problems is that the numerical approach actually addresses an LQR problem where $x(t-h)$ occurs in the costs. This issue, however, is already resolved by the splitting approach from Section 5.5.1, the effect of which will be discussed in the sequel.

7.1.3 Effect of the Splitting Approach

Recap that the resulting operator-valued ARE (5.65) from the splitting approach reads

$$\begin{aligned}\langle \mathscr{P}_0 \mathscr{A}\mathsf{x}, \mathsf{x}\rangle_{M_2} + \langle \mathscr{A}^* \mathscr{P}_0 \mathsf{x}, \mathsf{x}\rangle_{M_2} \\ + \langle \mathscr{Q}\mathsf{x}, \mathsf{x}\rangle_{M_2} + \langle (-\Pi_{aa})^{-1} \mathscr{B}^* \mathscr{P}_0 \mathsf{x}, \mathscr{B}^* \mathscr{P}_0 \mathsf{x}\rangle_{\mathbb{C}^m} = 0,\end{aligned} \tag{7.20}$$

$\forall \mathsf{x} \in D(\mathscr{A})$, where $\mathscr{Q}\left[\begin{smallmatrix}\phi\\ r\end{smallmatrix}\right] = \left[\begin{smallmatrix}0_{L_2}\\ (Q_0+Q_1)r\end{smallmatrix}\right]$ with Q_0 and Q_1 from (5.52) and $\mathscr{B}^*\left[\begin{smallmatrix}\phi\\ r\end{smallmatrix}\right] = B^\top r$. Based on a solution $\mathscr{P}_0$ of (7.20), the first part of the split functional $V = V_0 + V_1$

$$V_0(\phi) = \left\langle \mathscr{P}_0 \left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right], \left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right] \right\rangle_{M_2} \tag{7.21}$$

is defined in (5.62).

7.1.3.1 The Infinite-Dimensional Indefinite LQR Problem from the Splitting Approach

The operator-valued ARE (7.20), respectively the Lur'e equation from Lemma 5.5.7, cf. [131], is associated with the LQR problem

$$-V_0(\phi) = \inf_{u\in L_2} \int_0^\infty \Big(-x^\top(t)(Q_0+Q_1)x(t) + u^\top(t)(-\Pi_{aa})u(t)\Big)\,\mathrm{d}t$$

$$\text{subject to } \begin{cases}\dot{x}(t) = A_0^{\mathrm{I}}x(t) + A_1^{\mathrm{I}}x(t-h) + Bu(t),\\ x_0 = \phi,\end{cases} \tag{7.22a}$$

$$V(\phi) = V_0(\phi) + V_1(\phi), \quad \text{with } V_1(\phi) = \int_{-h}^0 \phi^\top(\theta) Q_1 \phi(\theta)\,\mathrm{d}\theta. \tag{7.22b}$$

The latter is closer to the standard LQR problem from (7.1) than (7.18) since, in fact, only

$$Q^{\mathrm{LQR}} = -(Q_0+Q_1) \preceq 0_{n\times n}, \quad \text{and} \quad R^{\mathrm{LQR}} = -\Pi_{aa} \succ 0_{m\times m} \tag{7.23}$$

is substituted in (7.1). However, in contrast to the standard LQR problem (7.1), for which in-depth considerations are available, (7.22a) is still a non-standard LQR problem with indefinite costs as $Q^{\mathrm{LQR}} \preceq 0_{n\times n}$.

The next section has to show that applying the numerical approach without a splitting (which will be done in practice) and after that splitting approach (which will consequently be analyzed theoretically) does not make any difference in the result.

7.1.3.2 Splitting Equivalence in the Discretization

Applying the Legendre-tau-based approach only to the transformed and split problem (7.22a) would require to calculate in a second step $\tilde{V}(\tilde{x}) = \tilde{V}_0(\tilde{x}) + \tilde{V}_1(\tilde{x})$ by adding the Legendre-tau approximation for $\tilde{V}_1$ from (7.22b) that has been described in Section 4.6.1.2. In the following, it is shown that such a result coincides with $\tilde{V}(\tilde{x}) = \tilde{x}^\top \tilde{P} \tilde{x}$ with $\tilde{P}$ directly obtained from the untransformed ARE (7.12), which is rather the numerical approach that has been proposed in Section 6.1. In other words, the spitting and the discretization commute.

Consider the splitting approach from Section 5.5.1. Before the actual splitting, a transformation that eliminates the state-and-input cross term $\Pi_{\zeta a}$ has been applied in Lemma 5.5.1. First, the following lemma confirms that the equality between the LK functionals V and V^{I} before and after that transformation is mirrored by their numerical approximations.

Lemma 7.1.2 (Transformation I for the finite-dimensional ARE). *$\tilde{P}$ is a solution of the ARE (7.12) if and only if $\tilde{P} = \tilde{P}^{\mathrm{I}}$ solves the transformed ARE*

$$\begin{aligned} &\tilde{P}\tilde{A}^{\mathrm{I}} + (\tilde{A}^{\mathrm{I}})^\top \tilde{P} + \tilde{C}^\top \Pi^{\mathrm{I}}_{\zeta\zeta} \tilde{C} + \tilde{P}\tilde{B}(-\Pi_{aa})^{-1}\tilde{B}^\top \tilde{P} = 0_{\tilde{n}\times\tilde{n}} \\ &\quad \textit{with} \quad \tilde{A}^{\mathrm{I}} = \tilde{A} - \tilde{B}(-\Pi_{aa})^{-1}\Pi_{\zeta a}^\top \tilde{C}, \\ &\qquad\quad \Pi^{\mathrm{I}}_{\zeta\zeta} = \Pi_{\zeta\zeta} + \Pi_{\zeta a}(-\Pi_{aa})^{-1}\Pi_{\zeta a}^\top \end{aligned} \tag{7.24}$$

Proof. Rewriting (7.12) as (7.24). □

Second, the next lemma shows that the equality between V and $V_0 + V_1$ before and after the splitting also holds for the respective numerical approximations $\tilde{V}$ and $\tilde{V}_0 + \tilde{V}_1$. Recap that the block diagonal structure assumed in (5.51) led to the introduction of Q_0, Q_1 in (5.52). Based on the latter, the $\tilde{P}$-independent term in (7.24) can be rewritten as

$$\tilde{C}^\top \Pi^{\mathrm{I}}_{\zeta\zeta} \tilde{C} = \tilde{N}^\top \begin{bmatrix} C_1^\top \Pi^{\mathrm{I},11}_{\zeta\zeta} C_1 & 0_{p_1 \times p_0} \\ 0_{p_0 \times p_1} & C_0^\top \Pi^{\mathrm{I},11}_{\zeta\zeta} C_0 \end{bmatrix} \tilde{N} = \tilde{N}^\top \begin{bmatrix} Q_1 & 0_{p_1 \times p_0} \\ 0_{p_0 \times p_1} & Q_0 \end{bmatrix} \tilde{N} \quad (7.25)$$

$$\text{with } \tilde{N} = \tilde{C}|_{\{C_0 = I_n, C_1 = I_n\}}, \quad (7.26)$$

where, along the lines of (6.9), $\tilde{N}\tilde{x}(t) \approx \begin{bmatrix} x(t-h) \\ x(t) \end{bmatrix}$ addresses (5.8) under the substitution $C_0 = C_1 = I_n$.

Lemma 7.1.3 (Splitting and discretization commute). *Assume $\tilde{A}^I$ is Hurwitz, or more generally, $\sigma(\tilde{A}^I) \cap \sigma(-\tilde{A}^I) = \emptyset$. Then $\tilde{P}$ is a solution of the transformed ARE (7.24), i.e., the Legendre-tau-based discretization of (5.56), if and only if*

$$\tilde{P} = \tilde{P}_0 + \tilde{P}_1, \quad (7.27)$$

where $\tilde{P}_0$ solves the Legendre-tau-based discretization of the modified defining equation (5.55) given by

$$\tilde{P}_0 \tilde{A}^{\mathrm{I}} + (\tilde{A}^{\mathrm{I}})^\top \tilde{P}_0 + \tilde{N}^\top \begin{bmatrix} 0 & 0 \\ 0 & Q_0 + Q_1 \end{bmatrix} \tilde{N} + \tilde{P}_0 \tilde{B} (-\Pi_{aa})^{-1} \tilde{B}^\top \tilde{P}_0 = 0_{\tilde{n} \times \tilde{n}} \quad (7.28)$$

and $\tilde{P}_1$ is the solution of the Lyapunov equation representing the Legendre-tau discretization of (5.57) given by

$$\tilde{P}_1 \tilde{A}^{\mathrm{I}} + \tilde{(A^{\mathrm{I}})}^\top \tilde{P}_1 + \tilde{N}^\top \begin{bmatrix} Q_1 & 0 \\ 0 & -Q_1 \end{bmatrix} \tilde{N} = 0_{\tilde{n} \times \tilde{n}}. \quad (7.29)$$

Proof. The assumption on the eigenvalues of $\tilde{A}^I$ ensures uniqueness of the Lyapunov equation solution $\tilde{P}_1$ by Lemma 4.1.2. To recognize that the sum of (7.29) and (7.28)

$$(\tilde{P}_0+\tilde{P}_1)\tilde{A}^{\mathrm{I}}+(\tilde{A}^{\mathrm{I}})^\top(\tilde{P}_0+\tilde{P}_1)+\tilde{N}^\top\begin{bmatrix}Q_1 & 0\\ 0 & Q_0\end{bmatrix}\tilde{N}+\tilde{P}_0\tilde{B}(-\Pi_{aa})^{-1}\tilde{B}^\top\tilde{P}_0 = 0_{\tilde{n}\times\tilde{n}} \tag{7.30}$$

coincides with (7.24) in terms of $\tilde{P}=\tilde{P}_0+\tilde{P}_1$, it must be shown that the quadratic last term in (7.30) and $(\tilde{P}_1+\tilde{P}_0)\tilde{B}(-\Pi_{aa})^{-1}\tilde{B}^\top(\tilde{P}_1+\tilde{P}_0)$ coincide. According to (4.40), $\tilde{P}_1$ has a zero last block column in Legendre coordinates. The latter is multiplied with $\tilde{B}$, which according to (6.8), in Legendre coordinates is nonzero only in the last block row. Hence, $\tilde{P}_1\tilde{B}$ vanishes (analogously to (5.59)) and thus $(\tilde{P}_1+\tilde{P}_0)\tilde{B}(-\Pi_{aa})^{-1}\tilde{B}^\top(\tilde{P}_1+\tilde{P}_0)=\tilde{P}_0\tilde{B}(-\Pi_{aa})^{-1}\tilde{B}^\top\tilde{P}_0$. □

Note that $\tilde{V}_1(\tilde{x})=\tilde{x}^\top\tilde{P}_1\tilde{x}$ has no counterpart in the cited analysis of LQR problems. However, it already turned out to provide a suitable approximation of $V_1(\phi)$ in Section 4.6.1.2. As a consequence, knowing that the result $\tilde{V}(\tilde{x})=\tilde{x}^\top\tilde{P}\tilde{x}$ from the numerical approach proposed in (6.1) equals $\tilde{V}(\tilde{x})=\tilde{V}_0(\tilde{x})+\tilde{V}_1(\tilde{x})$, and that the searched overall LK functional of robust type is $V(\phi)=V_0(\phi)+V_1(\phi)$, it suffices to focus on V_0. It only needs to be confirmed that $\tilde{V}_0(\tilde{x})=\tilde{x}^\top\tilde{P}_0\tilde{x}$, based on a positive semidefinite stabilizing solution $\tilde{P}_0$ of the ARE (7.28), converges to $V_0(\phi)$ from the indefinite LQR problem (7.22a), with convergence being meant in the same sense as in Section 4.6.2.

To sum up, for the standard LQR problem (7.1), the literature offers in-depth convergence results for a Legendre-tau-based early lumping approach. The splitting approach for LK functionals of robust type, which has already been introduced in Section 5.5.1, achieves that the resulting operator-valued ARE is associated with an LQR problem (7.22a) that has a very similar structure as these standard problems. Still, the involved indefinite costs are a significant difference to these standard problems and call for caution in the convergence analysis.

7.2 Convergence

The following convergence statement in Lemma 7.2.1 is borrowed from the Legendre-tau-based treatment of standard LQR problems in [99]. Whether the ARE stems from a standard LQR problem or an indefinite one does not make any difference in its proof (however, it will make a difference in the proofs of the involved conditions).

Analogously to (4.56), the lemma is expressed in terms of $\mathscr{P}_0^{[N]}$. The latter is an operator on M_2 that is represented by the $n(N+1) \times n(N+1)$ matrix $\tilde{P}_0$ from the ARE (7.28) such that

$$V_0(\phi) = \left\langle \mathscr{P}_0 \begin{bmatrix} \phi \\ \phi(0) \end{bmatrix}, \begin{bmatrix} \phi \\ \phi(0) \end{bmatrix} \right\rangle_{M_2} \approx \left\langle \mathscr{P}_0^{[N]} \begin{bmatrix} \phi \\ \phi(0) \end{bmatrix}, \begin{bmatrix} \phi \\ \phi(0) \end{bmatrix} \right\rangle_{M_2} = \tilde{x}^\top \tilde{P}_0 \tilde{x}. \tag{7.31}$$

In particular, $\mathscr{P}_0^{[N]}$ solves an operator-valued ARE that relies on the Legendre-tau based approximation $\mathscr{A}^{[N]}$ instead of $\mathscr{A}$. See Appendix A on the precise meaning of the statement that the matrices $\tilde{A}$ and $\tilde{P}_0$ are coordinate representations of the operators $\mathscr{A}^{[N]}$ and $\mathscr{P}_0^{[N]}$ and how the associated operator-valued ARE gives rise to the matrix-valued ARE. Whenever the discretization resolution has to be made explicit, a superscript $[N]$ will also be added at the matrices, e.g., $\tilde{A} = \tilde{A}^{[N]}$ or $\tilde{P}_0 = \tilde{P}_0^{[N]}$.

Lemma 7.2.1. *Assume existence and uniqueness hold for both the stabilizing ARE solution $\mathscr{P}_0$ from (5.65) and its Legendre-tau-based approximations $\mathscr{P}_0^{[N]}$ from any sufficiently large ($\exists N_{\min} > 1$) discretization resolution $N \geq N_{\min}$. Moreover, assume that $\{\|\mathscr{P}_0^{[N]}\|\}_{N \geq N_{\min}}$ is bounded. Then $\mathscr{P}_0^{[N]}$ converges weakly to $\mathscr{P}_0$ as $N \to \infty$.*

Proof. See [99, Thm. 5.1 (i)]. For the sake of plausibility, the main points of this known proof have already been outlined below Lemma 4.6.5 on the convergence of Lyapunov-equation solutions. The only discarded aspect was the quadratic term $\langle(-\Pi_{aa})^{-1}\mathscr{B}^*\mathscr{P}_0\mathsf{x}, \mathscr{B}^*\mathscr{P}_0\mathsf{x}\rangle_{\mathbb{C}^m}$ that, in fact, makes the difference between a

Lyapunov equation (4.55) and the ARE (7.20). This quadratic term, however, is unproblematic, since, due to the input dimension m being finite, $\mathscr{B}^* \mathscr{P}_0^{[N]} \overset{\text{strongly}}{\to} \mathscr{B}^* \mathscr{P}_0$ [99]. An alternative proof of the lemma is provided in [70, Thm. 6.7] (to be combined with the convergence results on the Legendre-tau-based approximation of the solution operator from [99, Thm. 3.6]). □

In the present section, the thus required properties of uniqueness, existence, and uniform boundedness will be discussed. If all these properties are ensured, a convergence statement on the derived approximations of the LK functionals of robust type can be concluded based on the above lemma. Thus, the procedure is similar to Section 4.6.2, where an adaption of the above lemma to Lyapunov equations has been used to derive a convergence statement for the numerical approximation of LK functionals of complete type.

7.2.1 Uniqueness

In an LQR problem, the optimum value function is unique. However, the corresponding ARE does not only have one unique solution, unless being restricted to the solution that gives rise to a stabilizing controller—which is the only solution of interest in LQR problems. Such a stabilization property is actually not required for the construction of Lyapunov functions, respectively LK functionals. However, uniqueness is required for convergence considerations from a specified sequence of finite-dimensional approximations towards a specified corresponding LK functional. Moreover, only for the unique stabilizing solution $\mathscr{P}_0$ of the operator-valued ARE, existence is guaranteed by the infinite-dimensional KYP lemma.

7.2.1.1 A Unique Solution of the Infinite-Dimensional ARE

The following uniqueness statement, which is also a part of [131, Thm. 3], was omitted in Lemma 5.5.6.

Lemma 7.2.2 (Uniqueness in the infinite-dimensional KYP lemma). *If, in Lemma 5.5.6, the requirement that $\mathscr{A} - \mathscr{B}\mathscr{K}$ generates an exponentially stable semigroup is additionally imposed, then solutions $\mathscr{X}_0$ and $\mathscr{K}$ exist if and only if $\alpha_3 > 0$, and these solutions are unique.*

Recap from Lemma 5.5.7 that $\mathscr{P}_0 = -\mathscr{X}_0$ and $\mathscr{K} = -(-\Pi_{aa})^{-1}\mathscr{B}^*\mathscr{P}_0$. Consequently, under the additionally imposed stabilization condition on $\mathscr{K}$ described in Lemma 7.2.2 above, the self-adjoint ARE solution $\mathscr{P}_0$ is uniquely defined. Thus, also $V_0(\phi)$ from (7.21) and the resulting LK functional of robust type $V(\phi) = V_0(\phi) + V_1(\phi)$ from (7.22b) are uniquely defined. In accordance with the common wording of a stabilizing ARE solution, the attribute stabilizing will henceforth be used to mark that unique LK functional of robust type.

Definition 7.2.3 (Stabilizing LK functional of robust type). *For any given LK functional V having the structure (5.5), define*[6] *$\mathcal{K}_v\colon C \to \mathbb{R}^m$;*

$$\mathcal{K}_v\phi = -(-\Pi_{aa})^{-1}[B^\top v(\phi) - \Pi_{\zeta a}^\top(\mathcal{C}\phi)] \tag{7.32}$$

based on $v(\phi)$ from (5.6). A functional $V(\phi)$ is called a stabilizing LK functional of robust type if $V(\phi)$ is an LK functional of robust type, and if the thus defined $\mathcal{K}_v$ renders the zero equilibrium of

$$\dot{x}(t) = A_0x(t) + A_1x(t-h) - B\mathcal{K}_vx_t \tag{7.33}$$

exponentially stable.

[6] Lemma 7.2.2 rather addresses $\mathscr{K}\left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right] = \mathcal{K}_v^{\mathrm{I}}\phi = -(-\Pi_{aa})^{-1}B^\top v(\phi)$ in the transformed system $\dot{x}(t) = A_0^{\mathrm{I}}x(t) + A_1^{\mathrm{I}}x(t-h) - B\mathcal{K}_v^{\mathrm{I}}x_t = A_0x(t) + A_1x(t-h) - (-\Pi_{aa})^{-1}\Pi_{\zeta a}^\top\left[\begin{smallmatrix}C_1x(t-h)\\ C_0x(t)\end{smallmatrix}\right] - B\mathcal{K}_v^{\mathrm{I}}x_t$ from Section 5.5.1. These closed loop dynamics, however, coincide with the untransformed $\dot{x}(t) = A_0x(t) + A_1x(t-h) - B\mathcal{K}_vx_t$. Moreover, $V_1(\phi)$ in $V(\phi) = V_0(\phi) + V_1(\phi)$ does not affect $v(\phi)$.

7.2.1.2 A Unique Solution of the Finite-Dimensional ARE

Consider the finite-dimensional ARE. The stabilizing ARE solution is the unique[7] smallest[8] symmetric matrix $\tilde{P}$ in terms of the Loewner partial order [120, Thm. 8.5.1]. In contrast to the infinite-dimensional KYP lemma, the limit case $W_{\tilde{G}} \succeq 0$ with $W_{\tilde{G}} \not\succ 0$ is included in Lemma 6.3.1. In that limit case, the associated controller matrix $\tilde{K}$ can lead to a closed loop matrix $\tilde{A} - \tilde{B}\tilde{K}$ that has eigenvalues on the imaginary axis. For the first part of the following statement, see [120, Thm. 7.9.3 and Thm. 8.5.1] (applied to (7.24)). For the second part, see [78, Sec. 3.6] and Lemma 7.2.2 applied to the special case of a finite-dimensional state space (cf. (7.15)).

Lemma 7.2.4 (Uniqueness in the finite-dimensional KYP lemma). *Assume $(\tilde{A}, \tilde{B})$ is stabilizable. For any $\tilde{P}$, consider the corresponding control matrix*

$$\tilde{K} = -(-\Pi_{aa})^{-1}(\tilde{B}^\top \tilde{P} - \Pi_{\zeta a}^\top \tilde{C})$$

from (7.11), and denote by σ_{cl} the set of eigenvalues of the closed loop matrix

$$\sigma_{cl} := \sigma(\tilde{A} - \tilde{B}\tilde{K}). \tag{7.34}$$

1. *If at least one solution $\tilde{P} = \tilde{P}^\top \in \mathbb{R}^{\tilde{n}\times\tilde{n}}$ of (7.12) exists, then among these solutions there is a unique one with*

$$\sigma_{cl} \subset \overline{\mathbb{C}^-}. \tag{7.35}$$

[7] A stabilizing solution is always unique [120, Prop. 7.9.2]. If the stabilizability assumption on $(\tilde{A}, \tilde{B})$ is not imposed, then an almost stabilizing solution might be non-unique, see [120, Example 3.9.1].

[8] In contrast to the minimum $\tilde{P} = \tilde{P}^\top$, a maximum $\tilde{P} = \tilde{P}^\top$ need not exist (i.e., the supremum w.r.t the Loewner partial order is not necessarily finite) if only a stabilizability but no controllability assumption is imposed, cf. [120, Example 7.9.1]. Note that, in (7.24), the linear term $\tilde{P}\tilde{A} + \tilde{A}^\top\tilde{P}$ and the quadratic term $\tilde{P}\Theta\tilde{P}$ with $\Theta = \tilde{B}(-\Pi_{aa})^{-1}\tilde{B}^\top \succeq 0$ occur with equal signs (written on the same side of the equation), whereas in the literature, frequently $\tilde{X} = -\tilde{P}$ is considered, and therefore the stabilizing solution is the maximum $\tilde{X}$.

2. *If the conditions of the KYP lemma stated in Lemma 6.3.1 hold with (6.32) being strengthened to a strict inequality, i.e.,*

$$0 < -w_{\mathbb{C}}(\tilde{C}\tilde{x}_f, -u), \tag{7.36}$$

then a unique stabilizing solution $\tilde{P} = \tilde{P}^\top$ of (7.12) exists, i.e., a unique solution with

$$\sigma_{cl} \subset \mathbb{C}^-. \tag{7.37}$$

7.2.2 Existence

A main difference between the AREs in the present context and AREs from standard LQR problems lies in the question of existence. For the former, the existence of solutions $\tilde{P} = \tilde{P}^\top \in \mathbb{R}^{\tilde{n}\times\tilde{n}}$ is a matter of the KYP lemma (Lemma 6.3.1). The KYP lemma relies on $-w(\tilde{C}\tilde{x}, -u)$, which coincides with the integrand of the costs in (7.14). As $-w_{\mathbb{C}}(\tilde{C}\tilde{x}, -u) = -\left[\begin{smallmatrix}\tilde{x}\\u\end{smallmatrix}\right]^H \left[\begin{smallmatrix}\tilde{C}^\top & 0\\0 & -I\end{smallmatrix}\right] \left[\begin{smallmatrix}\Pi_{\zeta\zeta} & \Pi_{\zeta a}\\\Pi_{\zeta a}^\top & \Pi_{aa}\end{smallmatrix}\right] \left[\begin{smallmatrix}\tilde{C} & 0\\0 & -I\end{smallmatrix}\right] \left[\begin{smallmatrix}\tilde{x}\\u\end{smallmatrix}\right]$ is an indefinite quadratic form, the KYP criterion $0 \leq -w_{\mathbb{C}}(\tilde{C}\tilde{x}_f, -u)$ with $\tilde{x}_f = (\mathrm{i}\omega I_{\tilde{n}} - \tilde{A})^{-1}\tilde{B}u$ might or might not be satisfied. In contrast, costs in standard LQR problems are nonnegative with the integrand $\left[\begin{smallmatrix}\tilde{x}\\u\end{smallmatrix}\right]^H \left[\begin{smallmatrix}Q & 0\\0 & R\end{smallmatrix}\right] \left[\begin{smallmatrix}\tilde{x}\\u\end{smallmatrix}\right]$ being a positive semidefinite quadratic form. Thus, the KYP lemma is trivially satisfied, and a real symmetric ARE solution always exists in standard LQR problems under the usual stabilizability assumption.

Therefore, the considered AREs differ from those of standard LQR problems in having a solvability bound in terms of the involved parameterization of $(\Pi_{\zeta\zeta}, \Pi_{\zeta a}, \Pi_{aa})$, which in fact amounts to the robustness bound of interest (cf. Table 5.2)—respectively, due to Lemma 5.5.1 and 7.1.2, to a transformation thereof. The following condition addresses the convergence of that solvability bound in the finite-dimensional ARE towards the solvability bound in the operator-valued ARE, as N increases. Condition 7.2.5 thus establishes that the Legendre-tau-based approximations exist for sufficiently large discretization resolutions N, whenever the functional that shall be approximated exists. Moreover, the proof

of the uniform boundedness assumption given Section 7.2.3 will also rely on this condition. The proof that the Condition 7.2.5 is indeed valid is postponed to Section 7.3, where the Legendre-tau-based transfer function is discussed.

Condition 7.2.5 (Convergence of the solvability bound). *It holds*

$$\gamma_{\max,Q}^{[N]} \to \gamma_{\max,Q} \quad \text{as } N \to \infty, \tag{7.38}$$

where $\gamma_{\max,Q}^{[N]}$ and $\gamma_{\max,Q} < \infty$ are defined as follows: Introduce a scalar factor[9] *$\gamma_Q > 0$ in the operator-valued ARE (7.20)*

$$\begin{aligned} &\langle \mathscr{P}_0 \mathscr{A}\mathsf{x}, \mathsf{x}\rangle_{M_2} + \langle \mathscr{A}^* \mathscr{P}_0 \mathsf{x}, \mathsf{x}\rangle_{M_2} \\ &+ \gamma_Q^2 \langle \mathscr{Q}\mathsf{x}, \mathsf{x}\rangle_{M_2} + \langle (-\Pi_{aa})^{-1} \mathscr{B}^* \mathscr{P}_0 \mathsf{x}, \mathscr{B}^* \mathscr{P}_0 \mathsf{x}\rangle_{\mathbb{C}^m} = 0 \end{aligned} \tag{7.39}$$

such that $\gamma_Q = 1$ gives the original equation (7.20). Then the value $\gamma_{Q,\max}$ is chosen such that a self-adjoint solution $\mathscr{P}_0$ exists if $\gamma_Q < \gamma_{Q,\max}$, whereas no such solution exists if $\gamma_Q > \gamma_{Q,\max}$. Analogously, introducing $\gamma_Q^{[N]} > 0$ in the finite-dimensional ARE (7.28)

$$\begin{aligned} &\tilde{P}_0^{[N]} \tilde{A}^{\mathrm{I}[N]} + (\tilde{A}^{\mathrm{I}[N]})^\top \tilde{P}_0^{[N]} \\ &+ (\gamma_Q^{[N]})^2 \underbrace{(\tilde{N}^{[N]})^\top \left[\begin{smallmatrix} 0 & 0 \\ 0 & Q_0+Q_1 \end{smallmatrix}\right] \tilde{N}^{[N]}}_{=:(\tilde{C}_Q^{[N]})^\top \tilde{C}_Q^{[N]}} + \tilde{P}_0^{[N]} \underbrace{\tilde{B}^{[N]}(-\Pi_{aa})^{-1}(\tilde{B}^{[N]})^\top}_{=:\tilde{B}_Q^{[N]}(\tilde{B}_Q^{[N]})^\top} \tilde{P}_0^{[N]} = 0_{\tilde{n}^{[N]}\times\tilde{n}^{[N]}} \end{aligned} \tag{7.40}$$

(where (7.28) is restored by $\gamma_Q^{[N]} = 1$), a real symmetric solution exists if and only if $\gamma_Q^{[N]} \le \gamma_{Q,\max}^{[N]}$.

The condition, i.e., the convergence (7.38) of the solvability bound, will be proven in Theorem 7.3.6. Until then it is considered as a standing assumption.

[9] By the definition of Q_0, Q_1 in (5.52), a scaling of $\mathscr{Q}$ in (5.66) amounts to $\gamma_Q^2(Q_0+Q_1) = \gamma_Q^2(C_0^\top \Pi_{\zeta\zeta}^{\mathrm{I},00} C_0 + C_1^\top \Pi_{\zeta\zeta}^{\mathrm{I},11} C_1)$ with $\Pi_{\zeta\zeta}^{\mathrm{I}} = \Pi_{\zeta\zeta} + \Pi_{\zeta a}(-\Pi_{aa})^{-1}\Pi_{\zeta a}^\top$ from (5.50a). Thus, in the case of the linear norm bound from row (I|a) in Table 5.1, where $\Pi_{\zeta\zeta} = \Pi_{\zeta\zeta}^{\mathrm{I}} = \gamma^2 I_p$, the factor γ_Q simply amounts to a scaling of the linear norm bound γ, and the maximum admissible scaling becomes $\gamma_{Q,\max} = \gamma_{\max}/\gamma$.

7.2.3 Uniform Boundedness

As a next step, the uniform boundedness assumption from Lemma 7.2.1 has to be proven. Proofs of that property for AREs from standard LQR problems in [70, 99] do not apply in the present situation: In standard LQR problems (7.7), the ARE solution $\tilde{X}^{\mathrm{LQR}} = -\tilde{P}^{\mathrm{LQR}} \succeq 0$ gives the infimum costs $x_0^\top \tilde{X}^{\mathrm{LQR}} x_0$ and thus can be upper bounded by some non-optimal costs. In contrast, in (7.14), the ARE solution is $\tilde{P} \succeq 0$ and the infimum costs are $-\tilde{x}_0^\top \tilde{P} \tilde{x}_0$, for which an upper bound only gives an upper bound on $-\tilde{P}$. The following lemma will be expedient instead. For the sake of readability, it is stated in terms of a general LTI system (A, B, C) instead of $(\tilde{A}^{\mathrm{I}[N]}, \tilde{B}_Q^{[N]}, \tilde{C}_Q^{[N]})$ from (7.40) to which it will be applied.

Lemma 7.2.6 (Upper bound on ARE solutions). *Let $A \in \mathbb{R}^{n\times n}$ be Hurwitz, and $B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{p\times n}$. Assume that the ARE*

$$PA + A^\top P + \gamma^2 C^\top C + PBB^\top P = 0_{n\times n} \tag{7.41}$$

has a solution $P = P^\top \in \mathbb{R}^{n\times n}$ if and only if $\gamma \leq \gamma_{\max}$. Let $\gamma = 1$ be the parameter choice of interest (w.l.o.g., always achievable by rescaling C), and denote by $P_{\gamma=1}$ a corresponding positive semidefinite stabilizing solution of the ARE (7.41), assuming $\gamma_{\max} > \gamma$. Let $P_{lyap} = P_{lyap}^\top$ solve the Lyapunov equation

$$P_{lyap} A + A^\top P_{lyap} + \gamma_{\max}^2 C^\top C = 0_{n\times n}. \tag{7.42}$$

Then the ARE solution $P_{\gamma=1}$ is upper bounded by

$$P_{\gamma=1} \preceq b(\gamma_{\max})\, P_{lyap}, \tag{7.43}$$

with $b\colon (1,\infty) \to \mathbb{R}_{>0}$ being a continuous function on the domain $\gamma_{\max} > \gamma = 1$.

Proof. The proof proceeds along the lines of [32, Thm. 25-1][10]. If there exists a stabilizing positive semidefinite solution P of (7.41), then this solution satisfies

$$-x_0^\top P x_0 = \inf_{u\in L_2} \int_0^\infty \big(-\gamma^2 x^\top(t) C^\top C x(t) + u^\top(t) u(t)\big)\,\mathrm{d}t \qquad (7.44)$$
$$\text{subject to } \begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ x(0) = x_0 \in \mathbb{R}^n \end{cases}$$

according to Lemma 7.1.1. Moreover, the minimum is attained by $u_{opt}(t) = -Kx_{opt}(t)$ with $K = -B^\top P$, and the closed loop dynamics $\dot{x} = (A - BK)x$ shows that $x_{opt}(t) = \mathrm{e}^{(A-BK)t}x$. Consider again the open loop system $\dot{x}(t) = Ax(t) + Bu(t)$, for which the solution $x(t) = \mathrm{e}^{At}x_0 + \int_0^t \mathrm{e}^{A(t-\eta)}Bu(\eta)\,\mathrm{d}\eta = \mathrm{e}^{At}\tilde{x}_0 + x_u(t)$ is the superposition from the autonomous system response and x_u, which denotes the response to the given input u if the initial value was zero. Thus,

$$-x_0^\top P x_0 = \int_0^\infty \Big(-\gamma^2 \big(\mathrm{e}^{At}x_0 + x_{u_{opt}}(t)\big)^\top C^\top C \big(\mathrm{e}^{At}x_0 + x_{u_{opt}}(t)\big) + u_{opt}^\top(t) u_{opt}(t)\Big)\,\mathrm{d}t. \qquad (7.45)$$

As A is assumed to be Hurwitz, the solution P_{lyap} of the Lyapunov equation (7.42) can be expressed via an improper integral. Based on the latter, introduce

$$V_{lyap} := x_0^\top P_{lyap}\, x_0 = \gamma_{\max}^2 \int_0^\infty x_0^\top (\mathrm{e}^{At})^\top C^\top C \mathrm{e}^{At} x_0\,\mathrm{d}t. \qquad (7.46)$$

[10] The ARE in [32, Thm. 25-1] is a standard-LQR-type ARE, and thus a solution always exists no matter how large the involved parameter is. Instead, the frequency-domain inequality $I_m - (G(\mathrm{i}\omega))^H G(\mathrm{i}\omega) \succeq 0_{m\times m}$ amounts to the L_2-gain of the system being less than or equal to one.

Expanding the product in (7.45) and using V_{lyap} for the autonomous part, gives

$$x_0^\top P x_0 = \frac{\gamma^2}{\gamma_{\max}^2} V_{lyap} + 2\gamma^2 \int_0^\infty x_0^\top (\mathrm{e}^{At})^\top C^\top C x_{u_{opt}}(t)\,\mathrm{d}t + \gamma^2 \int_0^\infty x_{u_{opt}}^\top(t) C^\top C x_{u_{opt}}(t)\,\mathrm{d}t - \int_0^\infty u_{opt}^\top(t) u_{opt}(t)\,\mathrm{d}t. \tag{7.47}$$

Concerning the second term in (7.47), the Cauchy–Schwarz inequality yields

$$\begin{aligned} |\langle C(\mathrm{e}^{A\cdot})x_0, C x_{u_{opt}}\rangle_{L_2}| &\le \|C(\mathrm{e}^{A\cdot})x_0\|_{L_2}\, \|C x_{u_{opt}}\|_{L_2} \\ &= \frac{1}{\gamma_{\max}^2}\sqrt{V_{lyap}}\, \|C x_{u_{opt}}\|_{L_2}. \end{aligned} \tag{7.48}$$

From the dissipativity considerations in Remark 6.3.3, it is known that

$$\|C x_{u_{opt}}\|_{L_2} \le \gamma_{\min}^{sys} \|u_{opt}\|_{L_2} = \frac{1}{\gamma_{\max}} \|u_{opt}\|_{L_2}. \tag{7.49}$$

Abbreviating $\beta = \frac{\gamma^2}{\gamma_{\max}^2}$, (7.47) becomes

$$\begin{aligned} x_0^\top P x_0 &\le \beta V_{lyap} + 2\beta\sqrt{V_{lyap}}\, \frac{1}{\gamma_{\max}} \|u_{opt}\|_{L_2} + (\beta - 1)\|u_{opt}\|_{L_2}^2 \\ &\le \sup_{z \ge 0} \left(\beta V_{lyap} + 2\frac{\beta\sqrt{V_{lyap}}}{\gamma_{\max}} z + (\beta - 1) z^2 \right). \end{aligned} \tag{7.50}$$

If $\beta < 1$, the maximum is attained at $z = \frac{\beta\sqrt{V_{lyap}}}{(1-\beta)\gamma_{\max}}$. As a result,

$$x_0^\top P x_0 \le \left(\beta + \frac{\beta^2}{(1-\beta)\gamma_{\max}^2} \right) V_{lyap}, \tag{7.51}$$

or, if $\gamma = 1$ and thus $\beta = 1/\gamma_{\max}^2$,

$$x_0^\top P x_0 \le \frac{\gamma_{\max}^4 - \gamma_{\max}^2 + 1}{\gamma_{\max}^4(\gamma_{\max}^2 - 1)} V_{lyap} =: b(\gamma_{\max}) V_{lyap} \tag{7.52}$$

for all $x_0 \in \mathbb{R}^n$. □

Lemma 7.2.7 (Uniform boundedness). *Assume the zero equilibrium of the transformed nominal RFDE (5.48) is exponentially stable, and the operator-valued ARE (5.65) has a self-adjoint stabilizing solution $\mathscr{P}_0$. Let $\mathscr{P}_0^{[N]}$ be the operator that is represented by the stabilizing matrix solution $\tilde{P}_0^{[N]}$ from the finite-dimensional ARE (7.28) with the discretization resolution N. If Condition 7.2.5 holds, then $\{\|\mathscr{P}_0^{[N]}\|\}_{N\geq N_{\min}}$ is bounded (assuming $N_{\min}$ is some sufficiently large discretization resolution).*

Proof. According to Condition 7.2.5, $\gamma_{Q,\max}^{[N]} \to \gamma_{\max,Q}$, where $\gamma_{\max,Q} > 1$ due to the assumed existence of $\mathscr{P}_0$. Thus, $\exists N_{\min} > 0, \forall N \geq N_{\min} : \gamma_{Q,\max}^{[N]} > 1$, which ensures the existence of the unique stabilizing ARE solutions $\tilde{P}_0^{[N]} = (\tilde{P}_0^{[N]})^\top$. Due to the stability preservation from Condition 3.7.1 (which is proven in [97, Thm. 5.3]), $\tilde{A}^{\mathrm{I}[N]}$ is Hurwitz if N is sufficiently large. Therefore, it follows from Lemma 6.2.7 (with $K_C = 0_{m\times p}$) that $\tilde{P}_0^{[N]}$ is positive semidefinite. According[11] to (A.112), the operator norm of $\mathscr{P}_0^{[N]}$ is related to $\tilde{P}_0^{[N]}$ by

$$\begin{aligned}\|\mathscr{P}_0^{[N]}\| &= \|((H^{[N]})^{-1/2} \otimes I_n)\tilde{P}_0^{[N]}((H^{[N]})^{-1/2} \otimes I_n)\|_2 \\ &= \sup_{\|\tilde{x}\|\leq 1} |\tilde{x}^\top((H^{[N]})^{-1/2} \otimes I_n)\tilde{P}_0^{[N]}((H^{[N]})^{-1/2} \otimes I_n)\tilde{x}| =: |\tilde{x}_{\mathrm{o}}^\top \tilde{P}_0^{[N]}\tilde{x}_{\mathrm{o}}|,\end{aligned} \tag{7.53}$$

with $H^{[N]} = (H^{[N]})^\top$ being the $(N+1)\times(N+1)$ matrix of metric coefficients depending on the basis, see, e.g., (A.21). In the following, Lemma 7.2.6 is applied to (7.40). The solution of the Lyapunov equation

$$\tilde{P}_{0,lyap}^{[N]}\tilde{A}^{\mathrm{I}[N]} + (\tilde{A}^{\mathrm{I}[N]})^\top \tilde{P}_{0,lyap}^{[N]} = -\tilde{C}_Q^{[N]}\tilde{C}_Q^{[N]} \tag{7.54}$$

[11] In (A.112), only $n = 1$ is considered. The extension to $n \geq 1$ gives rise to the Kronecker product of the metric coefficient matrix with an identity matrix in (7.53).

is positive semidefinite by Proposition 4.1.4. In particular, $\tilde{P}^{[N]}_{0,lyap}$ depends linearily on the right-hand side term of (7.54), which, in (7.42), is scaled by $(\gamma^{[N]}_{Q,\max})^2$. Thus, by (7.43), and due to the positive semidefiniteness,

$$\forall \tilde{x} \in \mathbb{R}^{\tilde{n}^{[N]}}: \quad |\tilde{x}^\top \tilde{P}^{[N]}_0 \tilde{x}| \leq b(\gamma^{[N]}_{Q,\max}) \left|\tilde{x}^\top \big((\gamma^{[N]}_{Q,\max})^2 \tilde{P}^{[N]}_{0,lyap}\big)\tilde{x}\right|. \tag{7.55}$$

Moreover, since $\gamma^{[N]}_{Q,\max}$ converges by Condition 7.2.5, and since b in (7.43) is a continuous function, $(\gamma^{[N]}_{Q,\max})^2\, b(\gamma^{[N]}_{Q,\max}) \leq \bar{b}$ is bounded by some $\bar{b} > 0$, which leads to the conclusion

$$\forall \tilde{x} \in \mathbb{R}^{\tilde{n}^{[N]}}: \quad |\tilde{x}^\top \tilde{P}^{[N]}_0 \tilde{x}| \leq \bar{b}\, |\tilde{x}^\top \tilde{P}^{[N]}_{0,lyap} \tilde{x}|. \tag{7.56}$$

The represented self-adjoint operators accordingly satisfy

$$\|\mathscr{P}^{[N]}_0\| \leq \bar{b}\, \|\mathscr{P}^{[N]}_{0,lyap}\|, \tag{7.57}$$

for which boundedness follows from Lemma 4.6.6. □

7.2.4 Convergence Statement

Based on these results, a convergence statement for the numerical approach proposed in Section 6.1 can be derived, provided Condition 7.2.5 is valid. The statement is analogous to what has been proven for the Legendre-tau-based approximation of LK functionals of complete type in Section 4.6.2.

Theorem 7.2.8. *Let the conditions in Theorem 5.5.11 hold and denote by $V(\phi)$ the thus existing unique stabilizing (see Definition 7.2.3) LK functional of robust type. For any given $\phi \in C$ and any discretization resolution N, let*

$$\tilde{x}^{[N]} = \tilde{\pi}^{[N]}(\phi) \in \mathbb{R}^{\tilde{n}^{[N]}} = \mathbb{R}^{n(N+1)} \tag{7.58}$$

denote the discretization of ϕ according to (3.47) (respectively a coordinate transformation thereof). Moreover, let

$$\tilde{V}^{[N]}(\tilde{x}^{[N]}) = (\tilde{x}^{[N]})^\top \tilde{P}^{[N]} \tilde{x}^{[N]} \tag{7.59}$$

be built from the stabilizing solution $\tilde{P}^{[N]}$ of the ARE (6.18) with $\tilde{E} = 0$. Assume Condition 7.2.5 holds (which, in fact, will be proven in Theorem 7.3.6 below), then

$$\forall \phi \in C: \quad \tilde{V}^{[N]}(\tilde{\pi}^{[N]}(\phi)) \to V(\phi), \quad (N \to \infty). \tag{7.60}$$

Proof. According to Lemma 7.1.3 and Lemma 7.1.2 from Section 7.1.3.2, the splitting $V = V_0 + V_1$ can be considered. Convergence towards V_0 follows from the weak operator convergence $\mathscr{P}_0^{[N]} \overset{\text{weakly}}{\to} \mathscr{P}_0$ stated in Lemma 7.2.1. The latter applies due to the uniqueness statements from Lemma 7.2.2 and Lemma 7.2.4, the existence statements from the presupposed conditions in Theorem 5.5.11 and Condition 7.2.5, and the uniform boundedness statement from Lemma 7.2.7. Convergence towards V_1 in (7.22b) has already been proven in Theorem 4.6.8. □

Due to the KYP lemma, the not yet proven Condition 7.2.5 on the maximum admissible additional gain γ_Q translates to a frequency-domain condition. Let, in accordance with $\tilde{C}_Q^{[N]}$ and $\tilde{B}_Q^{[N]}$, which are already introduced in the matrix-valued ARE (7.40) for $\tilde{P}_0^{[N]}$, the underlying terms in $\mathscr{Q}$ and $\mathscr{B}$ of the operator-valued ARE (7.39) be decomposed as

$$Q_0 + Q_1 =: C_{0,Q}^\top C_{0,Q} \tag{7.61}$$

(which occurs in $\mathscr{Q}\left[\begin{smallmatrix}\phi\\ r\end{smallmatrix}\right] = \begin{bmatrix} 0_{L_2} \\ (Q_0+Q_1)r \end{bmatrix}$) and

$$B(-\Pi_{aa})^{-1}B^\top =: B_Q B_Q^\top. \tag{7.62}$$

Then Corollary 5.5.13 gives for the maximum admissible gain in Condition 7.2.5

$$\gamma_{Q,\max} = \frac{1}{\|G_Q\|_\infty}, \tag{7.63}$$

relying on the time-delay-system transfer function

$$G_Q(s) := C_{0,Q}\,(sI_n - A_0^{\mathrm{I}} - A_1^{\mathrm{I}}\mathrm{e}^{-sh})^{-1}B_Q. \tag{7.64}$$

Analogously, Corollary 6.3.2 gives $\gamma_{Q,\max}^{[N]} = 1/\|\tilde{G}_Q^{[N]}\|_\infty$ relying on the Legendre-tau-based approximation

$$\tilde{G}_Q^{[N]}(s) := \tilde{C}_Q^{[N]}(sI_{\tilde{n}^{[N]}} - \tilde{A}^{\mathrm{I}[N]})^{-1}\tilde{B}_Q^{[N]}. \tag{7.65}$$

As a result, Condition 7.2.5 only requires that the H_∞-norm converges,

$$\|\tilde{G}_Q^{[N]}\|_\infty \to \|G_Q\|_\infty \quad \text{as } N \to \infty, \tag{7.66}$$

which motivates a closer look on the transfer function generated from the Legendre-tau-based ODE approximation of a time-delay system. The considerations will directly address general transfer functions (5.77), which include the splitting-related transfer function (7.64) as a special case.

7.3 The Legendre-Tau-Based Transfer Function

Let $G(s)$ be an RFDE transfer function (5.77). In this section, the equivalence between the transfer function $\tilde{G}(s)$ from the Legendre-tau-based approximation (6.13) on the one hand and the common rational approximation of $G(s)$ that

results from replacing e^{-sh} by an all-pass Padé approximation[12] on the other hand is proven. It is known that an analogous equivalence also holds for the characteristic equation, see [97, Thm. 3.1]. Compared to the proof of the latter, a useful relation between the Padé approximation of e^{-sh} and Bessel polynomials will simplify the considerations in this section. Note that the considered time-delay system (6.3), with $G(s)$ given by (5.77), involves, besides of the state delay, also delayed output terms $C_1 x(t-h)$, alongside undelayed output terms $C_0 x(t)$.

7.3.1 The Padé Approximation of Order $[N/N]$ for e^{-hs}

Let $\text{padé}_{[N/N]}(\mathrm{e}^z)$ denote the Padé approximant of order $[N/N]$ for the exponential function, i.e., its Padé approximant with equal numerator and denominator polynomial degree, which is also known as diagonal or all-pass Padé approximant. The common formula, see [153, §75 / p. 431],

$$\text{padé}_{[N/N]}(\mathrm{e}^z) = \frac{\upsilon_N(z)}{\upsilon_N(-z)} \tag{7.67}$$

is based on the polynomials

$$\begin{aligned}\upsilon_N(z) &:= \sum_{k=0}^{N} \frac{(2N-k)!}{(2N)!} \binom{N}{k} z^k, \qquad \text{with } \binom{N}{k} = \frac{N!}{k!(N-k)!} \\ &= 1 + \tfrac{1}{2N}\tfrac{N}{1}\, z + \tfrac{1}{(2N)(2N-1)}\tfrac{N(N-1)}{2\cdot 1}\, z^2 + \tfrac{1}{(2N)(2N-1)(2N-2)}\tfrac{N(N-1)(N-2)}{3\cdot 2\cdot 1}\, z^3 \\ &\quad + \ldots + \tfrac{1}{(2N)(2N-1)\cdots(N+1)}\tfrac{N!}{N!}\, z^N.\end{aligned} \tag{7.68}$$

However, it can also be expressed via Bessel polynomials, cf. [34, eq. (2.8)].

[12] If the description "Padé approximation of a model with time delay" (e.g., in Matlab) is used, it is just this substitution that is meant, although, strictly spoken, substituting in $G(s)$ the individual exponential terms by their Padé approximation does not coincide with the Padé approximation of the overall transfer function $G(s)$.

Lemma 7.3.1. *Let $\theta_N(z)$ denote the N-th reverse Bessel polynomial and $b_N(z)$ the N-th Bessel polynomial, related by $\theta_N(z) = z^N b_N(\frac{1}{z})$, see (7.69) and (7.70) below. The numerator polynomial of* $\text{padé}_{[N/N]}(\mathrm{e}^z)$ *can be written as*

$$\upsilon_N(z), \quad \textit{according to (7.68) given by} \quad \upsilon_N(z) \stackrel{\text{def}}{=} \sum_{k=0}^{N} \frac{\frac{1}{k!}\frac{N!}{(N-k)!}}{\frac{(2N)!}{(N+(N-k))!}} z^k$$

$$= \frac{N!}{(2N)!} 2^N \, \theta_N(\tfrac{1}{2}z), \qquad \theta_N(z) \stackrel{\text{def}}{=} \sum_{j=0}^{N} \frac{1}{j!}\frac{(N+j)!}{(N-j)!}\frac{z^{N-j}}{2^j} \tag{7.69}$$

$$= \frac{N!}{(2N)!} 2^N \, (\tfrac{1}{2}z)^N b_N((\tfrac{1}{2}z)^{-1}), \qquad b_N(z) \stackrel{\text{def}}{=} \sum_{j=0}^{N} \frac{1}{j!}\frac{(N+j)!}{(N-j)!}\big(\tfrac{z}{2}\big)^j. \tag{7.70}$$

Proof. By inspection. □

Consequently, with the exponent $z = -hs$, which is relevant in the transfer function of time-delay systems, the following alternative descriptions evolve.

Lemma 7.3.2. *The Padé approximant of* e^{-hs} *can equivalently*[13] *be described via*

$$\text{padé}_{[N/N]}(\mathrm{e}^{-hs}) = \frac{\upsilon_N(-hs)}{\upsilon_N(hs)} = \frac{\theta_N(-\frac{h}{2}s)}{\theta_N(\frac{h}{2}s)} = (-1)^N \frac{b_N(-(\frac{h}{2}s)^{-1})}{b_N((\frac{h}{2}s)^{-1})}. \tag{7.71}$$

[13] ignoring the removable singularity at $s = 0$ that is induced by the reciprocal arguments in b_N

Example 7.3.3. *To get a better impression of the relations, consider* $N = 4$*, where* $\frac{N!}{(2N)!}2^N = \frac{1}{105}$ *in (7.69) and* $(\frac{1}{2}z)^N = (-\frac{h}{2}s)^N = (\frac{h}{2}s)^4$ *in (7.70). The numerators in (7.71) become*

$$\upsilon_4(-hs) = 1 - \tfrac{1}{2}hs + \tfrac{3}{28}(hs)^2 - \tfrac{1}{84}(hs)^3 + \tfrac{1}{1680}(hs)^4 \tag{7.72}$$

$$\theta_4(-\tfrac{h}{2}s) = (\tfrac{h}{2}s)^4 - 10(\tfrac{h}{2}s)^3 + 45(\tfrac{h}{2}s)^2 - 105\tfrac{h}{2}s + 105 \tag{7.73}$$

$$\begin{aligned} b_4(-(\tfrac{h}{2}s)^{-1}) &= 105(\tfrac{h}{2}s)^{-4} - 105(\tfrac{h}{2}s)^{-3} + 45(\tfrac{h}{2}s)^{-2} - 10(\tfrac{h}{2}s)^{-1} + 1 \\ &= 1680\tfrac{1}{h^4s^4} - 840\tfrac{1}{h^3s^3} + 180\tfrac{1}{h^2s^2} - 20\tfrac{1}{hs} + 1 \end{aligned} \tag{7.74}$$

and the resulting Padé approximation is

$$\text{padé}_{[4/4]}(\mathrm{e}^{-hs}) = \frac{h^4s^4 - 20h^3s^3 + 180h^2s^2 - 840hs + 1680}{h^4s^4 + 20h^3s^3 + 180h^2s^2 + 840hs + 1680}. \tag{7.75}$$

7.3.2 Legendre-Tau ODE as a State Space Realization of a Padé Approximation

The following proof relies on a relation between Bessel polynomials and a certain three-banded system of linear equations that is derived in Lemma B.3.2 in the appendix.

Theorem 7.3.4. *The transfer function from the Legendre-tau-based discretization*

$$\tilde{G}(s) = \tilde{C}(sI_{\tilde{n}} - \tilde{A})^{-1}\tilde{B}$$

described in (6.13) is equal to the RFDE transfer function with e^{-hs} *being replaced by its Padé approximant of order* $[N/N]$

$$\tilde{G}(s) = \begin{bmatrix} C_1\,\text{padé}_{[N/N]}(\mathrm{e}^{-hs}) \\ C_0 \end{bmatrix} \Big[sI_n - A_0 - \text{padé}_{[N/N]}(\mathrm{e}^{-hs})A_1 \Big]^{-1} B. \tag{7.76}$$

Proof. Introducing $R_c \in (\mathbb{R}(s))^{\tilde{n}\times m}$, the transfer function (6.13) results from

$$(sI_{\tilde{n}} - A_c)R_c(s) = B_c \tag{7.77a}$$

$$\tilde{G}(s) = C_c R_c(s). \tag{7.77b}$$

Based on

$$\beta_j(z) := (2j+1)z, \quad \text{and} \quad \gamma_k(z) := -\tfrac{k(k+1)}{2}z, \tag{7.78}$$

$j \in \{0,\dots,N-1\}$, $k \in \{0,\dots,N\}$, the system matrix (3.28) can be written as

$$A_c = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & & 0 & 0 \\ A_0{+}A_1 & A_0{-}A_1 & \cdots & A_0{+}A_1 & A_0{-}A_1 \end{bmatrix} + \begin{bmatrix} 0 & \beta_0(\frac{2}{h}) & 0 & \beta_0(\frac{2}{h}) & 0 & \cdots & \beta_0(\frac{2}{h}) & 0 \\ & 0 & \beta_1(\frac{2}{h}) & 0 & \beta_1(\frac{2}{h}) & \cdots & 0 & \beta_1(\frac{2}{h}) \\ & & 0 & \ddots & \ddots & & & \\ & & & & \ddots & & \ddots & \\ & & & & & \ddots & & \\ & & & & & 0 & \beta_{N-2}(\frac{2}{h}) & 0 \\ & & & & & & 0 & \beta_{N-1}(\frac{2}{h}) \\ \gamma_0(\frac{2}{h}) & \gamma_1(\frac{2}{h}) & \gamma_2(\frac{2}{h}) & \cdots & & \cdots & \gamma_{N-1}(\frac{2}{h}) & \gamma_N(\frac{2}{h}) \end{bmatrix} \otimes I_n \tag{7.79}$$

(exemplary shown for N even[14]). Introduce $R_b \in (\mathbb{R}(s))^{\tilde{n}\times m}$ such that

$$R_c(s) = T_{cb}R_b(s), \quad T_{cb} = \begin{bmatrix} 1 & 0 & -1 & & & \\ & 1 & 0 & -1 & & \\ & & 1 & & \ddots & \\ & & & \ddots & & -1 \\ & & & & & 0 \\ & & & & & 1 \end{bmatrix} \otimes I_n. \tag{7.80}$$

[14] The shown structure in the right upper part corresponds to N even, otherwise the last right-upper side-diagonals are nonzero, zero, and nonzero.

As a consequence, (7.77) transforms to

$$(sI_{\tilde{n}} - A_c)T_{cb}R_b(s) = B_c \tag{7.81a}$$
$$\tilde{G}(s) = C_c T_{cb} R_b(s). \tag{7.81b}$$

Concerning the last row of $A_c T_{cb}$, note that, for $k \geq 2$,

$$\begin{aligned}\gamma_k(\tfrac{2}{h}) - \gamma_{k-2}(\tfrac{2}{h}) &= (-\tfrac{k(k+1)}{2} + \tfrac{(k-2)(k-1)}{2})\tfrac{2}{h} \\ &= -\beta_{k-1}(\tfrac{2}{h}).\end{aligned} \tag{7.82}$$

As a result, (7.81a) becomes

$$\left(s \begin{bmatrix} 1 & -\beta_0(\frac{1}{s}\frac{2}{h}) & -1 & & & & \\ & 1 & -\beta_1(\frac{1}{s}\frac{2}{h}) & -1 & & & \\ & & 1 & \ddots & \ddots & & \\ & & & & & -1 & \\ & & & & 1 & -\beta_{N-1}(\frac{1}{s}\frac{2}{h}) & \\ 0 & \beta_0(\frac{1}{s}\frac{2}{h}) & \beta_1(\frac{1}{s}\frac{2}{h}) & & & (1+\beta_{N-1}(\frac{1}{s}\frac{2}{h})) & \end{bmatrix} \otimes I_n \right. \\ \left. - \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & & 0 & 0 \\ A_0+A_1 & A_0-A_1 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \right) R_b(s) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ B \end{bmatrix}, \tag{7.83}$$

and (7.81b) becomes

$$\tilde{G}(s) = \begin{bmatrix} C_1 & -C_1 & 0_{p_1\times n} & \cdots & 0_{p_1\times n} \\ C_0 & C_0 & 0_{p_0\times n} & \cdots & 0_{p_0\times n} \end{bmatrix} R_b(s) = \begin{bmatrix} C_1\big(R_b^0(s) - R_b^1(s)\big) \\ C_0\big(R_b^0(s) + R_b^1(s)\big) \end{bmatrix}. \tag{7.84}$$

Ignoring the last row in (7.83), Lemma B.3.2 shows that

$$R_b^0(s) + R_b^1(s) = b_N(\tfrac{1}{s}\tfrac{2}{h})R_b^N(s) \tag{7.85}$$
$$R_b^0(s) - R_b^1(s) = (-1)^N b_N(-\tfrac{1}{s}\tfrac{2}{h})R_b^N(s). \tag{7.86}$$

To obtain R_b^N, all rows of (7.83) are added

$$s(R_b^0 + R_b^1) - A_0(R_b^0 + R_b^1) - A_1(R_b^0 - R_b^1) = B, \tag{7.87}$$

yielding with (7.85) and (7.86)

$$R_b^N(s) = \left[b_N(\tfrac{1}{s}\tfrac{2}{h})sI_n - b_N(\tfrac{1}{s}\tfrac{2}{h})A_0 - (-1)^N b_N(-\tfrac{1}{s}\tfrac{2}{h})A_1\right]^{-1} B. \tag{7.88}$$

Consequently, (7.84) with (7.85) and (7.86) gives

$$\tilde{G}(s) = \tag{7.89}$$
$$\begin{bmatrix}(-1)^N C_1 b_N(-\frac{1}{s}\frac{2}{h}) \\ C_0 b_N(\frac{1}{s}\frac{2}{h})\end{bmatrix} \left[b_N(\tfrac{1}{s}\tfrac{2}{h})sI_{\tilde{n}} - b_N(\tfrac{1}{s}\tfrac{2}{h})A_0 - (-1)^N b_N(-\tfrac{1}{s}\tfrac{2}{h})A_1\right]^{-1} B,$$

which by (7.71) completes the proof. □

7.3.3 Completing the Proof of Convergence for the Numerical Approach to the LK Functional

As a consequence, existing results on the convergence of these common rational approximations for time-delay transfer functions apply, see, e.g., [71]. It is well known that, as the order N increases, the error bound[15] for the Padé approximation of $\mathrm{e}^{-\mathrm{i}\omega h}$ goes towards zero on an increasing range $|\omega| < \nu N$ of frequencies (with $\nu \approx 1.4426\frac{1}{h}$), whereas, beyond this linearly growing frequency window, the error bound[16] does not improve until being reached by the window. Nevertheless, for the strictly proper overall transfer function $G(s)$, the asymptotic behavior $\|G(\mathrm{i}\omega)\|_\infty \to 0$, as $\omega \to \infty$, is retained when substituting the exponential terms by their Padé approximations. In the end, the derived rational approximation of

[15] if $|\omega| < \nu N : |\mathrm{e}^{-\mathrm{i}\omega h} - \text{padé}(\mathrm{e}^{-\mathrm{i}\omega h})| \leq 2|\omega|^{2N+1}(\nu N)^{-(2N+1)}$ with $\nu = 2(\frac{\sqrt{2}}{\mathrm{e}})^{\frac{1}{2}}\frac{1}{h}$, see [71, Prop. 3.3]

[16] if $|\omega| \geq \nu N : |\mathrm{e}^{-\mathrm{i}\omega h} - \text{padé}(\mathrm{e}^{-\mathrm{i}\omega h})| \leq 2$, see [71, Prop. 3.3].

$G(\mathrm{i}\omega)$ converges uniformly on the whole unbounded domain $\omega \in (-\infty, \infty)$ to the original $G(\mathrm{i}\omega)$ as the order N increases (presuming there are no poles on the imaginary axis).

Lemma 7.3.5 (Convergence of the approximated transfer function, [71, Prop. 4.2]). *Consider $G(s)$ from (5.77), assuming that $\dot{x}(t) = A_0 x(t) + A_1 x(t-h)$ has no characteristic roots on the imaginary axis, and let $\tilde{G}^{[N]}(s)$ denote its rational approximation (7.76). Then*[17] $\sup_{\omega\in\mathbb{R}} \|\tilde{G}^{[N]}(\mathrm{i}\omega) - G(\mathrm{i}\omega)\|_2 \to 0$ *as* $N \to \infty$.

Proof. [71, Prop. 4.2], which, if $m > 1$ or $p > 1$, can elementwise be applied. □

Thus the asymptotic transition of $\tilde{G}^{[N]}(\mathrm{i}\omega)$ in Corollary 6.3.2 from the finite-dimensional KYP lemma to $G(\mathrm{i}\omega)$ in Table 5.2 is established. In particular, the lemma can be used to prove the convergence of the splitting-based solvability bound described by Condition 7.2.5.

Theorem 7.3.6. *Assume that $\dot{x}(t) = f^{\mathrm{I}}(x_t)$ defined in (5.48) has an exponentially stable equilibrium and $G_Q(s) \not\equiv 0$ in (7.64). Then Condition 7.2.5 holds.*

Proof. In Condition 7.2.5, $\gamma_{Q,\max} = \frac{1}{\|G_Q\|_\infty}$ can be derived from $G_Q(s)$ defined in (7.64) due to Corollary 5.5.13. The above Lemma 7.3.5 and the reverse triangle inequality in

$$\left|\|G_Q^{[N]}\|_\infty - \|G_Q\|_\infty\right| \leq \|G_Q^{[N]} - G_Q\|_\infty \to 0, \quad \text{as } N \to \infty \tag{7.90}$$

confirm (7.66), which concludes the proof. □

Note that the involved stability assumption on $\dot{x}(t) = f^{\mathrm{I}}(x_t)$ (which is not restrictive, see Remark 5.5.10), is already presumed in Theorem 7.2.8. Thus, the

[17] $\sup_{\omega\in\mathbb{R}} \|E(\mathrm{i}\omega)\|_2 = \|E\|_\infty$ is the H_∞-norm if $E \in H_\infty$ and otherwise still the peak gain (L_∞-norm, not to be confused with the peak-to-peak or induced L_∞-gain of the system).

provisionally imposed Condition 7.2.5 in Theorem 7.2.8 is always valid and can be dropped.

As a result, Theorem 7.3.6 completes the proof of the convergence statement in Theorem 7.2.8, which is

$$\forall \phi \in C: \quad \tilde{V}^{[N]}(\tilde{\pi}^{[N]}(\phi)) \to V(\phi), \quad (N \to \infty), \tag{7.91}$$

without imposing additional assumptions. Consequently, the convergence, which for the Legendre-tau-based approximation of LK functionals of complete type has already be shown in Section 4.6.2 of Chapter 4, also holds for the Legendre-tau-based approximation of the LK functionals of robust type.

7.4 Revisiting the Main Points of the Chapter

- The ARE is associated to an LQR problem. Therefore, the convergence proof can benefit from known results on a Legendre-tau-based approach to optimal control problems in time-delay systems.
- However, there are two major differences to standard LQR problems.
  - First, the associated LQR problem is an LQR problem with indefinite costs, whereas, in standard LQR problems, costs are always nonnegative.
  - Second, the matrix-valued ARE actually addresses an LQR problem with costs that involve delayed terms. However, this issue is already resolved by the splitting approach from Section 5.5.1, on which the proofs in this chapter rely.
- Because of the indefinite costs, aspects like the existence of solutions and the uniform boundedness of solutions have to be proven differently than for the AREs from standard LQR problems (considering the sequence of finite-dimensional AREs that arise by increasing the discretization resolution N).

- For the considered LTI system (6.3) with a state delay and both delayed and non-delayed output terms, it is shown that the Legendre-tau-based finite-dimensional LTI system (6.10) is a state space realization of the common Padé-based rational approximation of the transfer function $G(s)$. Therefore, known results on these wide-spread rational approximations apply. They ensure that if the frequency-domain existence condition for the LK functional of robust type $V(\phi)$ is satisfied, then the approximation $\tilde{V}(\tilde{x})$ of the LK functional also exists if N is sufficiently large.
- The obtained convergence statement (Theorem 7.2.8) equals the one already obtained for complete-type LK functionals in Section 4.6.2: Under the imposed assumptions, for any given argument $\phi \in C$, it is proven that the resulting $\tilde{V}(\tilde{x}) = \tilde{x}^\top \tilde{P} \tilde{x}$ (where $\tilde{x}$ represents coordinates of the polynomial approximation of ϕ in the chosen polynomial basis) converges to the value of the LK functional of robust type $V(\phi)$ as the discretization resolution N increases.

8 Conclusion

The guiding theme of the present thesis is given by the simple delay-free template from Section 1.2. The template breaks down the known approach of complete-type Lyapunov–Krasovskii (LK) functionals to its counterpart in delay-free ODEs, namely the well-known construction of Lyapunov functions from a Lyapunov equation.

The numerical approach to complete-type LK functionals that is proposed in Section 4.1 relies on an ODE approximation of the time-delay system. Only a finite-dimensional Lyapunov equation has to be solved. As a consequence, it brings the template (the Lyapunov-equation-based construction of a Lyapunov function for an ODE) and its counterpart (the complete-type LK functional for a time-delay system) closer together. In fact, the approach merges two seemingly disjunctive schools concerning stability in linear time-invariant time-delay systems: constructing an LK functional versus numerically computing characteristic roots via discretization (to be more precise, computing eigenvalues of a matrix that actually belongs to an ODE approximation). Indeed, the presented finite-dimensional approximation of the LK functional is shown to be a partial Lyapunov function that proves partial asymptotic stability of the ODE approximation. For a suitably chosen approximation scheme, the latter in turn means that all eigenvalues of the ODE system matrix have negative real parts.

Additionally, the numerical approach opens up new freedoms in how to construct the defining equation of the LK functional. Clearly, at the level of the involved ODE, a more general Lyapunov equation could be solved as well, without requiring any additional effort. At the same time, the numerical approach makes the importance of the structure of the problem apparent, thus calling for a construction

that can incorporate structural information. This is achieved by replacing the Lyapunov-equation template by an algebraic-Riccati-equation template. From the theory on absolute stability in ODEs, algebraic Riccati equations are known to be more appropriate when questions of robustness shall be tackled. The reader might rather associate LMIs or frequency-domain methods with robustness theory for ODEs, which is why the thesis also intends to explain the relations, and known equivalences, in a concise manner. Most importantly, a time-delay counterpart to the template of an algebraic-Riccati-equation-based Lyapunov function does not exist in the same spirit as complete-type LK functionals represent a time-delay counterpart to the template of a Lyapunov-equation-based Lyapunov function. The newly introduced Lyapunov–Krasovskii functionals of robust type fill that gap.

8.1 Summary of the Contributions

In view of the objectives declared in Section 1.5, the following has been achieved.

I The first part of the thesis (Chapter 2) is devoted to the issue of how to prove stability in a linear time-delay system.

- A necessary and sufficient criterion for delay-independent exponential stability in $\dot{x}(t) = A_0x(t) + A_1x(t-h)$ is proposed (Theorem 2.3.4). Merely the spectral abscissa of $M(\varphi) = A_0 + \mathrm{e}^{\mathrm{i}\varphi}A_1$ must be plotted over $\varphi \in [0, \pi]$.
- For delay-dependent stability, a constrained minimization problem is formulated, which describes the exact critical delay value (Theorem 2.2.2).

- An underlying framework of three possible perspectives on the two-variable formulation of the characteristic equation is introduced (Section 2.1.2). In particular, this framework makes the parallels between the proposed criteria and the wide-spread so-called frequency-sweeping tests apparent.

II The second part (Chapter 3 and 4) is devoted to the issue of how to compute Lyapunov–Krasovskii functionals of complete type numerically.

- A numerical approach is proposed that relies on an ODE approximation of the time-delay system. Only a finite-dimensional Lyapunov equation must be solved to obtain an approximation of the complete-type Lyapunov–Krasovskii functional (Section 4.1.1).
- A formula for the coefficient of a quadratic lower bound on the numerical result of the Lyapunov–Krasovskii functional is derived (Theorem 4.3.2). Examples confirm that this lower-bound coefficient is significantly less conservative than known formulas (Table 4.1).
- For both the value of the functional as well as the lower-bound coefficient convergence is proven (Theorem 4.6.8 and Theorem 4.6.10).
- An insightful finding is that the Lyapunov–Krasovskii theorem is the counterpart of a theorem for Lyapunov–Rumyantsev partial asymptotic stability in ODEs (Section 4.5). The proposed finite-dimensional approximation of the Lyapunov–Krasovskii functional is shown to be a partial Lyapunov function for the underlying finite-dimensional ODE (Theorem 4.5.12).
- Basic concepts from tensor algebra are exploited to make precise in which sense the involved matrices represent operators and how to handle these coordinate representations (Appendix A).

III The final part (Chapter 5, 6, 7) is devoted to a new class of Lyapunov–Krasovskii functionals.

- Lyapunov–Krasovskii functionals of robust type are proposed (Definition 5.2.4). Instead of relying on an arbitrarily chosen desired derivative of the functional along solutions of the nominal system, the proposed construction incorporates decisive information about the nonlinear term that perturbs the linear nominal system: the structure of the injection into the system and a sector restriction within which the nonlinearity is allowed to reside.
- Important properties of these functionals are proven (Theorem 5.3.2 and Theorem 5.4.1).
- The robustness statements that can be obtained from these functionals are derived (Theorem 5.5.11, Corollary 5.5.13, Corollary 5.5.14, and Corollary 5.5.17). The results are significantly less restrictive than the ones arising from Lyapunov–Krasovskii functionals of complete type (Table 5.3).
- A numerical approach for Lyapunov–Krasovskii functionals of robust type is proposed, which is along the lines of the proposed approach for Lyapunov–Krasovskii functionals of complete type. Instead of the matrix-valued Lyapunov equation, a matrix-valued algebraic Riccati equation must be solved (Section 6.1.3).
- Concerning the underlying algebraic Riccati equation, various interrelations are clarified (Section 6.2 and Section 6.3).
- A convergence statement on the proposed numerical approach is proven (Theorem 7.2.8).

8.2 Outlook

Various further developments and generalizations are in order.

- Regional stability statements in terms of an estimation of the domain of attraction: If the nonlinearity resides only locally within the considered sector, then, based on a (restricted) sublevel set of the Lyapunov–Krasovskii (LK) functional, a guaranteed subset of the domain of attraction can be concluded. If $C_0 = C_1 = I_n$, the approach that is known from complete-type LK functionals can be adopted directly to LK functionals of robust type. Still, it remains to discuss the more general cases.

- Using the numerical approach to derive other bounds on the LK functional (for both LK functionals of complete and robust type): So far, the numerical approach is only used for the coefficient of the global quadratic lower bound on the LK functional. This coefficient is needed in the above mentioned estimation of the domain of attraction. However, already if, for the sake of comparability, a norm ball within the domain of attraction shall be given, a quadratic upper bound on the LK functional is required as well. Besides, in more general cases where no quadratic lower bound exists, an estimation of the domain of attraction will be based on the quadratic lower bound on a Razumikhin-like set. Such types of bounds can also be derived from the finite-dimensional approximation of the LK functional.

- Application to further example systems and comparison with results from complete-type LK functionals: The example in Section 5.6 relies on a linear norm bound, which makes the result well comparable to the results from complete-type LK functionals. However, in many applications, the additional freedom in choosing the perturbation restriction that is offered by LK functionals of robust type is expected to come along with further improvements.

- Generalizations of the presented results to more general system classes:

  - The necessary and sufficient criterion for delay-independent stability as well as the constrained minimization problem for delay-dependent stability turn out to be particularly insightful when it comes to multiple

incommensurate delays, multiple commensurate delays and perturbations thereof.

  - The concept of LK functionals of robust type should be generalized in various directions. Note that the concept of complete-type LK functionals has been extended in the literature, for instance, to

    * RFDEs with multiple commensurate delays,
    * RFDEs with multiple incommensurate delays,
    * RFDEs with distributed delays, and
    * neutral functional differential equations (NFDEs),

    see, e.g., the monograph [110].

  - The numerical approach also has to be considered for these more general system classes, which concerns both the approximation of LK functionals of robust type and the approximations of LK functionals of complete type.

- Using the LK functional of robust type as an initialization in an iterative approach to improved estimations of the domain of attraction: A truncation of the obtained LK functional of robust type could be combined with a sum-of-squares (SOS)-based approach.
- Further development of LK functionals of robust type: The mature theory on absolute stability for ODEs comes along with many developments that motivate various possible extensions of the proposed concept.
- Coping with time-varying time delays: Results based on LK functionals of complete type could be adopted to LK functionals of robust type.

Moreover, the theory of absolute stability is closely related to integral-quadratic-constraints (IQCs), which, however, usually only cope with finite-dimensional nominal systems and consider the overall delay as an uncertainty. Incorporating infinite-dimensional nominal systems in IQC-based approaches appears to be promising.

A Appendix: A Tensorial Point of View on Polynomials

The present thesis draws heavily on coordinate representations of polynomials and operators: be it, e.g., Legendre coordinates $c(t)$ that represent a polynomial approximation of the state x_t, be it the corresponding system matrix A_c that represents an operator $\mathscr{A}^{[N]}$, or be it the resulting matrix P_c from a Lyapunov or algebraic Riccati equation that represents an operator $\mathscr{P}^{[N]}$. The present appendix focuses on the results that arise from the Legendre tau method. The objective is to clarify how the operators are precisely defined by their coordinate representations and how to compute inner products, norms, operator norms, or adjoints from the coordinates, taking into account that the considered bases of the arising polynomial subspaces of M_2 are nonorthonormal.

Clearly, the coordinate representation of a polynomial can be transformed by changing the basis, see Section 3.6, while the represented object—the polynomial—remains the same. This, however, is exactly what tensor calculus is about. Since polynomials of degree at most N form a finite-dimensional vector space, basic concepts from tensor algebra, like metric coefficients, dual bases and dyadic products, can be applied to polynomials in M_2, which is the leading idea in the present appendix.

The present appendix is organized as follows. First, **Appendix A.1** introduces the basis functions that are associated with Legendre coordinates and with mixed coordinates for the considered polynomial subspaces of M_2. Then, **Appendix A.2.1**

explains how to compute inner products, how to represent operators in that finite-dimensional polynomial subspace, and how the adjoint looks like. Finally, a projection operator is incorporated in **Appendix A.3** such that the finite-dimensional domain is left and the operators $\mathscr{A}^{[N]}$ and $\mathscr{P}^{[N]}$ that occur in the main part of the present thesis can be described. **Appendix A.4** revisits important results of this appendix.

To simplify the considerations for readers not familiar with tensor calculus, the present appendix employs two deviations from the common notation [180, p. 14-15], [128, 86, 37].

- Summation symbols are explicitly given instead of using Einstein's summation convention. According to the latter, any index that occurs two times in a term would be recognized to belong to a summation without writing the summation symbol. Nevertheless, looking at the occurring sums, the reader will probably quickly understand why that convention is convenient.
- An underline is added to variables that deviate from the standard coordinates or basis elements used in the main part of this thesis. That is, the dual basis and its coordinates are underlined. In the usual index notation, the original basis and its (contravariant) coordinates are only distinguished from the dual basis and its (covariant) coordinates by the position of the indices. These index positions are still also employed—which is why the reader will probably quickly understand that the underline is actually a redundant information in a pure index notation.

Moreover, for the sake of notational compactness, the present section only considers scalar polynomials,

$$n = 1. \tag{A.1}$$

Nevertheless, the results can straightforwardly be extended to vector-valued polynomials.

A.1 Basis Functions for Polynomials in M_2

The present section discusses for Legendre coordinates c and for mixed coordinates χ the corresponding bases in M_2 and their metric coefficients. Both interpretations encountered in Section 3.5.2.3 are discussed: the standard continuous interpretation, where the Legendre coordinates are understood as coefficients of the 0-th to N-th Legendre polynomial, and the alternative discontinuous interpretation where the last Legendre coordinate c^N instead amounts to a discontinuous end point that is appended to the remaining $(N-1)$-th degree polynomial.

A.1.1 Legendre Coordinates

Legendre Coordinates, Continuous Consider the standard interpretation of Legendre coordinates from (3.46). The resulting function and its (continuous) end point give rise to the following element in M_2

$$\left[\begin{smallmatrix}\phi^{[N]}(\cdot)\\ \phi^{[N]}(0)\end{smallmatrix}\right] = \sum_{k=0}^{N} c^k \left[\begin{smallmatrix}p_k(\vartheta(\cdot))\\ p_k(\vartheta(0))\end{smallmatrix}\right] =: \sum_{k=0}^{N} c^k \mathrm{g}_{c,k}. \tag{A.2}$$

Consequently, since $p_k(\vartheta(0)) = p_k(1) = 1$ (see Figure 3.5), the basis functions of the associated polynomial subspace of M_2 read

$$\mathrm{g}_{c,0} = \left[\begin{smallmatrix}p_0(\vartheta(\cdot))\\ 1\end{smallmatrix}\right], \quad \ldots, \quad \mathrm{g}_{c,N-1} = \left[\begin{smallmatrix}p_{N-1}(\vartheta(\cdot))\\ 1\end{smallmatrix}\right], \quad \mathrm{g}_{c,N} = \left[\begin{smallmatrix}p_N(\vartheta(\cdot))\\ 1\end{smallmatrix}\right]. \tag{A.3}$$

The metric coefficients are nothing more than the inner products of (A.3).

Definition A.1.1 (Metric coefficients). *Let $\{\mathrm{g}_k\}_{k\in\{0,\ldots,N\}}$ be the basis of a subspace in M_2. The metric coefficients G_{jk}, with $j,k \in \{0,\ldots,N\}$, are*

$$G_{jk} = \langle \mathrm{g}_j, \mathrm{g}_k \rangle_{M_2}. \tag{A.4}$$

If $G_{jk} = \delta_{jk}$, the basis is said to be orthonormal.

For the above defined basis elements $\{\mathsf{g}_{c,k}\}_k$,

$$G_{c,jk} = \langle \mathsf{g}_{c,j}, \mathsf{g}_{c,k} \rangle_{M_2} = \underbrace{\langle p_j(\vartheta(\cdot)), p_k(\vartheta(\cdot)) \rangle_{L_2([-h,0],\mathbb{R})}}_{\int_{-h}^{0} p_j(\vartheta(\theta)) p_k(\vartheta(\theta)) \,\mathrm{d}\theta} + \underbrace{p_j(\vartheta(0))}_{1} \underbrace{p_k(\vartheta(0))}_{1}$$
$$= \begin{cases} \frac{h}{2} \frac{2}{2k+1} + 1 & \text{if } k = j, \\ 1 & \text{otherwise} \end{cases} \tag{A.5}$$

gives a dense matrix of metric coefficients

$$G_c = \begin{bmatrix} (\frac{h}{2}2 + 1) & 1 & \cdots & & 1 \\ 1 & (\frac{h}{2}\frac{2}{3} + 1) & 1 & & \\ \vdots & & \ddots & & \vdots \\ & & 1 & (\frac{h}{2}\frac{2}{2(N-1)+1} + 1) & 1 \\ 1 & & \cdots & 1 & (\frac{h}{2}\frac{2}{2N+1} + 1) \end{bmatrix}. \tag{A.6}$$

Not having an orthonormal basis is the raison d'être of the present appendix.

Legendre Coordinates, Discontinuous In Section 3.5.2.2, the interpretation of the Legendre coordinates in terms of a polynomial of degree at most $N-1$ and a discontinuous end point occurs. This interpretation amounts to

$$\begin{bmatrix} \bar{\phi}_{\mathrm{d}}^{[N]}(\cdot) \\ \phi_{\mathrm{d}}^{[N]}(0) \end{bmatrix} = \sum_{k=0}^{N-1} c^k \begin{bmatrix} p_k(\vartheta(\cdot)) \\ 1 \end{bmatrix} + c^N \begin{bmatrix} 0 \\ 1 \end{bmatrix} =: \sum_{k=0}^{N} c^k \mathsf{h}_{c,k} \tag{A.7}$$

(where $\phi_{\mathrm{d}}^{[N]}(0) = \sum_{k=0}^{N} c^k$ still applies), i.e., the basis functions

$$\mathsf{h}_{c,0} = \begin{bmatrix} p_0(\vartheta(\cdot)) \\ 1 \end{bmatrix}, \quad \ldots, \quad \mathsf{h}_{c,N-1} = \begin{bmatrix} p_{N-1}(\vartheta(\cdot)) \\ 1 \end{bmatrix}, \quad \mathsf{h}_{c,N} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{A.8}$$

span the corresponding subspace of M_2.

Since, compared to $\{\mathsf{g}_k\}_k$, only the last basis element has changed, only the last row and last column of the metric coefficients G_c can be affected. In fact, only the lower diagonal element $H_{c,NN} = \langle \mathsf{h}_N, \mathsf{h}_N \rangle_{M_2} = 1$ is altered, giving rise to

$$H_{c,jk} = \langle \mathsf{h}_{c,j}, \mathsf{h}_{c,k} \rangle_{M_2}, \tag{A.9}$$

$$H_c = \begin{bmatrix} (\frac{h}{2}2+1) & 1 & & \cdots & & 1 \\ 1 & (\frac{h}{2}\frac{2}{3}+1) & 1 & & & \vdots \\ \vdots & & \ddots & & & \\ & & & 1 & (\frac{h}{2}\frac{2}{2(N-1)+1}+1) & 1 \\ 1 & & \cdots & & 1 & 1 \end{bmatrix}. \tag{A.10}$$

A.1.2 Change of Basis, Mixed Coordinates

Change of Basis Clearly, a change of basis from Legendre coordinates c to mixed coordinates χ discussed in Section 3.6.2 must leave the polynomial unaltered

$$\begin{bmatrix} \phi^{[N]}(\cdot) \\ \phi^{[N]}(0) \end{bmatrix} = \sum_{k=0}^{N} c^k \mathsf{g}_{c,k} = \sum_{k=0}^{N} \chi^k \mathsf{g}_{\chi,k}. \tag{A.11}$$

Using the transformation law of the coordinates $c = T_{c\chi}\chi$ (relying on the inverse $T_{c\chi} = T_{\chi c}^{-1}$ of the transformation matrix (3.64) from $\chi = T_{\chi c}c$) in the above expression

$$\sum_{k=0}^{N} c^k \mathsf{g}_{c,k} = \begin{bmatrix} \mathsf{g}_{c,0} & \cdots & \mathsf{g}_{c,N} \end{bmatrix} \underbrace{\begin{bmatrix} c^0 \\ \vdots \\ c^N \end{bmatrix}}_{T_{c\chi}\chi} = \begin{bmatrix} \mathsf{g}_{\chi,0} & \cdots & \mathsf{g}_{\chi,N} \end{bmatrix} \underbrace{\begin{bmatrix} \chi^0 \\ \vdots \\ \chi^N \end{bmatrix}}_{\chi} \tag{A.12}$$

shows that the basis elements are related by this inverse transformation matrix via

$$\begin{bmatrix} \mathsf{g}_{\chi,0} & \cdots & \mathsf{g}_{\chi,N} \end{bmatrix} = \begin{bmatrix} \mathsf{g}_{c,0} & \cdots & \mathsf{g}_{c,N} \end{bmatrix} T_{c\chi} \tag{A.13}$$

$$= \begin{bmatrix} \mathsf{g}_{c,0} & \cdots & \mathsf{g}_{c,N} \end{bmatrix} \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ -1 & \cdots & -1 & 1 \end{bmatrix}. \tag{A.14}$$

Mixed Coordinates, Continuous As a result of (A.14), the basis that is associated with the mixed coordinates χ in the continuous interpretation is given by

$$\mathsf{g}_{\chi,0} = \mathsf{g}_{c,0} - \mathsf{g}_{c,N} = \begin{bmatrix} p_0(\vartheta(\cdot)) - p_N(\vartheta(\cdot)) \\ 0 \end{bmatrix}, \quad \ldots, \tag{A.15}$$

$$\mathsf{g}_{\chi,N-1} = \mathsf{g}_{c,N-1} - \mathsf{g}_{c,N} = \begin{bmatrix} p_{N-1}(\vartheta(\cdot)) - p_N(\vartheta(\cdot)) \\ 0 \end{bmatrix}, \quad \mathsf{g}_{\chi,N} = \mathsf{g}_{c,N} = \begin{bmatrix} p_N(\vartheta(\cdot)) \\ 1 \end{bmatrix}.$$

With

$$\langle p_N(\vartheta(\cdot)), p_N(\vartheta(\cdot)) \rangle_{L_2([-h,0],\mathbb{R})} = \frac{h}{2} \frac{2}{2N+1} =: \nu, \tag{A.16}$$

the metric coefficients are

$$G_{\chi,jk} = \langle \mathsf{g}_{\chi,j}, \mathsf{g}_{\chi,k} \rangle_{M_2}$$

$$G_\chi = \begin{bmatrix} (\frac{h}{2}2 + \nu) & \nu & \cdots & \nu & -\nu \\ \nu & (\frac{h}{2}\frac{2}{3} + \nu) & \nu & \vdots & \vdots \\ \vdots & & \ddots & \nu & \\ \nu & \cdots & \nu & (\frac{h}{2}\frac{2}{2(N-1)+1} + \nu) & -\nu \\ -\nu & \cdots & & -\nu & (\nu+1) \end{bmatrix}. \tag{A.17}$$

The same result can also be derived via

$$G_\chi = T_{c\chi}^\top G_c T_{c\chi}, \tag{A.18}$$

which will become clear from (A.26) below.

Mixed Coordinates, Discontinuous Analogously to (A.14), the basis $\{\mathsf{h}_{\chi,k}\}_k$ associated with the mixed coordinates χ in the discontinuous interpretation can be deduced from $\{\mathsf{h}_{c,k}\}_k$ in (A.8) and $T_{c\chi}$ as

$$\begin{aligned} \mathsf{h}_{\chi,0} = \mathsf{h}_{c,0} - \mathsf{h}_{c,N} = \left[\begin{smallmatrix} p_0(\vartheta(\cdot)) \\ 0 \end{smallmatrix}\right], \quad \ldots, \\ \mathsf{h}_{\chi,N-1} = \mathsf{h}_{c,N-1} - \mathsf{h}_{c,N} = \left[\begin{smallmatrix} p_{N-1}(\vartheta(\cdot)) \\ 0 \end{smallmatrix}\right], \qquad \mathsf{h}_{\chi,N} = \mathsf{h}_{c,N} = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]. \end{aligned} \tag{A.19}$$

The metric coefficients

$$H_{\chi,jk} = \langle \mathsf{h}_{\chi,j}, \mathsf{h}_{\chi,k} \rangle_{M_2}, \tag{A.20}$$

$$H_\chi = \begin{bmatrix} \frac{h}{2}2 & 0 & \cdots & & & 0 \\ 0 & \frac{h}{2}\frac{2}{3} & 0 & & & \\ \vdots & & \ddots & & & \vdots \\ & & & 0 & \frac{h}{2}\frac{2}{2(N-1)+1} & 0 \\ 0 & & \cdots & & 0 & 1 \end{bmatrix} \tag{A.21}$$

$$= T_{c\chi}^\top H_c T_{c\chi} \tag{A.22}$$

show that, in the discontinuous interpretation, the basis that is associated with the mixed coordinates χ is an orthogonal but still no orthonormal basis (which, however, could simply be achieved by scaling each basis function).

A.2 Operators on the Considered Polynomial Subspaces of M_2

Henceforth, instead of denoting Legendre coordinates by $c^k \in \mathbb{R}$, $k \in \{0, \dots, N\}$, and mixed coordinates by $\chi^k \in \mathbb{R}$, general variable names are introduced, where instead a subscript indicates the chosen basis. For instance, an element $\mathsf{b} \in \operatorname{span}(\{\mathsf{g}_{c,k}\}_k) = \operatorname{span}(\{\mathsf{g}_{\chi,k}\}_k) \subset M_2$ (i.e., a polynomial of degree N and its end point) that is represented in the basis from (A.3), respectively in the basis from (A.15), is denoted by

$$\mathsf{b} = \sum_{k=0}^{N} b_c^k \mathsf{g}_{c,k} = \sum_{k=0}^{N} b_\chi^k \mathsf{g}_{\chi,k}. \tag{A.23}$$

The involved Legendre coordinates $b_c^k \in \mathbb{R}$ and mixed coordinates $b_\chi^k \in \mathbb{R}$ are related by

$$\begin{bmatrix} b_c^0 \\ \vdots \\ b_c^N \end{bmatrix} = T_{c\chi} \begin{bmatrix} b_\chi^0 \\ \vdots \\ b_\chi^N \end{bmatrix} \tag{A.24}$$

according to Section 3.6. Moreover, if the basis can arbitrarily (but of course consistently) be chosen, this subscript is omitted, i.e.,

$$\mathsf{b} = \sum_{k=0}^{N} b^k \mathsf{g}_k.$$

Henceforth, $a^j, b^k \in \mathbb{R}$ and $j, k \in \{0, \dots, N\}$, unless otherwise stated.

To make the appendix self-contained, the present section involves the derivations of some elementary relations in Appendix A.2.1.

A.2.1 Preliminaries

Inner Product in a Non-Orthonormal Basis The metric coefficients are important when computing inner products. Consider two functions $\mathsf{a} = \sum_{k=0}^{N} a^k \mathsf{g}_k$ and $\mathsf{b} = \sum_{k=0}^{N} b^k \mathsf{g}_k$ in an arbitrary basis $\{\mathsf{g}_k\}_{k\in\{0,\dots,N\}}$ of the considered[1] $(N+1)$-dimensional subspace of M_2. Then

$$\begin{aligned}\langle \mathsf{a}, \mathsf{b}\rangle_{M_2} &= \left\langle \sum_{j=0}^{N} a^j \mathsf{g}_j, \sum_{k=0}^{N} b^k \mathsf{g}_k \right\rangle_{M_2} = \sum_{j=0}^{N}\sum_{k=0}^{N} a^j \langle \mathsf{g}_j, \mathsf{g}_k\rangle_{M_2} b^k = \sum_{j,k} a^j G_{jk} b^k \\ &= a^\top G\, b, \qquad a = \begin{bmatrix} a^0 \\ \vdots \\ a^N \end{bmatrix}, \quad b = \begin{bmatrix} b^0 \\ \vdots \\ b^N \end{bmatrix}. \end{aligned} \tag{A.25}$$

Moreover, the congruence transformation law for metric coefficients stated in (A.18) is immediately explained by recognizing that the scalar result $\langle \mathsf{a}, \mathsf{b}\rangle_{M_2}$ must be invariant with respect to the chosen basis. Thus, considering a and b in Legendre coordinates or in mixed coordinates related by (A.24) may not make a difference in

$$\langle \mathsf{a}, \mathsf{b}\rangle_{M_2} = a_c^\top G_c b_c = (T_{c\chi} a_\chi)^\top G_c (T_{c\chi} b_\chi) = a_\chi^\top \underbrace{T_{c\chi}^\top G_c T_{c\chi}}_{=:G_\chi} b_\chi. \tag{A.26}$$

The Dual Basis Functions The dual basis functions are functions that are orthogonal to the original basis functions $\{\mathsf{g}_k\}_k$. In a tensor calculus setting, it is common practice to distinguish the original basis and the dual basis only by the position of the indices, notating the latter with upper indices. In the present appendix an underline $\underline{\mathsf{g}}^k$ intends to to emphasize the difference more clearly.

[1] For notational convenience, the subspace is denoted by $\mathrm{span}(\{\mathsf{g}_k\}_k)$, but the results in this section hold equally well for $\mathrm{span}(\{\mathsf{h}_k\}_k)$ or any other $(N+1)$-dimensional subspace of M_2.

Definition A.2.1 (Dual basis functions). *Let $\{\mathsf{g}_k\}_k$ be a basis of the $(N+1)$-dimensional space $V = \operatorname{span}(\{\mathsf{g}_k\}_k) \subset M_2$. The dual basis functions $\underline{\mathsf{g}}^k \in V$ are uniquely defined by requiring that $\forall j,k \in \{0,\dots,N\}$:*

$$\langle \mathsf{g}_j, \underline{\mathsf{g}}^k \rangle_{M_2} = \delta_j^k, \qquad \delta_j^k \stackrel{\text{def}}{=} \delta_{jk} = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases} \tag{A.27}$$

With a slight abuse of nomenclature, the term dual basis will henceforth refer to $\{\underline{\mathsf{g}}^k\}_k$, although the latter is a basis of the original space V rather than a basis of the dual space V^*, which, however, is isomorphic to V (see Remark A.2.2).

As a consequence, any function $\mathsf{b} \in V$ can equally well be represented in the original basis or in the dual basis, for which the coordinates in turn are denoted by lower indices

$$\mathsf{b} = \sum_{k=0}^{N} b^k \mathsf{g}_k = \sum_{k=0}^{N} \underline{b}_k \underline{\mathsf{g}}^k. \tag{A.28}$$

The transformation matrix, which transforms $b \in \mathbb{R}^{N+1}$ to $\underline{b} \in \mathbb{R}^{N+1}$ will be discussed in (A.31) below.

Remark A.2.2 (Alternative nomenclature referring to the dual space). *The dual space V^* of $V = \operatorname{span}(\{\mathsf{g}_k\}_k) \subset M_2$ is spanned by the associated functionals*

$$\langle \underline{\mathsf{g}}^k, \cdot \rangle_{M_2}. \tag{A.29}$$

In the literature, the term dual basis is frequently taken literally and refers to these functionals. Moreover, this reading motivates the use of row vectors for the coordinates in the dual basis, and these row vectors are also called covectors or one-forms as a synonym for linear functionals. Since the present section does not follow this interpretation, coordinates in the dual basis are—such as coordinates in any other basis of V—collected in column vectors. In the mere index notation,

there is no need to distinguish between row and column vectors anyway, see the calculations below.

Inner Product without Metric Coefficients The sense behind the dual basis is that, if either a or b is represented in the dual basis, then no metric coefficients occur in (A.25). Instead, using that $\sum_j \delta_j^k x^j = x^k$,

$$\langle \mathsf{a}, \mathsf{b} \rangle_{M_2} = \left\langle \sum_{j=0}^{N} a^j \mathsf{g}_j, \sum_{k=0}^{N} \underline{b}_k \underline{\mathsf{g}}^k \right\rangle_{M_2} = \sum_{j,k} a^j \underbrace{\langle \mathsf{g}_j, \underline{\mathsf{g}}^k \rangle_{M_2}}_{\delta_j^k} \underline{b}_k = \sum_k a^k \underline{b}_k$$

$$= a^\top \underline{b}, \qquad a = \begin{bmatrix} a^0 \\ \vdots \\ a^N \end{bmatrix}, \quad \underline{b} = \begin{bmatrix} \underline{b}_0 \\ \vdots \\ \underline{b}_N \end{bmatrix}. \tag{A.30}$$

Associated Transformation of the Coordinates Comparing (A.25) and (A.30) already shows that both coordinate representations are related by

$$\underline{b} = Gb. \tag{A.31}$$

The same conclusion can also be drawn from noting that the i-th coordinate $\underline{b}_i$ in the dual basis results from taking an inner product with the i-th original basis element g_i

$$\langle \mathsf{g}_i, \mathsf{b} \rangle_{M_2} = \langle \mathsf{g}_i, \sum_{k=0}^{N} \underline{b}_k \underline{\mathsf{g}}^k \rangle_{M_2} = \sum_{k=0}^{N} \underbrace{\langle \mathsf{g}_i, \underline{\mathsf{g}}^k \rangle_{M_2}}_{\delta_i^k} \underline{b}_k = \underline{b}_i, \tag{A.32}$$

and, at the same time,

$$\langle \mathsf{g}_i, \mathsf{b} \rangle_{M_2} = \langle \mathsf{g}_i, \sum_{k=0}^{N} b^k \mathsf{g}_k \rangle_{M_2} = \sum_{k=0}^{N} \langle \mathsf{g}_i, \mathsf{g}_k \rangle_{M_2} b^k = \sum_{k=0}^{N} G_{ik} b^k. \tag{A.33}$$

Comparing both also shows the known relation $\underline{b}_i = \sum_{k=0}^{N} G_{ik} b^k$.

Computing the Dual Basis Functions Analogously to the change of basis discussed in (A.13),

$$\sum_{k=0}^{N} b^k \mathsf{g}_k = \begin{bmatrix} \mathsf{g}_0 & \cdots & \mathsf{g}_N \end{bmatrix} \underbrace{\begin{bmatrix} b^0 \\ \vdots \\ b^N \end{bmatrix}}_{b} = \sum_{k=0}^{N} \underline{b}_k \underline{\mathsf{g}}^k = \begin{bmatrix} \underline{\mathsf{g}}^0 & \cdots & \underline{\mathsf{g}}^N \end{bmatrix} \underbrace{\begin{bmatrix} \underline{b}_i \\ \vdots \\ \underline{b}_N \end{bmatrix}}_{\underline{b}=Gb}$$

shows that the dual basis functions $\underline{\mathsf{g}}^k$ can be computed from the original basis functions g_k via

$$\begin{bmatrix} \underline{\mathsf{g}}^0 & \cdots & \underline{\mathsf{g}}^N \end{bmatrix} = \begin{bmatrix} \mathsf{g}_0 & \cdots & \mathsf{g}_N \end{bmatrix} G^{-1}. \tag{A.34}$$

Clearly, all these results hold for arbitrary basis functions, including the ones that are associated with the discontinuous interpretation, where as well

$$\begin{bmatrix} \underline{\mathsf{h}}^0 & \cdots & \underline{\mathsf{h}}^N \end{bmatrix} = \begin{bmatrix} \mathsf{h}_0 & \cdots & \mathsf{h}_N \end{bmatrix} H^{-1} \tag{A.35}$$

gives the dual basis such that $\langle \mathsf{h}_j, \underline{\mathsf{h}}^k \rangle_{M_2} = \delta_j^k$.

Change of Coordinates in the Dual Basis Similarly to $b_c = T_{c\chi} b_\chi$ from (A.24), if $\underline{b}_\chi$ are coordinates in $\{\underline{\mathsf{g}}_\chi^k\}_k$, then

$$\begin{aligned} \underline{b}_c &\overset{(A.31)}{=} G_c b_c = G_c T_{c\chi} b_\chi \overset{(A.31)}{=} G_c T_{c\chi} G_\chi^{-1} \underline{b}_\chi \overset{(A.26)}{=} (T_{\chi c}^\top G_\chi T_{\chi c}) T_{c\chi} G_\chi^{-1} \underline{b}_\chi \\ &= T_{\chi c}^\top \underline{b}_\chi \end{aligned} \tag{A.36}$$

are the coordinates in $\{\underline{\mathsf{g}}_c^k\}_k$. The same transformation applies if $\underline{b}_\chi$ are coordinates in $\{\underline{h}_\chi^k\}_k$ (where G_χ and G_c above are replaced by H_χ and H_c).

Using the Dual Basis to Obtain the Coordinates in the Original Basis Analogously to (A.32),

$$\langle \underline{\mathsf{g}}^i, \mathsf{b} \rangle_{M_2} = \langle \underline{\mathsf{g}}^i, \sum_{k=0}^{N} b^k \mathsf{g}_k \rangle_{M_2} = \sum_{k=0}^{N} \underbrace{\langle \underline{\mathsf{g}}^i, \mathsf{g}_k \rangle_{M_2}}_{\delta^i_k} b^k = b^i. \tag{A.37}$$

Metric Coefficients The matrix of metric coefficients of the dual basis are

$$\underline{G} = G^{-1}, \tag{A.38}$$

which becomes obvious from comparing

$$\langle \mathsf{a}, \mathsf{b} \rangle_{M_2} = \left\langle \sum_{j=0}^{N} \underline{a}_j \underline{\mathsf{g}}^j, \sum_{k=0}^{N} \underline{b}_k \underline{\mathsf{g}}^k \right\rangle_{M_2} = \sum_{j=0}^{N} \sum_{k=0}^{N} \underline{a}_j \underbrace{\langle \underline{\mathsf{g}}^j, \underline{\mathsf{g}}^k \rangle_{M_2}}_{\underline{G}^{jk}} \underline{b}_k = \underline{a}^\top \underline{G}\, \underline{b}$$

with

$$\langle \mathsf{a}, \mathsf{b} \rangle_{M_2} = a^\top G\, b \overset{(A.31)}{=} (G^{-1}\underline{a})^\top G (G^{-1}\underline{b}) = \underline{a}^\top G^{-1} \underline{b}. \tag{A.39}$$

Transformation Laws of Basis Functions Since $G = G^\top$, (A.34) combined with (A.38) gives

$$\mathsf{g}_k = \sum_{j=0}^{N} G_{jk} \underline{\mathsf{g}}^j = \sum_{j=0}^{N} G_{kj} \underline{\mathsf{g}}^j, \qquad \underline{\mathsf{g}}^k = \sum_{j=0}^{N} \underline{G}^{jk} \mathsf{g}_j = \sum_{j=0}^{N} \underline{G}^{kj} \mathsf{g}_j. \tag{A.40}$$

A.2.2 Representation of Operators on the Polynomial Subspaces of M_2

In view of the operators $\mathscr{A}^{[N]}$ and $\mathscr{P}^{[N]}$ that occur in the present thesis, the following section takes a closer look on two related types of operators, which,

however, do not yet incorporate a projection. Instead, they are only defined on the polynomial subspaces that are spanned by the basis functions from Appendix A.1.

Linear Operator $\mathscr{L}$ from $\mathbf{span}(\{\mathsf{g}_j\}_j)$ to $\mathbf{span}(\{\mathsf{h}_j\}_j)$ Consider a linear operator $\mathscr{L}$ that maps a function $\mathsf{x} \in \mathrm{span}(\{\mathsf{g}_k\}_k)$ to a function $\mathsf{y} \in \mathrm{span}(\{\mathsf{h}_k\}_k)$,

$$\mathsf{x} = \sum_k x^k \mathsf{g}_k \quad \mapsto \quad \mathscr{L}\mathsf{x} := \mathsf{y} = \sum_j y^j \mathsf{h}_j \quad \text{with } y^j = L^j{}_k \, x^k. \tag{A.41}$$

The matrix that maps the coordinate representation of x to the coordinate representation of y in

$$\begin{bmatrix} y^0 \\ \vdots \\ y^N \end{bmatrix} = \begin{bmatrix} L^0{}_0 & \cdots & L^0{}_N \\ \vdots & & \vdots \\ L^N{}_0 & \cdots & L^N{}_N \end{bmatrix} \begin{bmatrix} x^0 \\ \vdots \\ x^N \end{bmatrix} \tag{A.42}$$

is a coordinate representation of that operator. The operator is explicitly given by

$$\mathscr{L} = \sum_{j=0}^{N} \sum_{k=0}^{N} L^j{}_k \; \mathsf{h}_j \otimes \underline{\mathsf{g}}^k, \tag{A.43}$$

relying on the following definition.

Definition A.2.3 (Dyadic product in M_2). *The dyadic product (or tensor product) $\otimes$ is defined by*[2]

$$(\mathsf{a} \otimes \mathsf{b}) \cdot \stackrel{\text{def}}{=} \mathsf{a} \langle \mathsf{b}, \cdot \rangle_{M_2}. \tag{A.44}$$

[2] The resulting operator can be applied to elements of the overall space M_2 rather than only the polynomial subspace considered in the present section (see Appendix A.3.1).

Indeed, (A.43) applied to x

$$\mathscr{L}\mathsf{x} = \Big(\sum_{j,k} L^j{}_k \mathsf{h}_j \otimes \underline{\mathsf{g}}^k\Big)\Big(\sum_l x^l \mathsf{g}_l\Big) \overset{(A.44)}{=} \sum_{j,k,l} L^j{}_k \mathsf{h}_j \underbrace{\langle \underline{\mathsf{g}}^k, \mathsf{g}_l\rangle_{M_2}}_{\delta^k_l} x^l \tag{A.45}$$

$$= \sum_{j,k} L^j{}_k x^k \mathsf{h}_j = \sum_j y^j \mathsf{h}_j \tag{A.46}$$

gives y defined in (A.41). The latter can also directly be recognized from the definition (A.44)

$$\mathscr{L}\cdot = \sum_{j,k} L^j{}_k \langle \underline{\mathsf{g}}^k, \cdot\rangle_{M_2}\, \mathsf{h}_j \tag{A.47}$$

by noticing that, according to (A.32), $\langle \underline{\mathsf{g}}^k, \mathsf{x}\rangle_{M_2} = x^k$ extracts the k-th component in the original basis.

Alternative Representations Since $\mathrm{span}(\{\mathsf{h}_j\}_j) = \mathrm{span}(\{\underline{\mathsf{h}}^j\}_j)$ and also $\mathrm{span}(\{\mathsf{g}_k\}_k) = \mathrm{span}(\{\underline{\mathsf{g}}^k\}_k)$, the same operator could as well be rewritten as

$$\mathscr{L} = \sum_{j,k} L^j{}_k\, \mathsf{h}_j \otimes \underline{\mathsf{g}}^k = \sum_{j,k} L_{jk}\, \underline{\mathsf{h}}^j \otimes \underline{\mathsf{g}}^k = \sum_{j,k} L^{jk}\, \mathsf{h}_j \otimes \mathsf{g}_k = \sum_{j,k} L_j{}^k\, \underline{\mathsf{h}}^j \otimes \mathsf{g}_k, \tag{A.48}$$

where in general $L^j{}_k \neq L_{jk} \neq L^{jk} \neq L_j{}^k$ holds for nonorthonormal bases. The conversion laws between the various matrices can be concluded from (A.40).

Operator $\mathscr{W}$ in a Bilinear Form on $\mathbf{span}(\{\mathsf{h}_j\}_j)$ Let the bilinear form

$$(\mathsf{z}, \mathsf{y}) = \left(\sum_k z^k \mathsf{h}_k, \sum_k y^k \mathsf{h}_k\right) \quad \mapsto \quad \langle \mathsf{z}, \mathscr{W}\mathsf{y}\rangle_{M_2} := \sum_{j,k} z^j W_{jk} y^k \tag{A.49}$$

map two functions that are represented by the coordinate vectors z and y to the scalar value

$$\sum_{j,k} z^j W_{jk} y^k = \begin{bmatrix} z^0 \\ \vdots \\ z^N \end{bmatrix}^\top \begin{bmatrix} W_{00} & \cdots & W_{0N} \\ \vdots & & \vdots \\ W_{N0} & \cdots & W_{NN} \end{bmatrix} \begin{bmatrix} y^0 \\ \vdots \\ y^N \end{bmatrix}. \tag{A.50}$$

The involved operator is

$$\mathscr{W} = \sum_{j=0}^{N} \sum_{k=0}^{N} W_{jk}\, \underline{\mathsf{h}}^j \otimes \underline{\mathsf{h}}^k \tag{A.51}$$

since then $\underline{w}_j := \sum_k W_{jk} y^k$ are the coordinates of

$$\begin{aligned} \mathsf{w} := \mathscr{W}\mathsf{y} &= \Big(\sum_{j,k} W_{jk} \underline{\mathsf{h}}^j \otimes \underline{\mathsf{h}}^k \Big) \Big(\sum_l y^l \mathsf{h}_l \Big) \overset{(A.44)}{=} \sum_{j,k,l} W_{jk} \underline{\mathsf{h}}^j \underbrace{\langle \underline{\mathsf{h}}^k, \mathsf{h}_l \rangle_{M_2}}_{\delta_l^k} y^l \\ &= \sum_{j,k} W_{jk} y^k \underline{\mathsf{h}}^j = \sum_j \underline{w}_j \underline{\mathsf{h}}^j \end{aligned} \tag{A.52}$$

in the dual basis (in contrast to the result from (A.46)). As a consequence, the computation of $\langle \mathsf{z}, \mathsf{w} \rangle_{M_2}$ in (A.49) does not introduce additional metric coefficients but, cf. (A.30), simply becomes

$$\langle \mathsf{z}, \mathscr{W}\mathsf{y} \rangle_{M_2} = \Big\langle \sum_j z^j \mathsf{h}_j, \sum_k \underline{w}_k \underline{\mathsf{h}}^k \Big\rangle_{M_2} = \sum_{j,k} z^j \underbrace{\langle \mathsf{h}_j, \underline{\mathsf{h}}^k \rangle_{M_2}}_{\delta_j^k} \underline{w}_k = \sum_j z^j \underline{w}_j$$

as desired in (A.49).

Relation Between the Operator and its Coordinates Conversely to the construction of the operator from the coordinates in (A.51), the coordinates are recovered from the operator (no matter in which form it is given) via

$$W_{jk} = \langle \mathsf{h}_j, \mathscr{W} \mathsf{h}_k \rangle_{M_2} \tag{A.53}$$

since

$$\langle \mathsf{h}_j, \mathscr{W}\mathsf{h}_k\rangle_{M_2} = \Big\langle \mathsf{h}_j, \Big(\sum_{l,m} W_{lm}\, \underline{\mathsf{h}}^l \otimes \underline{\mathsf{h}}^m\Big)\mathsf{h}_k\Big\rangle_{M_2} \tag{A.54}$$

$$= \sum_{l,m} \underbrace{\langle \mathsf{h}_j, \underline{\mathsf{h}}^l\rangle_{M_2}}_{\delta_j^l} W_{lm} \underbrace{\langle \underline{\mathsf{h}}^m, \mathsf{h}_k\rangle_{M_2}}_{\delta_k^m} = W_{jk}. \tag{A.55}$$

Similarly, concerning (A.43),

$$L^j{}_k = \langle \underline{\mathsf{h}}^j, \mathscr{L}\mathsf{g}_k\rangle_{M_2}. \tag{A.56}$$

A.2.3 Norm and Operator Norm

M_2-Norm Consider a function $\mathsf{x} = \sum_j x^j \mathsf{g}_j$. Because of (A.25), the norm $\|\mathsf{x}\|_{M_2} = \sqrt{\langle \mathsf{x}, \mathsf{x}\rangle_{M_2}}$ is derived from the coordinates $x \in \mathbb{R}^{N+1}$ via

$$\|\mathsf{x}\|_{M_2} = \sqrt{x^\top G x} = \|G^{1/2}x\|_2, \tag{A.57}$$

where the matrix square root can equally well be replaced by an arbitrary decomposition $G = C^\top C$.

Operatornorm of $\mathscr{L}$ Let $\mathscr{L}\colon \operatorname{span}(\{\mathsf{g}_k\}_k) \to \operatorname{span}(\{\mathsf{h}_k\}_k)$ be described by $\mathscr{L} = \sum_{j,k} L^j{}_k\, \mathsf{h}_j \otimes \underline{\mathsf{g}}^k$. The operator norm is computed from the matrix L with entries $L^j{}_k$ via

$$\|\mathscr{L}\| = \|H^{1/2} L G^{-1/2}\|_2 \tag{A.58}$$

since

$$\|\mathscr{L}\|^2 = \sup_{\mathsf{x}\in\operatorname{span}(\{\mathsf{g}_k\}_k)\setminus\{0_{M_2}\}} \frac{\|\mathscr{L}\mathsf{x}\|_{M_2}^2}{\|\mathsf{x}\|_{M_2}^2} \overset{(A.57)}{=} \sup_{x\in\mathbb{R}^{(N+1)}\setminus\{0_{N+1}\}} \frac{x^\top L^\top H L x}{x^\top G x}$$
$$= \sup_{u\in\mathbb{R}^{(N+1)}\setminus\{0_{N+1}\}} \frac{u^\top G^{-1/2} L^\top H L G^{-1/2} u}{u^\top u}. \tag{A.59}$$

Operatornorm of $\mathscr{W}$ In the representation $\mathscr{W} = \sum_{j,k} W_{jk}\ \underline{\mathsf{h}}^j \otimes \underline{\mathsf{h}}^k$, the matrix W with entries W_{jk} maps to coordinates in the dual basis $\{\underline{\mathsf{h}}^j\}$. Therefore, the metric coefficients of $\{\underline{\mathsf{h}}^j\}$, which due to (A.38) are given by H^{-1}, occur in

$$\|\mathscr{W}\| = \sqrt{\sup_{x\in\mathbb{R}^{(N+1)}\setminus\{0_{N+1}\}} \frac{x^\top W^\top H^{-1} W x}{x^\top H x}} = \|H^{-1/2} W H^{-1/2}\|_2. \tag{A.60}$$

A.2.4 Adjoint Operator

Adjoint of $\mathscr{L}$ The adjoint of $\mathscr{L}\colon \operatorname{span}(\{\mathsf{g}_k\}_k) \to \operatorname{span}(\{\mathsf{h}_k\}_k)$,

$$\mathscr{L} = \sum_{j,k} L^j{}_k\ \mathsf{h}_j \otimes \underline{\mathsf{g}}^k, \quad L := \begin{bmatrix} L^0{}_0 & \cdots & L^0{}_N \\ \vdots & & \\ L^N{}_0 & \cdots & L^N{}_N \end{bmatrix}, \tag{A.61}$$

becomes $\mathscr{L}^*\colon \operatorname{span}(\{\mathsf{h}_k\}_k) \to \operatorname{span}(\{\mathsf{g}_k\}_k)$,

$$\mathscr{L}^* = \sum_{j,k} S^j{}_k \mathsf{g}_j \otimes \underline{\mathsf{h}}^k, \quad S = \begin{bmatrix} S^0{}_0 & \cdots & S^0{}_N \\ \vdots & & \\ S^N{}_0 & \cdots & S^N{}_N \end{bmatrix} := G^{-1} L^\top H, \tag{A.62}$$

and thus (in this representation) the coordinate matrix S of the adjoint is not the matrix transpose of L unless both $\{\mathsf{g}_k\}_k$ and $\{\mathsf{h}_k\}_k$ are orthonormal bases.

The reasoning is as follows: The adjoint $\mathscr{L}^*$ must satisfy

$$\langle \mathsf{z}, \mathscr{L}\mathsf{x}\rangle_{M_2} = \langle \mathscr{L}^*\mathsf{z}, \mathsf{x}\rangle_{M_2}. \tag{A.63}$$

The left-hand side of (A.63) becomes

$$\begin{aligned}\langle \mathsf{z}, \mathscr{L}\mathsf{x}\rangle_{M_2} &= \Big\langle \sum_j z^j \mathsf{h}_j, \Big(\sum_{k,l} L^k_l \mathsf{h}_k \otimes \underline{\mathsf{g}}^l\Big) \sum_m x^m \mathsf{g}_m \Big\rangle_{M_2} = \sum_{j,k,m} z^j H_{jk} L^k{}_m x^m \\ &= z^\top H L x. \end{aligned} \tag{A.64}$$

The right-hand side of (A.63) becomes

$$\begin{aligned}\langle \mathscr{L}^*\mathsf{z}, \mathsf{x}\rangle_{M_2} &= \Big\langle \Big(\sum_{j,k} S^j{}_k \mathsf{g}_j \otimes \underline{\mathsf{h}}^k\Big) \sum_l z^l \mathsf{h}_l, \sum_m x^m \mathsf{g}_m \Big\rangle_{M_2} = \sum_{j,k,m} z^k S^j{}_k G_{jm} x^m \\ &= z^\top S^\top G x. \end{aligned} \tag{A.65}$$

Therefore, (A.63) requires that (by $G = G^\top$ and $H = H^\top$)

$$S := G^{-1} L^\top H. \tag{A.66}$$

If the same operator $\mathscr{L}^*$ is instead represented as

$$\mathscr{L}^* = \sum_{j,k} \underline{S}_j{}^k \, \underline{\mathsf{g}}^j \otimes \mathsf{h}_k, \qquad \underline{S} = \begin{bmatrix} \underline{S}_0{}^0 & \cdots & \underline{S}_0{}^N \\ \vdots & & \\ \underline{S}_N{}^0 & \cdots & \underline{S}_N{}^N \end{bmatrix}, \tag{A.67}$$

then the right-hand side becomes

$$\begin{aligned}\langle \mathscr{L}^*\mathsf{z}, \mathsf{x}\rangle_{M_2} &= \Big\langle \Big(\sum_{k,l} \underline{S}_k{}^l \underline{\mathsf{g}}^k \otimes \mathsf{h}_l\Big) \sum_j z^j \mathsf{h}_j, \sum_m x^m \mathsf{g}_m \Big\rangle_{M_2} = \sum_{j,k,m} z^j H_{jl} \underline{S}_m{}^l x^m \\ &= z^\top H \underline{S}^\top x \end{aligned} \tag{A.68}$$

and thus, in view of (A.64), (A.63) is established by

$$\underline{S} := L^\top. \tag{A.69}$$

The formerly used coordinates S map coordinates z of z in the original basis $\{\mathsf{h}^k\}_k$ to coordinates Sz of $\mathscr{L}^*\mathsf{z}$ in the original basis $\{\mathsf{g}^k\}_k$, whereas $\underline{S}$ maps

coordinates $\underline{z}$ of z in the dual basis $\{\underline{\mathsf{h}}^k\}_k$ to coordinates $\underline{S}\underline{z}$ of $\mathscr{L}^*\mathsf{z}$ in the dual basis $\{\underline{\mathsf{g}}^k\}_k$. Thus, the latter is even more convenient if z is already represented by $\mathsf{z} = \sum_j \underline{z}_j \underline{\mathsf{h}}^j$. Then the left-hand side of (A.63) simplifies to $\underline{z}^\top L x$ and the right-hand side of (A.63) simplifies to $\underline{z}^\top \underline{S}^\top x$.

The transformation $S = G^{-1}\underline{S}H$ between (A.67) and (A.62), where $G^{-1} = \underline{G}$, can also directly be recognized from (A.40) in

$$\mathscr{L}^* = \sum_{j,k} S^j{}_k \mathsf{g}_j \otimes \underline{\mathsf{h}}^k = \sum_{j,k} \underline{S}_j{}^k \underbrace{\underline{\mathsf{g}}^j}_{\sum_{l=0}^N \underline{G}^{jl}\mathsf{h}_l} \otimes \underbrace{\mathsf{h}_k}_{\sum_{m=0}^N H_{km}\underline{\mathsf{h}}^m}. \tag{A.70}$$

Adjoint of $\mathscr{W}$ The adjoint $\mathscr{W}^*$ of $\mathscr{W}$ must satisfy

$$\langle \mathsf{z}, \mathscr{W}\mathsf{x}\rangle_{M_2} = \langle \mathscr{W}^*\mathsf{z}, \mathsf{x}\rangle_{M_2}. \tag{A.71}$$

According to (A.49), the left-hand side of (A.71) is $z^\top W x$. If $\mathscr{W}^*$ is represented in the same basis as $\mathscr{W}$ in (A.51), i.e.,

$$\mathscr{W}^* = \sum_{j,k} R_{jk} \underline{\mathsf{h}}^j \otimes \underline{\mathsf{h}}^k, \tag{A.72}$$

then the right-hand side of (A.71) becomes

$$\begin{aligned}\langle \mathscr{W}^*\mathsf{z}, \mathsf{x}\rangle_{M_2} &= \Big\langle \Big(\sum_{j,k} R_{jk}\underline{\mathsf{h}}^j \otimes \underline{\mathsf{h}}^k\Big) \sum_l z^l \mathsf{h}_l, \sum_m x^m \mathsf{h}_m \Big\rangle_{M_2} = \sum_{j,k} z^k R_{jk} x^j \\ &= z^\top R^\top x, \end{aligned} \tag{A.73}$$

which is why the coordinates R_{jk} of the adjoint operator must be

$$R := W^\top, \quad \text{where } W = (W_{jk})_{jk} \text{ and } R = (R_{jk})_{jk}. \tag{A.74}$$

Thus, despite of a possibly nonorthonormal basis, self-adjointness $\mathscr{W} = \mathscr{W}^*$ simply requires that W with entries W_{jk} is a symmetric matrix. Note that, for entries W^j_k along the lines of (A.48), this statement would not be true.

A.3 The Operators in the Present Thesis

Recap that, for notational simplicity, the present appendix considers only the scalar case $n = 1$ (but the following results can easily be extended to $n > 1$). The approximated infinitesimal generator $\mathscr{A}^{[N]}$ from (3.55) that is represented by the Legendre-tau-based ODE system matrix is—in contrast to the operators from the last section—not only defined for polynomial arguments. Due to the involved projection, it can be applied to arbitrary arguments in M_2. If, however, only the restriction to the polynomial subspace $\mathrm{span}(\{\mathsf{g}_k\}_k) \subset M_2$ is considered, that projection is not needed and $\mathscr{A}^{[N]}$ behaves like $\mathscr{L}$ discussed above: It maps any such $\mathsf{x} \in \mathrm{span}(\{\mathsf{g}_k\}_k)$ to some $\mathsf{y} \in \mathrm{span}(\{\mathsf{h}_k\}_k)$. Thus, $A^j{}_k$ in the correspondingly defined $\sum_{j=0}^{N} \sum_{k=0}^{N} A^j{}_k \mathsf{h}_j \otimes \underline{\mathsf{g}}^k$ is the (j,k)-th entry of the ODE system matrix, which in terms of Legendre coordinates is given by (3.28). According to Section 3.5.3, on the restriction to that polynomial subspace, $\mathscr{A}^{[N]}$ even coincides with the exact infinitesimal generator $\mathscr{A}$, and thus

$$\mathscr{A}|_{\mathrm{span}(\{\mathsf{g}_k\}_{k\in\{0,\dots,N\}})} = \sum_{j=0}^{N} \sum_{k=0}^{N} A^j{}_k \, \mathsf{h}_j \otimes \underline{\mathsf{g}}^k. \tag{A.75}$$

A.3.1 The Projection Operator in the Legendre Tau Method

The present section discusses the projection operator $\mathrm{Proj}_{\mathrm{cont}}$ from (3.51). In accordance with the ansatz made in the Legendre tau method, this projection maps to a function in the polynomial subspace $\mathrm{span}(\{\mathsf{g}_k\}_k)$. The resulting polynomial is represented by the coordinates from the discretization that is used in the Legendre tau method.

Orthogonality Between the Continuous Basis and the Dual Discontinuous Basis In the previous sections, the relation $b^k = \langle \underline{\mathsf{g}}^k, \mathsf{b} \rangle_{M_2}$ is used to obtain the coordinates b^k from a given function $\mathsf{b} \in \mathrm{span}(\{\mathsf{g}_k\}_{k\in\{0,\dots,N\}})$. This

relation from (A.37) is only based on the orthogonality $\langle \underline{\mathsf{g}}^j, \mathsf{g}_k \rangle_{M_2} = \delta_k^j$. In fact, the dual basis elements $\underline{\mathsf{g}}^k$ are even uniquely defined by that relation. However, this uniqueness only holds in the polynomial space $\operatorname{span}(\{\mathsf{g}_k\}_{k\in\{0,\dots,N\}})$. In the overall space M_2, $\{\underline{\mathsf{g}}^k\}_{k\in\{0,\dots,N\}}$ is not the only set of functions orthogonal to $\{\mathsf{g}_k\}_{k\in\{0,\dots,N\}}$. In particular, the discontinuous functions $\{\underline{\mathsf{h}}^k\}_{k\in\{0,\dots,N\}}$ also satisfy the orthogonality relation

$$\forall j,k \in \{0,\dots,N\}: \qquad \langle \underline{\mathsf{h}}^j, \mathsf{g}_k \rangle_{M_2} = \delta_k^j. \tag{A.76}$$

The latter can be recognized best in mixed coordinates, where, by (A.35),

$$\underline{\mathsf{h}}^k_\chi = \begin{cases} \left[\begin{smallmatrix} \frac{2}{h}\frac{2j+1}{2} p_k(\vartheta(\cdot)) \\ 0 \end{smallmatrix}\right], & k<N, \\ \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right], & k=N, \end{cases} \quad \mathsf{g}_{\chi,k} = \begin{cases} \left[\begin{smallmatrix} p_k(\vartheta(\cdot)) - p_N(\vartheta(\cdot)) \\ 0 \end{smallmatrix}\right], & k<N, \\ \left[\begin{smallmatrix} p_N(\vartheta(\cdot)) \\ 1 \end{smallmatrix}\right], & k=N. \end{cases} \tag{A.77}$$

The orthogonality, however, applies independently from the chosen coordinates since the thus achieved

$$\underbrace{\Big\langle \sum_{j=0}^N \underline{a}_j \underline{\mathsf{g}}^j, \sum_{k=0}^N b^k \mathsf{g}_k \Big\rangle_{M_2}}_{\langle \mathsf{a},\mathsf{b}\rangle_{M_2}} = \underline{a}^\top b = \underbrace{\Big\langle \sum_{j=0}^N \underline{a}_j \underline{\mathsf{h}}^j, \sum_{k=0}^N b^k \mathsf{g}_k \Big\rangle_{M_2}}_{=:\langle \mathsf{d},\mathsf{b}\rangle_{M_2}}, \tag{A.78}$$

also holds after a change of coordinates, which due to (A.36) equally transforms[3] the coordinates of a in $\{\underline{\mathsf{g}}^j\}$ and the coordinates of d in $\{\underline{\mathsf{h}}^j\}$.

[3] The coincidence only holds in the dual basis. In the original basis, by (A.31), $d = H^{-1}\underline{d}$ and $a = G^{-1}\underline{a}$ do no longer coincide and therefore $\langle \mathsf{h}_j, \underline{\mathsf{g}}^k \rangle_{M_2} \neq \delta_j^k$.

Obtaining the Coefficients of $\{\mathsf{g}_k\}_k$ via $\underline{\mathsf{h}}^j$ Instead of $\underline{\mathsf{g}}^j$ As a result, instead of (A.37), the coordinates b of $\mathsf{b} = \sum_{k=0}^{N} b^k \mathsf{g}_k$ can also be obtained via

$$\langle \underline{\mathsf{h}}^i, \mathsf{b} \rangle_{M_2} = \Big\langle \underline{\mathsf{h}}^i, \sum_{k=0}^{N} b^k \mathsf{g}_k \Big\rangle_{M_2} = \sum_{k=0}^{N} \underbrace{\langle \underline{\mathsf{h}}^i, \mathsf{g}_k \rangle_{M_2}}_{\delta^i_k} b^k = b^i. \tag{A.79}$$

The Projection Operator The projection operator

$$\mathrm{Proj}_{\mathrm{cont}}^{[N]} : M_2 \to \mathrm{span}(\{g_k\}_{k\in\{0,\dots,N\}}) \tag{A.80}$$

shall be defined on the overall space M_2 and map to the continuous interpretation of the coordinates from the Legendre tau method as described in (3.51). In terms of the above defined basis functions, it becomes

$$\mathrm{Proj}_{\mathrm{cont}}^{[N]} = \sum_{j=0}^{N}\sum_{k=0}^{N} \delta^j{}_k \mathsf{g}_j \otimes \underline{\mathsf{h}}^k = \sum_{j=0}^{N} \mathsf{g}_j \otimes \underline{\mathsf{h}}^j = \sum_{j=0}^{N} \langle \underline{\mathsf{h}}^j, \cdot \rangle_{M_2}\, \mathsf{g}_j. \tag{A.81}$$

Henceforth, the result of that projection will be denoted by a tilde,

$$\begin{aligned} \tilde{\mathsf{x}} := \mathrm{Proj}_{\mathrm{cont}}^{[N]} \mathsf{x} &= \Big(\sum_{j=0}^{N} \mathsf{g}_j \otimes \underline{\mathsf{h}}^j \Big) \sum_{l=0}^{\infty} x^l \mathsf{g}_l \\ &\overset{(A.76)}{=} \sum_{j=0}^{N} x^j \mathsf{g}_j + \Big(\sum_{j=0}^{N} \mathsf{g}_j \otimes \underline{\mathsf{h}}^j \Big) \sum_{l=N+1}^{\infty} x^l \mathsf{g}_l \quad =: \sum_{j=0}^{N} \tilde{x}^j \mathsf{g}_j. \end{aligned} \tag{A.82}$$

Note that the first term in (A.82) is the series truncation discussed in Section 3.5.2.1. The second term, when taking mixed coordinates for the projection

operator, with $\underline{\mathsf{h}}^j_\chi$ given in (A.78), and taking Legendre coordinates for the series remainder, becomes

$$\Big(\sum_{j=0}^{N} \mathsf{g}_{\chi,j} \otimes \underline{\mathsf{h}}^j_\chi\Big) \sum_{l=N+1}^{\infty} x^l \mathsf{g}_{c,l} \tag{A.83}$$

$$= \sum_{j=0}^{N-1} \mathsf{g}_{\chi,j} \underbrace{\Big\langle \underline{h}^j_\chi, \sum_{l=N+1}^{\infty} x^l \mathsf{g}_{c,l} \Big\rangle_{M_2}}_{0} + \underbrace{\mathsf{g}_{\chi,N}}_{=\mathsf{g}_{c,N}} \Big\langle \underline{h}^N_\chi, \sum_{l=N+1}^{\infty} x^l \mathsf{g}_{c,l} \Big\rangle_{M_2}$$

$$= \left\langle \begin{bmatrix} 0_{n_{[-h,0]}} \\ 1 \end{bmatrix}, \sum_{l=N+1}^{\infty} x^l_c \underline{\mathsf{g}}_{c,l} \right\rangle_{M_2} \mathsf{g}_{c,N}. \tag{A.84}$$

Thus, it extracts the bottom component of the remainder term and provides the correction that achieves for $\tilde{\mathsf{x}}$ and $\mathsf{x} = \sum_{l=0}^{\infty} x^l \mathsf{g}_l$ coincidence in the bottom components of the M_2-tuples (in Legendre coordinates, these are $\sum_{j=0}^{N} \tilde{x}^j_c$ and $\sum_{l=0}^{\infty} x^l_c =: \hat{x}$), see (3.47).

A.3.2 The Legendre-Tau-Based Approximation of the Infinitesimal Generator

The Operator $\mathscr{A}^{[N]}$ As a result, the operator

$$\mathscr{A}^{[N]} \colon M_2 \to \mathrm{span}(\{\mathsf{h}_j\}_{j\in\{0,\dots,N\}}) \tag{A.85}$$

introduced in (3.55) can be written as

$$\mathscr{A}^{[N]} = \mathscr{A}|_{\mathrm{span}(\{\mathsf{g}_k\}_{k\in\{0,\dots,N\}})} \mathrm{Proj}^{[N]}_{\mathrm{cont}} \overset{\substack{(A.75)\\(A.81)}}{=} \Big(\sum_{j,k} A^j{}_k \, \mathsf{h}_j \otimes \underline{\mathsf{g}}^k\Big)\Big(\sum_{l} \mathsf{g}_l \otimes \underline{\mathsf{h}}^l\Big)$$

$$= \sum_{j=0}^{N}\sum_{k=0}^{N} A^j{}_k \, \mathsf{h}_j \otimes \underline{\mathsf{h}}^k. \tag{A.86}$$

Thus, $\mathscr{A}^{[N]}\mathsf{x} = \mathscr{A}\tilde{\mathsf{x}}$. Note that $\mathscr{A}_{\mathrm{cc}}^{[N]}$ from (3.52) instead becomes

$$\mathscr{A}_{\mathrm{cc}}^{[N]} := \mathrm{Proj}_{\mathrm{cont}}^{[N]} \mathscr{A}|_{\mathrm{span}(\{\mathsf{g}_k\}_k)} \mathrm{Proj}_{\mathrm{cont}}^{[N]} = \sum_{j,k} A^j_{\ k}\ \mathsf{g}_j \otimes \underline{\mathsf{h}}^k. \tag{A.87}$$

The Adjoint The operator $(\mathscr{A}^{[N]})^* \colon \mathrm{span}(\{\mathsf{h}_j\}_{j\in\{0,\dots,N\}}) \to M_2$

$$(\mathscr{A}^{[N]})^* = \sum_{j,k} U^j_{\ k} \mathsf{h}_j \otimes \underline{\mathsf{h}}^k = \sum_{j,k} \underline{U}_j^{\ k}\ \underline{\mathsf{h}}^j \otimes \mathsf{h}_k \tag{A.88}$$

$$\text{with } U = H^{-1}A^\top H \quad \text{and} \quad \underline{U} = A^\top \tag{A.89}$$

achieves that

$$\begin{aligned}\langle \mathsf{z}, \mathscr{A}^{[N]}\mathsf{x}\rangle_{M_2} &= \Big\langle \sum_{j=0}^{N} z^j \mathsf{h}_j, \Big(\sum_{k,l=0}^{N} A^k_l \mathsf{h}_k \otimes \underline{\mathsf{h}}^l \Big) \sum_{m=0}^{\infty} x^m \mathsf{g}_m \Big\rangle_{M_2} \\ &= z^\top H A \tilde{x} \end{aligned} \tag{A.90}$$

equals

$$\begin{aligned}\langle (\mathscr{A}^{[N]})^* \mathsf{z}, \mathsf{x}\rangle_{M_2} &= \Big\langle \Big(\sum_{j,k=0}^{N} \underline{U}_j^{\ k} \underline{\mathsf{h}}^j \otimes \mathsf{h}_k \Big) \sum_{l=0}^{N} z^l \mathsf{h}_l, \sum_{m=0}^{\infty} x^m \mathsf{g}_m \Big\rangle_{M_2} \\ &= z^\top H \underline{U}^\top \tilde{x}. \end{aligned} \tag{A.91}$$

Remark A.3.1 (Relation to $\mathscr{A}^*$). *The adjoint of $\mathscr{A}$ from (3.12) is*

$$\mathscr{A}^* \begin{bmatrix} \psi \\ r \end{bmatrix} = \begin{bmatrix} -\psi' \\ \psi(0) + A_0^\top r \end{bmatrix}, \tag{A.92a}$$

$$D(\mathscr{A}^*) = \{ \left[\begin{smallmatrix} \psi \\ r \end{smallmatrix}\right] \in M_2 : \ \psi(-h) = A_1^\top r,\ \psi' \in L_2, \psi \in AC \}, \tag{A.92b}$$

[183, Thm. 5.1], [47, Lem. 3.3.9]. In particular, it entails the negative of a differentiation operator. How can this go together with a transposed differentiation matrix that occurs instead in $A_c^\top$ in (A.89)?

Without loss of generality, consider $\phi\colon [-h,0] \to \mathbb{R}$ *and* $\psi\colon [-h,0] \to \mathbb{R}$, *piecewise defined by* $\psi(\theta) = \begin{cases} \bar{\psi}(\theta), & \theta \in [-h,0), \\ r, & \theta = 0, \end{cases}$ *with* $\bar{\psi} \in AC$ *such that*

$$\mathsf{x} = \begin{bmatrix} \phi \\ \phi(0) \end{bmatrix} \in D(\mathscr{A}) \quad \textit{and} \quad \mathsf{z} = \begin{bmatrix} \psi \\ \psi(0^+) \end{bmatrix} \overset{(\textit{in } M_2)}{=} \begin{bmatrix} \bar{\psi} \\ r \end{bmatrix} \in D(\mathscr{A}^*) \tag{A.93}$$

($\psi(0^-) = \bar{\psi}(0)$ *does not have to equal* $\psi(0^+) = r$*). Compare the following.*

1a) The result of $\langle \mathsf{z}, \mathscr{A}\mathsf{x} \rangle_{M_2}$*:*

$$\begin{aligned} \langle \mathsf{z}, \mathscr{A}\mathsf{x} \rangle_{M_2} &= \left\langle \begin{bmatrix} \psi \\ \psi(0^+) \end{bmatrix}, \mathscr{A} \begin{bmatrix} \phi \\ \phi(0) \end{bmatrix} \right\rangle_{M_2} && \text{(A.94)} \\ &= \left\langle \begin{bmatrix} \psi \\ \psi(0^+) \end{bmatrix}, \begin{bmatrix} \phi' \\ A_0\phi(0) + A_1\phi(-h) \end{bmatrix} \right\rangle_{M_2} \\ &= \underbrace{\langle \psi, \phi' \rangle_{L_2}}_{=:\beta_0} + \underbrace{\langle \psi(0^+), A_0\phi(0) + A_1\phi(-h) \rangle_{\mathbb{R}}}_{=:\beta_A}. && \text{(A.95)} \end{aligned}$$

1b) Where $-\psi'$ *in* $\mathscr{A}^*$ *stems from:*
The term β_0 *in (A.95) is tackled by integration by parts,*

$$\begin{aligned} \beta_0 &= \langle \psi(0^-), \phi(0) \rangle_{\mathbb{R}} - \langle \psi(-h), \phi(-h) \rangle_{\mathbb{R}} - \langle \psi', \phi \rangle_{L_2} && \text{(A.96)} \\ &= \left\langle \begin{bmatrix} -\psi' \\ \psi(0^-) \end{bmatrix}, \begin{bmatrix} \phi \\ \phi(0) \end{bmatrix} \right\rangle_{M_2} - \langle \psi(-h), \phi(-h) \rangle_{\mathbb{R}}, && \text{(A.97)} \\ \beta_A &= \left\langle \begin{bmatrix} 0 \\ A_0^\top \psi(0^+) \end{bmatrix}, \begin{bmatrix} \phi \\ \phi(0) \end{bmatrix} \right\rangle_{M_2} + \langle A_1^\top \psi(0^+), \phi(-h) \rangle_{\mathbb{R}}. && \text{(A.98)} \end{aligned}$$

As a result, if the second terms cancel out in $\beta_0 + \beta_A$ *by*

$$\psi(-h) = A_1^\top \psi(0^+) \tag{A.99}$$

which becomes a requirement for z *being in the domain (A.92b), then*

$$\langle \mathsf{z}, \mathscr{A}\mathsf{x}\rangle_{M_2} \overset{(A.95)}{=} \left\langle \begin{bmatrix} -\psi' \\ \psi(0^-) + A_0^\top \psi(0^+) \end{bmatrix}, \begin{bmatrix} \phi \\ \phi(0) \end{bmatrix} \right\rangle_{M_2} \overset{!}{=} \langle \mathscr{A}^*\mathsf{z}, \mathsf{x}\rangle_{M_2}.$$

2a) The result of $\langle \mathsf{z}, \mathscr{A}^{[N]}\mathsf{x}\rangle_{M_2}$*:*
According to (A.90), the numerical result is simply

$$\langle \mathsf{z}, \mathscr{A}^{[N]}\mathsf{x}\rangle_{M_2} = z^\top H A \tilde{x}. \tag{A.100}$$

Consider $A_0 = A_1 = 0$*, for which* $\beta_A = 0$ *in (A.95). The result will correspondingly be denoted by* $\tilde{\beta}_0$*. The decomposition* $H_{jk} = \langle \mathsf{h}_j, \mathsf{h}_k\rangle_{M_2} = \langle (\mathsf{h}_j)_1, (\mathsf{h}_k)_1\rangle_{L_2} + \langle (\mathsf{h}_j)_2, (\mathsf{h}_k)_2\rangle_{\mathbb{R}}$ *is henceforth used to split the metric coefficient matrix into* $H = H_{L_2} + H_{\mathbb{R}}$*, where, in Legendre coordinates,* $H_{\mathbb{R},c} = 1_{(N+1)\times(N+1)}$ *and* $H_{L_2,c}$ *is the remaining diagonal matrix in (A.10). Then, in Legendre coordinates, with* A_c *from (3.28),*

$$\begin{aligned} \tilde{\beta}_0 &:= z_c^\top H_c A_c|_{(A_0=A_1=0)}\, \tilde{x}_c = z_c^\top (H_{L_2,c} + H_{\mathbb{R},c}) A_c|_{(A_0=A_1=0)}\, \tilde{x}_c \\ &= z_c^\top \left(\begin{bmatrix} 0 & 2 & 0 & \cdots & 2 \\ & 0 & 2 & \cdots & 0 \\ & & 0 & \cdots & 2 \\ & & & \ddots & \\ & & & & 0 \end{bmatrix} + \underbrace{H_{\mathbb{R},c} A_c|_{(A_0=A_1=0)}}_{0} \right) \tilde{x}_c. \end{aligned} \tag{A.101}$$

2b) Follow up how $-\psi'$ *emerged in 1b):*
Compare (A.101) with the result of the integration by parts (A.96), where x *is replaced by its polynomial projection* $\tilde{\mathsf{x}} \in \operatorname{span}(\{\mathsf{g}_k\}_{k\in\{0,\dots N\}})$*. As in (A.90), let* $\mathsf{z} \in \operatorname{span}(\{\mathsf{h}_k\}_{k\in\{0,\dots N\}})$*. Then Legendre coordinates*

$$\tilde{\mathsf{x}} = \operatorname{Proj}_{\text{cont}}^{[N]}\mathsf{x} = \begin{bmatrix} \tilde{\phi} \\ \tilde{\phi}(0) \end{bmatrix} = \sum_{k=0}^{N} \tilde{x}_c^k \mathsf{g}_{c,k}, \qquad \mathsf{z} = \begin{bmatrix} \psi \\ \psi(0^+) \end{bmatrix} = \sum_{k=0}^{N} z_c^k \mathsf{h}_{c,k}$$

$$\tilde{\phi}(0) = \sum_{k=0}^{N} \tilde{x}_c^k, \qquad \psi(0^-) = \sum_{k=0}^{N-1} z_c^k$$

$$\tilde{\phi}(-h) = \sum_{k=0}^{N} (-1)^k \tilde{x}_c^k, \qquad \psi(-h) = \sum_{k=0}^{N-1} (-1)^k z_c^k$$

give in (A.96)

$$\begin{aligned}
\beta_0 &= \langle \psi(0^-), \tilde{\phi}(0) \rangle_{\mathbb{R}} - \langle \psi(-h), \tilde{\phi}(-h) \rangle_{\mathbb{R}} - \langle \psi', \tilde{\phi} \rangle_{L_2} \\
&= \begin{bmatrix} z_{c,0} \\ \vdots \\ z_{c,N-1} \\ z_{c,N} \end{bmatrix}^\top \left(\begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \\ \vdots \\ (-1)^N \\ 0 \end{bmatrix} \begin{bmatrix} 1 & (-1) & \cdots & (-1)^N \end{bmatrix} \right) \begin{bmatrix} \tilde{x}_{c,0} \\ \vdots \\ \tilde{x}_{c,N} \end{bmatrix} \\
&\quad - \left(D_c \begin{bmatrix} z_{c,0} \\ \vdots \\ z_{c,N-1} \\ 0 \end{bmatrix} \right)^\top H_{L_2} \begin{bmatrix} \tilde{x}_{c,0} \\ \vdots \\ \tilde{x}_{c,N} \end{bmatrix} \\
&= \begin{bmatrix} z_{c,0} \\ \vdots \\ z_{c,N-1} \\ z_{c,N} \end{bmatrix}^\top \left(\begin{bmatrix} 0 & 2\ 0 & \cdots & 2 \\ 2 & 0\ 2 & \cdots & 0 \\ 0 & 2\ 0 & \cdots & 2 \\ \vdots & & & \\ 1-(-1)^{N-1} & & & 0 \\ 0 & 0\ 0 & \cdots & 0 \end{bmatrix} - \begin{bmatrix} 0 & & & \\ & \ddots & & \\ 2 & \cdots & 0 & \\ 0 & \cdots & 2\ 0 & \\ 2 & \cdots & 0\ 2\ 0 & \\ 0 & \cdots & 0\ 0\ 0\ 0 \end{bmatrix} \right) \begin{bmatrix} \tilde{x}_{c,0} \\ \vdots \\ \tilde{x}_{c,N} \end{bmatrix} \\
&= z_c^\top \begin{bmatrix} 0 & 2\ 0 & \cdots & 2 \\ & 0\ 2 & \cdots & 0 \\ & & 0 & \cdots\ 2 \\ & & & \ddots \\ & & & 0 \end{bmatrix} \tilde{x}_c.
\end{aligned} \tag{A.102}$$

The last step shows that the occurrence of a lower triangular matrix $-D_c^\top$ has its justification. Of course, $-\langle \psi(-h), \phi(-h) \rangle_{\mathbb{R}}$ from (A.97) may not be neglected, respectively (A.99) must hold.

A.3.3 Lyapunov and Algebraic Riccati Equations and the Involved Operator $\mathscr{P}^{[N]}$

The matrix P (denoted by $\tilde{P}$ in Section 6.1) from a matrix-valued Lyapunov or algebraic Riccati equation in this thesis (no matter in which coordinates P_c, P_χ

or P_y) is used in a quadratic form like (A.50). Therefore, P is considered to be a coordinate representation of

$$\mathscr{P}^{[N]} = \sum_{j=0}^{N} \sum_{k=0}^{N} P_{jk}\, \underline{\mathsf{h}}^j \otimes \underline{\mathsf{h}}^k \tag{A.103}$$

like $\mathscr{W}$ from (A.51), respectively like its extension $\mathrm{Proj}_{\mathrm{cont}}^{*} \mathscr{W} \mathrm{Proj}_{\mathrm{cont}}$ to M_2 which still looks the same.

Lyapunov Equation As a result, the operator-valued Lyapunov equation

$$\left\langle \mathsf{x}, \mathscr{P}^{[N]} \mathscr{A}^{[N]} \mathsf{x} \right\rangle_{M_2} + \left\langle \mathsf{x}, (\mathscr{A}^{[N]})^{*} \mathscr{P}^{[N]} \mathsf{x} \right\rangle_{M_2} = \left\langle \mathsf{x}, \mathscr{Q} \mathsf{x} \right\rangle_{M_2}, \tag{A.104}$$

$\forall \mathsf{x} \in D(\mathscr{A})$, yields in the first term

$$\begin{aligned}
&\left\langle \mathsf{x}, \mathscr{P}^{[N]} \mathscr{A}^{[N]} \mathsf{x} \right\rangle_{M_2} \\
&= \left\langle \sum_{j=0}^{\infty} x^j \mathsf{g}_j,\ \Big(\sum_{k,l=0}^{N} P_{kl} \underline{\mathsf{h}}^k \otimes \underline{\mathsf{h}}^l \Big) \Big(\sum_{o,p=0}^{N} A^o{}_p \mathsf{h}_o \otimes \underline{\mathsf{h}}^p \Big) \sum_{m=0}^{\infty} x^m \mathsf{g}_m \right\rangle_{M_2} \\
&= \sum_{j,l,m=0}^{N} \tilde{x}^j P_{jl} A^l{}_m \tilde{x}^m = \tilde{x}^\top P A \tilde{x}
\end{aligned} \tag{A.105}$$

and in the second term

$$\begin{aligned}
&\left\langle \mathsf{x}, (\mathscr{A}^{[N]})^{*} \mathscr{P}^{[N]} \mathsf{x} \right\rangle_{M_2} \\
&= \left\langle \sum_{j=0}^{\infty} x^j \mathsf{g}_j,\ \Big(\sum_{k,l=0}^{N} \underline{U}_k^{\,l} \underline{\mathsf{h}}^k \otimes \mathsf{h}_l \Big) \Big(\sum_{o,p=0}^{N} P_{op} \underline{\mathsf{h}}^o \otimes \underline{\mathsf{h}}^p \Big) \sum_{m=0}^{\infty} x^m \mathsf{g}_m \right\rangle_{M_2} \\
&= \sum_{j,l,m=0}^{N} \tilde{x}^j \underline{U}_j^{\,l} P_{lm} \tilde{x}^m = \tilde{x}^\top A^\top P \tilde{x}.
\end{aligned} \tag{A.106}$$

Therefore, the operator-valued Lyapunov equation leads to the matrix-valued Lyapunov equation considered in this thesis.

Note that, with $\mathscr{A}_{\mathrm{cc}}^{[N]}$ from (A.87), the same results would arise in (A.105) and (A.106).

Algebraic Riccati Equation The algebraic Riccati equation from (5.65) becomes in terms of the approximated operators $\mathscr{A}^{[N]}$ and $\mathscr{P}^{[N]}$

$$\begin{aligned}&\langle \mathsf{x}, \mathscr{P}^{[N]}\mathscr{A}^{[N]}\mathsf{x}\rangle_{M_2} + \langle \mathsf{x}, (\mathscr{A}^{[N]})^*\mathscr{P}^{[N]}\mathsf{x}\rangle_{M_2} \\ &= -\langle \mathsf{x}, \mathscr{Q}\mathsf{x}\rangle_{M_2} - \langle \mathscr{B}^*\mathscr{P}^{[N]}\mathsf{x}, (-\Pi_{aa})^{-1}\mathscr{B}^*\mathscr{P}^{[N]}\mathsf{x}\rangle_{\mathbb{R}^m}. \end{aligned} \tag{A.107}$$

In view of the common parts with (A.104), it only remains to discuss the last term on the right-hand side. Note that $\mathscr{B}\colon \mathbb{R}^m \to M_2$ from (5.66) can be written as (again considering only one component, $n = 1$, i.e., assuming that $B \in \mathbb{R}^{1\times m}$)

$$\mathscr{B} = \sum_{j=0}^{N}\sum_{k=1}^{m} L^j{}_k \mathsf{h}_j \otimes \underline{e}^k = \sum_{k=1}^{m} B_{\cdot k} \underbrace{\begin{bmatrix} 0_{[-h,0]} \\ 1 \end{bmatrix}}_{=\mathsf{h}_{\chi,N}=\mathsf{h}_{c,N}} \otimes \underline{e}^k, \tag{A.108}$$

$$\text{i.e., } L_\chi = L_c = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \\ L^N_{c,\,1} & \dots & L^N_{c,\,m} \end{bmatrix} = \begin{bmatrix} 0_{N\times m} \\ B \end{bmatrix}, \tag{A.109}$$

where $\{\underline{e}^k\}_{k\in\{1,\dots,m\}} = \{e_k\}_k$ is the canonical basis of $\mathbb{R}^m$. In this sense, (6.8) is a coordinate representation of $\mathscr{B}$. Analogously to (A.67), the adjoint is $\mathscr{B}^* = \sum_{j=1}^m \sum_{k=0}^N \underline{S}_j{}^k \underline{e}^j \otimes \mathsf{h}_k$ with $\underline{S} = L^\top$ and therefore

$$\begin{aligned}\mathscr{B}^*\mathscr{P}^{[N]}\mathsf{x} &= \Big(\sum_{j=1}^{m}\sum_{k=0}^{N} \underline{S}_j{}^k \underline{e}^j \otimes \mathsf{h}_k\Big)\Big(\sum_{o=0}^{N}\sum_{p=0}^{N} P_{op}\underline{\mathsf{h}}^o \otimes \underline{\mathsf{h}}^p\Big)\Big(\sum_{l=0}^{\infty} x^l \mathsf{g}_l\Big) \\ &= \sum_{j=1}^{m}\sum_{k=0}^{N}\sum_{l=0}^{N} \underline{S}_j{}^k P_{kl}\tilde{x}^l\, \underline{e}^j, \end{aligned} \tag{A.110}$$

respectively

$$\begin{aligned}\langle \mathscr{B}^* \mathscr{P}^{[N]} \mathsf{x}, (-\Pi_{aa})^{-1} \mathscr{B}^* \mathscr{P}^{[N]} \mathsf{x} \rangle_{\mathbb{R}^m} &= \tilde{x}^\top P^\top \underline{S}^\top (-\Pi_{aa})^{-1} \underline{S} P \tilde{x} \\ &= \tilde{x}^\top P L (-\Pi_{aa})^{-1} L^\top P \tilde{x}, \end{aligned} \tag{A.111}$$

which coincides with the corresponding term in (7.24).

Operator norm The operator norm of $\mathscr{P}^{[N]}$ is still the same as the operator norm of its restriction $\mathscr{P}^{[N]}|_{\mathrm{span}(\{\mathsf{h}_k\}_k)}$ to the finite-dimensional subspace $\mathrm{span}(\{\mathsf{h}_k\}_k) \subset M_2$, for which the result of Appendix A.2.3 applies,

$$\|\mathscr{P}^{[N]}\| = \|H^{-1/2} P H^{-1/2}\|_2. \tag{A.112}$$

The reason for the equality

$$\|\mathscr{P}^{[N]}\|^2 = \sup_{\mathsf{x} \in M_2 \setminus \{0_{M_2}\}} \frac{\|\mathscr{P}^{[N]} \mathsf{x}\|_{M_2}^2}{\|\mathsf{x}\|_{M_2}^2} = \sup_{\tilde{\mathsf{x}} \in \mathrm{span}(\{\mathsf{h}_k\}_k) \setminus \{0_{M_2}\}} \frac{\|\mathscr{P}^{[N]} \tilde{\mathsf{x}}\|_{M_2}^2}{\|\tilde{\mathsf{x}}\|_{M_2}^2}$$

is that the supremum over M_2 among all $\mathsf{x} \in M_2$ that have in common to be projected to $\tilde{\mathsf{x}} \in \mathrm{span}(\{\mathsf{h}_k\}_k)$ is attained by the one that achieves the infimum possible norm value $\|\tilde{\mathsf{x}}\|_{M_2}$ in the denominator, i.e., by $\mathsf{x} = \tilde{\mathsf{x}} \in \mathrm{span}(\{\mathsf{h}_k\}_k)$.

Written-Out Result of $\mathscr{P}^{[N]}$ In mixed coordinates, the basis functions $\underline{\mathsf{h}}_\chi^k$ that occur in $\mathscr{P}^{[N]}$ from (A.103) are very simple, see (A.77). In fact, $\mathscr{P}^{[N]}$ can even compactly be written out in the integral form of (4.48) with explicitly given kernel functions. To establish the separation between $P_{\mathrm{zz}}(\xi, \eta)$ and $P_{\mathrm{zz,diag}}(\eta)$ in (4.11a), the splitting approach from (4.46) should be applied and explicit kernel functions only be computed for the remaining operator $\mathscr{P}_0^{[N]}$. See also Remark 4.1.1.

To make the explicit integral formula of $\mathscr{P}^{[N]}$ visible, note that, by (A.103),

$$\mathscr{P}^{[N]}\left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right] = \Big(\sum_{j,k} P_{jk}\ \underline{\mathsf{h}}^j \otimes \underline{\mathsf{h}}^k\Big)\left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right] \overset{(A.44)}{=} \sum_{j=0}^{N}\sum_{k=0}^{N} P_{jk}\underline{\mathsf{h}}^j \big\langle \underline{\mathsf{h}}^k, \left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right]\big\rangle_{M_2}.$$

According to (A.77), the dual basis functions are, in mixed coordinates, given by $\underline{\mathsf{h}}^j_\chi = \begin{cases} \left[\begin{smallmatrix}\frac{2}{h}\frac{2j+1}{2}p_j(\vartheta(\cdot))\\ 0\end{smallmatrix}\right] & \text{if } j \neq N,\\ \left[\begin{smallmatrix}0\\ 1\end{smallmatrix}\right] & \text{if } j = N,\end{cases}$. Using the latter, in the end, yields the written out result

$$\mathscr{P}^{[N]}\left[\begin{smallmatrix}\phi\\ \phi(0)\end{smallmatrix}\right] = \begin{bmatrix} \overbrace{\int_{-h}^{0} \sum_{j=0}^{N-1}\sum_{k=0}^{N-1} P_{\chi,jk}\frac{2j+1}{h}p_j(\vartheta(\cdot))\frac{2k+1}{h}p_k(\vartheta(\eta))}^{P_{\mathrm{zz}}(\cdot,\eta)}\,\phi(\eta)\,\mathrm{d}\eta + \underbrace{\sum_{j=0}^{N-1} P_{\chi,jN}\frac{2j+1}{h}p_j(\vartheta(\cdot))}_{P_{\mathrm{zx}}(\cdot)}\,\phi(0) \\ \int_{-h}^{0} \underbrace{\sum_{k=0}^{N-1} P_{\chi,Nk}\frac{2k+1}{h}p_k(\vartheta(\eta))}_{P_{\mathrm{xz}}(\eta)}\,\phi(\eta)\,\mathrm{d}\eta + \underbrace{P_{\chi,NN}}_{P_{\mathrm{xx}}}\,\phi(0) \end{bmatrix}. \tag{A.113}$$

Thus, the structure is the same as the one encountered in (4.48).

A.4 Revisiting the Main Points of the Chapter

- The chapter uses basic concepts from tensor algebra like metric coefficients, dual bases, and dyadic products to explain how to deal with coordinate representations of polynomials and polynomial operators in M_2.
- The basis functions of two $(N+1)$-dimensional subspaces of M_2 are considered:
  - $\{\mathsf{g}_k\}_k$ amounts to a polynomial of degree at most N and its continuous end point, motivated by the polynomial ansatz in the Legendre tau method.

  - $\{\mathsf{h}_k\}_k$ amounts to the discontinuous interpretation of the coordinates, describing a polynomial of degree at most $N-1$ and a discontinuous end point. In Section 3.5.2.2, this interpretation might not have seemed compelling. However, due to the projection operator $\mathrm{Proj}_{\mathrm{cont}}$ from Appendix A.3.1, the associated dual functions $\{\underline{\mathsf{h}}^k\}_k$ turn out to be crucial. In the end, both operators $\mathscr{A}^{[N]}$ and $\mathscr{P}^{[N]}$ do not even rely on $\{\mathsf{g}_k\}_k$ but only on $\{\mathsf{h}_k\}_k$ and the associated dual functions $\{\underline{\mathsf{h}}^k\}_k$.

- The metric coefficients, which are the inner products of the basis functions, are given for Legendre coordinates and mixed coordinates in both interpretations (Appendix A.1).

- None of the considered bases is orthonormal. However, at least $\{\mathsf{h}_{\chi,k}\}_k$ from the discontinuous interpretation in mixed coordinates is orthogonal.

- The dual basis functions are defined as being orthogonal to the original basis functions. They can be computed from the original basis functions based on the inverse of the metric coefficient matrix. In mixed coordinates, the basis functions $\{\underline{\mathsf{h}}^k_\chi\}_k$ from the discontinuous interpretation are particularly simple, see (A.77).

- In the matrices that are coordinate representations of operators, each matrix entry is associated with a dyadic product $\otimes$ between a basis function of the codomain and a (dual) basis function of the domain of the operator.

  - The matrix $A \in \{A_c, A_\chi, A_y, \ldots\}$ from the Legendre tau method represents the operator $\mathscr{A}^{[N]} = \sum_{j,k} A^j_k \; \mathsf{h}_j \otimes \underline{\mathsf{h}}^k$.

  - The matrix $P \in \{P_c, P_\chi, P_y, \ldots\}$ in the quadratic form that approximates an LK functional represents $\mathscr{P}^{[N]} = \sum_{j,k} P_{jk} \; \underline{\mathsf{h}}^j \otimes \underline{\mathsf{h}}^k$.

- Despite of a nonorthonormal basis, the operator $\mathscr{P}^{[N]}$ is self-adjoint if the matrix P is symmetric. Due to the different representation, such a result would, e.g., not be true for the matrix A that describes the operator $\mathscr{A}^{[N]}$.

- The operator norm of $\mathscr{P}^{[N]}$, which becomes relevant in the statement of uniform boundedness in Lemma 4.6.5, can be computed from the coordinate matrix P via (A.112).
- The operator $\mathscr{P}^{[N]}$ becomes particularly simple in mixed coordinates P_χ, for which $\mathscr{P}^{[N]}$ is made explicit in (A.113).
- Appendix A.3.3 confirms that the matrix-valued Lyapunov and algebraic Riccati equation considered in the main part of the thesis exactly address the operator-valued Lyapunov and algebraic Riccati equation with $\mathscr{A}^{[N]}$ and $\mathscr{P}^{[N]}$.

B Appendix: Further Aspects

B.1 Addendum to the Numerical-Integration-Based Approach

Besides of the ODE-based approach, proposed in Section 4.1, a numerical-integration based approach is proposed in Section 4.2 for the sake of validation. The latter makes use of the known formula of complete-type and related LK functionals $V(\phi)$ given in (1.15).

B.1.1 Delay Lyapunov Matrix Function

In contrast to the ODE-based approach, the numerical integration of the LK functional formula (1.15) requires explicit knowledge of the delay Lyapunov matrix function Ψ on which (1.15) relies.

This delay Lyapunov matrix function, associated to a given matrix $Q \succ 0_{n\times n}$ and to the given time-delay system (1.4), is defined as being a solution $\Psi \in C([-h,h],\mathbb{R}^{n\times n})$ of

$$\tfrac{\mathrm{d}}{\mathrm{d}s}\Psi(s) = \Psi(s)A_0 + \Psi(s-h)A_1, \qquad s\in(0,h], \tag{B.1a}$$

$$\Psi(s) = \Psi^\top(-s), \qquad s\in[-h,0], \tag{B.1b}$$

$$\Psi(0)A_0 + \Psi(-h)A_1 = -A_0^\top\Psi(0) - A_1^\top\Psi(h) - Q, \tag{B.1c}$$

[110, Thm. 2.6]. These conditions are commonly referred to as the dynamic property (B.1a), the symmetry property (B.1b), and the algebraic property (B.1c).

Proposition B.1.1 (Lyapunov condition, unique solution, [110, Thm. 2.8]). *Let $\sigma(\mathcal{A}) = \{s \in \mathbb{C} : (1.6)\}$ denote the characteristic roots of (1.4). If $\sigma(\mathcal{A}) \cap (-\sigma(\mathcal{A})) = \emptyset$, then a unique solution $\Psi \in C([-h, h], \mathbb{R}^{n\times n})$ of (B.1) exists.*

For the considered system class, a semi-analytical solution, which still relies on a matrix exponential, can be derived in vectorized form.

Proposition B.1.2 (Semi-analytical solution for $\Psi(s)$, cf. [69]). *If Proposition B.1.1 holds, then the solution $\Psi\colon [-h, h] \to \mathbb{R}^{n\times n}$ of (B.1) is given by*

$$\begin{bmatrix} \operatorname{vec}(\Psi(s-h)) \\ \operatorname{vec}(\Psi(s)) \end{bmatrix} = E(s) y_0, \quad s \in [0, h] \tag{B.2}$$

where $E\colon [0, h] \to \mathbb{R}^{2n^2\times 2n^2}$ describes the matrix exponential

$$E(s) = \exp\left(\begin{bmatrix} -I_n \otimes A_0^\top & -I_n \otimes A_1^\top \\ A_1^\top \otimes I_n & A_0^\top \otimes I_n \end{bmatrix} s \right) \tag{B.3}$$

and $y_0 \in \mathbb{R}^{2n^2}$ is a solution of

$$\begin{aligned} &\left(\begin{bmatrix} I_{n^2} & 0_{n^2\times n^2} \\ 0_{n^2\times n^2} & I_n \otimes A_1^\top \end{bmatrix} E(h) + \begin{bmatrix} 0_{n^2\times n^2} & -I_{n^2} \\ A_1^\top \otimes I_n & I_n \otimes A_0^\top + A_0^\top \otimes I_n \end{bmatrix} \right) y_0 \\ &= \begin{bmatrix} 0_{n^2} \\ -\operatorname{vec}(Q) \end{bmatrix}. \end{aligned} \tag{B.4}$$

B.1.2 A Factorization of the Numerical Integration of Complete-Type LK Functionals

The numerical integration result from (4.16) reveals some further structure, when incorporating the kernel functions from (4.11b). The following proposition primarily addresses P_y^{quad} from Section 4.2. However, it is not restricted to polynomial quadrature rules, like Clenshaw–Curtis quadrature or Gauss with added

zero-weighted boundary nodes. It equally well holds for Riemann-sum approximations of the integrals.

In particular, the proposition makes clear that

$$(\Psi(\tilde{\theta}_j - \tilde{\theta}_k))_{jk} = \begin{bmatrix} \Psi(0) & \Psi(\tilde{\theta}_0-\tilde{\theta}_1) & \cdots & & \Psi(\tilde{\theta}_0-\tilde{\theta}_N) \\ \Psi(\tilde{\theta}_1-\tilde{\theta}_0) & \Psi(0) & & & \vdots \\ & & \ddots & & \\ \vdots & & & & \Psi(\tilde{\theta}_{N-1}-\tilde{\theta}_N) \\ \Psi(\tilde{\theta}_N-\tilde{\theta}_0) & \cdots & & \Psi(\tilde{\theta}_N-\tilde{\theta}_{N-1}) & \Psi(0) \end{bmatrix}, \quad \text{(B.5)}$$

which, due to (4.11b), is closely related to $(P_{\mathrm{zz}}(\tilde{\theta}_j, \tilde{\theta}_k))_{jk}$ in (4.16), is rather decisive. The matrix (B.5) is known to be positive semidefinite whenever the RFDE equilibrium is exponentially stable, no matter how coarse the grid is, see [62, Thm. 9] (also being valid for non-equidistant grids $\{\tilde{\theta}_k\}_k$). An immediate consequence of the following proposition is that the same holds for P_y^{quad} from Section 4.2.

Proposition B.1.3 (Factorization of the numerical integration result (4.16)). *For any $u \in C([-h, 0], \mathbb{R})$, let*

$$\int_{-h}^{0} u(\theta)\,\mathrm{d}\theta \approx \sum_{k=0}^{N} u(\tilde{\theta}_k) w_k$$

be a numerical integration rule relying on $\tilde{\theta}_0 = -h < \tilde{\theta}_1 < \ldots < \tilde{\theta}_{N-1} < \tilde{\theta}_N = 0$ and corresponding integration weights $(w_k)_{k\in\{0,\ldots,N\}} \in \mathbb{R}^{N+1}$. Applying such a rule to (4.11a) with (4.11b) gives

$$V(\phi) \approx \begin{bmatrix} \phi(\tilde{\theta}_0) \\ \vdots \\ \phi(\tilde{\theta}_N) \end{bmatrix}^{\top} P_y^{\mathrm{quad}} \begin{bmatrix} \phi(\tilde{\theta}_0) \\ \vdots \\ \phi(\tilde{\theta}_N) \end{bmatrix}, \qquad P_y^{\mathrm{quad}} = S^{\top} \big(\Psi(\tilde{\theta}_j - \tilde{\theta}_k)\big)_{jk} S + D,$$

where $S = \mathrm{diag}((w_k)_k) \otimes A_1 + \left[\begin{smallmatrix} 0_{n\times nN} & I_n \\ 0_{nN\times nN} & 0_{nN\times n} \end{smallmatrix}\right]$, and where the block diagonal matrix $D = \mathrm{blkdiag}((w_k(Q_1 + (h + \tilde{\theta}_k)Q_2))_k)$ vanishes if $Q_1 = Q_2 = 0_{n\times n}$ and otherwise inherits the positive semidefiniteness from Q_1, Q_2.

Proof. Let

$$y = \begin{bmatrix} \phi(\tilde\theta_0) \\ \vdots \\ \phi(\tilde\theta_N) \end{bmatrix} = \begin{bmatrix} \phi(-h) \\ \vdots \\ \phi(0) \end{bmatrix}, \quad W = \begin{bmatrix} w_0 A_1 & & \\ & \ddots & \\ & & w_N A_1 \end{bmatrix}, \quad E^\top = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & 0 & \\ I_n & 0 & \cdots\ 0 \end{bmatrix}.$$

Then ($\frac{1}{2}$ times) the second term of $V(\phi)$ in (4.11a) becomes (where, in fact, (B.6) only precises what is already given as the second term in (4.16))

$$\begin{aligned}
&\int_{-h}^{0} \phi^\top(0) P_{\mathrm{xz}}(\theta)\phi(\theta)\,\mathrm{d}\theta = \int_{-h}^{0} \underbrace{\phi^\top(0)\Psi(-h-\theta)A_1\phi(\theta)}_{u(\theta)}\,\mathrm{d}\theta \\
&\approx \sum_{k=0}^{N} \underbrace{\phi^\top(0)\Psi(-h-\tilde\theta_k)A_1\phi(\tilde\theta_k)}_{u(\tilde\theta_k)} w_k \\
&= \underbrace{\phi^\top(0)}_{\phi^\top(\tilde\theta_N)} \Big(\Psi(\underbrace{-h}_{\tilde\theta_0} - \tilde\theta_0)A_1\phi(\tilde\theta_0)w_0 + \ldots + \Psi(\underbrace{-h}_{\tilde\theta_0} - \tilde\theta_N)A_1\phi(\tilde\theta_N)w_N \Big) \\
&= \begin{bmatrix} \phi(\tilde\theta_0) \\ \vdots \\ \phi(\tilde\theta_N) \end{bmatrix}^\top \begin{bmatrix} 0 & \ldots & 0 \\ \vdots & & \vdots \\ w_0\Psi(\tilde\theta_0-\tilde\theta_0)A_1 & \ldots & w_N\Psi(\tilde\theta_0-\tilde\theta_N)A_1 \end{bmatrix} \begin{bmatrix} \phi(\tilde\theta_0) \\ \vdots \\ \phi(\tilde\theta_N) \end{bmatrix} \qquad \text{(B.6)} \\
&= \begin{bmatrix} \phi(\theta_0) \\ \vdots \\ \phi(\theta_N) \end{bmatrix}^\top E^\top \begin{bmatrix} \Psi(\tilde\theta_0-\tilde\theta_0) & \cdots & \Psi(\tilde\theta_0-\tilde\theta_N) \\ \vdots & \ddots & \vdots \\ \Psi(\tilde\theta_N-\tilde\theta_0) & \cdots & \Psi(\tilde\theta_N-\tilde\theta_N) \end{bmatrix} W \begin{bmatrix} \phi(\theta_0) \\ \vdots \\ \phi(\theta_N) \end{bmatrix} \\
&= y^\top E^\top \big(\Psi(\tilde\theta_j - \tilde\theta_k)\big)_{jk} W\, y. \qquad \text{(B.7)}
\end{aligned}$$

Being scalar, the latter is also equal to its transpose

$$\int_{-h}^{0} \phi^\top(\theta) P_{\mathrm{xz}}^\top(\theta)\phi(0)\,\mathrm{d}\theta = y^\top W^\top \big(\underbrace{\Psi^\top(\tilde\theta_k - \tilde\theta_j)}_{\overset{\text{(B.1b)}}{=}\Psi(\tilde\theta_j-\tilde\theta_k)} \big)_{jk} E\, y. \qquad \text{(B.8)}$$

The first term of $V(\phi)$ in (4.11a) is

$$\phi^\top(0) P_{\mathrm{xx}}\phi(0) = \phi^\top(0)\Psi(0)\phi(0) = y^\top E^\top \big(\Psi(\tilde\theta_j - \tilde\theta_k)\big)_{jk} E y, \qquad \text{(B.9)}$$

and the third one

$$\begin{aligned}\int_{-h}^{0}\int_{-h}^{0}\phi^\top(\xi)P_{\mathrm{zz}}(\xi,\theta)\phi(\theta)\,\mathrm{d}\theta\mathrm{d}\xi &= \int_{-h}^{0}\int_{-h}^{0}\phi^\top(\xi)A_1^\top\Psi(\xi-\theta)A_1\phi(\theta)\,\mathrm{d}\theta\mathrm{d}\xi\\ &\approx \sum_{j=0}^{N}w_j\sum_{k=0}^{N}w_k\phi^\top(\tilde\theta_j)A_1^\top\Psi(\tilde\theta_j-\tilde\theta_k)A_1\phi(\tilde\theta_k)\\ &= y^\top W^\top\big(\Psi(\tilde\theta_j-\tilde\theta_k)\big)_{jk}W\,y. \end{aligned}\tag{B.10}$$

Altogether, abbreviating $P=\big(\Psi(\tilde\theta_j-\tilde\theta_k)\big)_{jk}$,

$$W^\top PW+W^\top PE+E^\top PW+E^\top PE=(W+E)^\top P(W+E)=S^\top PS.$$

The remaining term D is as in (4.16). □

B.1.3 Some Remarks

Some remarks are in order.

Remark B.1.4 (ODE-based vs. numerical-integration-based approach). *Both the ODE-based approach from Section 4.1.1 and the numerical-integration-based approach from Section 4.2 provide an approximation $V_y(y)=y^\top P_y y$. The former seeks for an approximative solution of the defining equation (1.14). In contrast, the latter already starts with the exact knowledge of the LK functional (1.15), presupposing knowledge of Ψ, and only has to describe a discretization thereof. In so far, the numerical-integration-based approach is related to discretizations of the known $V(\phi)$ already proposed in the literature—be it based on piecewise cubic polynomials that approximate ϕ [136, 134] or, recently, on a Legendre series truncation of ϕ [10] (also used in [12, 11]), or a certain fundamental-matrix-dependent discontinuous approximation of ϕ [72, 60, 62]. With the exception of the latter discontinuous approximation (which, however, only addresses zero Q_0 and Q_2), integral terms with Ψ must still be evaluated in*

these approaches, which is not the case for the interpolatory quadrature proposed in Section 4.2.

Remark B.1.5 (The role of the delay Lyapunov matrix function). *In a common perception, the delay Lyapunov matrix function Ψ (defined in (B.1)) is considered as being the counterpart to the Lyapunov equation solution P from the finite-dimensional template ii in Section 1.2. The more direct counterpart, however, is addressed by the matrix P_y (in fact, P_y represents an approximation of an operator $\mathscr{P}$ on M_2 that is introduced in Section 4.6.2.1). Take a look at the ODE-based approach in Section 4.1.1. If A_y is Hurwitz, then the solution P_y of the Lyapunov equation (4.3) is well known to be expressible by an improper integral, yielding*

$$V_y(y) = y^\top P_y y, \quad \text{with } P_y = \int_0^\infty \mathrm{e}^{A_y^\top t} Q_y \mathrm{e}^{A_y t}\,\mathrm{d}t. \tag{B.11}$$

This formula for V_y and the LK functional formula (1.15) with $Q_1 = Q_2 = 0_{n\times n}$ can be proven quite analogously, see [110]. However, the derivation of (1.15) does not consider the solution operator $\mathcal{T}(t)\colon C \to C$, respectively $\mathscr{T}(t)\colon M_2 \to M_2$ in the Hilbert space setting, which would be the counterpart to $\mathrm{e}^{A_y t}$ in (B.11). Instead the solution formula for $x_0 \mapsto x(t)$ is used. The latter relies on the RFDE fundamental matrix function X. See Table B.1. Therefore, in the resulting formula for $V(\phi)$, terms of the form $\int_0^\infty X^\top(t)\tilde{Q}X(t+s)\,\mathrm{d}t$ occur. This is where the delay Lyapunov matrix function Ψ originates from. It is defined to encapsulate the decisive expression

$$\Psi(s;\tilde{Q}) = \int_0^\infty X^\top(t)\tilde{Q}X(t+s)\,\mathrm{d}t, \tag{B.12}$$

[110, Def. 2.4]. That substitution finally leads to the famous formula of the LK functional (1.15). As a consequence, the involved Ψ merely relies on X. However, $X(t)$ represents only a part of the whole solution operator, which is actually required (see (4.59) relying on $\mathscr{T}(t)$ in M_2) and readily approximated by the whole matrix exponential $\mathrm{e}^{A_y t}$ in (B.11). That is why the ODE-based approach neither needs to approximate the fundamental matrix function X nor the resulting

RFDE state at time $t \geq 0$ $x_t \in C([-h, 0], \mathbb{R}^n)$	ODE state at $t \geq 0$ $y(t) \in \mathbb{R}^{n(N+1)}$
Solution value at time $t \geq -h$ $x(t) = x_t(0) \in \mathbb{R}^n$	Lower subvector of $y(t)$ $\hat{x}(t) = y^N(t) \in \mathbb{R}^n$
Solution operator $\mathcal{T}(t)$ for $x_0 \mapsto x_t$ $x_t = \mathcal{T}(t)x_0$	Matrix exponential $y(t) = \mathrm{e}^{A_y t} y(0)$
Solution formula for $x_0 \mapsto x(t)$ $x(t) = \int_{-h}^{0} X(t-\eta-h)A_1 x_0(\eta)\,\mathrm{d}\eta + X(t)x_0(0)$	Lower block row of $\mathrm{e}^{A_y t}$ $y^N(t) = (\mathrm{e}^{A_y t})^{(N,:)} y(0)$
Fundamental matrix at $t \geq 0$ $X(t) \in \mathbb{R}^{n\times n}$, which is the matrix solution for the matrix-valued initial condition $X(t) = \begin{cases} 0_{n\times n} & \text{if } t < 0, \\ I_n & \text{if } t = 0 \end{cases}$	Right lower submatrix $(\mathrm{e}^{A_y t})^{(N,N)}$ if $t \geq 0$

Table B.1: Correspondences between the RFDE and the approximating ODE. (No claim of convergence of the solution operator approximation in C. However, for the Legendre tau method, a convergence statement in M_2 is available, see Lemma 3.5.6.)

delay Lyapunov matrix function Ψ as suggested by the LK functional formula (1.15). Instead, the ODE-based approach readily yields an approximation of the whole LK functional V in one step.

Remark B.1.6. *The numerical result for $\Psi(0;\tilde{Q})$ in [101, 140] is founded on a comparison between the improper integral for $\Psi(0;\tilde{Q})$ from (B.12) and the improper integral for P_y (with $Q_1 = Q_2 = 0_{n\times n}$ in (4.4)) from (B.11), incorporating the last row of Table B.1. These improper integrals, however, are only applicable if the solution decays exponentially. For that reason, the result has been restricted to exponentially stable equilibria. However, comparing $V_y(y)$ in (4.10) with $V(\phi)$ in (4.11), suggests as well that the right lower submatrix $(P_y)^{(N,N)} = P_{y,xx}$ approximates $\Psi(0;\tilde{Q}) = P_{\mathrm{xx}}$, even without a stability restriction.*

B.2 A Lemma Related to the Equicontinuity of Solutions for Linear RFDEs

The proof of Theorem 5.4.1 relies on the following lemma. Although the lemma intends to express the explicit dependence of $\|x(t_1) - x(t_0)\|$ on the individual norm of the initial function $\|x_{t_0}\|_C$, it is closely related to equicontinuity of RFDE solutions that share to have x_{t_0} from a common bounded set. In fact, fixing a norm ball of initial functions with $\|x_{t_0}\|_C < r$ for some given $r > 0$, this equicontinuity at t_0 for the resulting family of solutions $x\colon \mathbb{R}_{\geq t_0} \to \mathbb{R}^n$ could directly be concluded from the lemma.

Lemma B.2.1. *Solutions x of (5.36) satisfy*

$$\begin{aligned} &\forall \alpha \geq 0, \exists \delta(\alpha) \geq 0, \forall t_1 \geq t_0 \geq 0 : \\ &t_1 - t_0 \leq \delta(\alpha) \Longrightarrow \|x(t_1) - x(t_0)\| \leq \alpha \|x_{t_0}\|_C, \end{aligned} \tag{B.13}$$

and $\delta\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}; \alpha \mapsto \delta(\alpha)$ can be chosen linear.

Proof. Let (5.36) be $\dot{x}(t) = \tilde{A}_0 x(t) + \tilde{A}_1 x(t-h)$. Then

$$\begin{aligned} \|x(t_1) - x(t_0)\| &= \left\| \int_{t_0}^{t_1} \left(\tilde{A}_0 x(t) + \tilde{A}_1 x(t-h) \right) \mathrm{d}t \right\| \\ &\leq (t_1 - t_0) \left(\|\tilde{A}_0\| + \|\tilde{A}_1\| \right) \max_{t \in [t_0 - h, t_1]} \|x(t)\|. \end{aligned} \tag{B.14}$$

Due to the uniform stability, $\forall \varepsilon_{\mathrm{s}} > 0, \exists \delta_{\mathrm{s}}(\varepsilon_{\mathrm{s}}) > 0, \forall t_0 \geq 0 : \|\phi\|_C \leq \delta_{\mathrm{s}}(\varepsilon_{\mathrm{s}}) \Longrightarrow \forall t \geq t_0 - h : \|x(t; t_0, \phi)\| \leq \varepsilon_{\mathrm{s}}$. Thus, $\psi = \frac{1}{b}\phi$ with $b := \frac{\|\phi\|_C}{\delta_{\mathrm{s}}(1)}$, i.e., $\|\psi\|_C = \delta_{\mathrm{s}}(1)$, implies $\|x(t; t_0, \psi)\| \leq 1$. By linearity, $x(t; t_0, \psi) = \frac{1}{b} x(t; t_0, \phi)$. Hence, $\forall t \geq t_0 - h : \|x(t; t_0, \phi)\| \leq b = \frac{\|\phi\|_C}{\delta_{\mathrm{s}}(1)}$ with $\phi = x_{t_0}$. Consequently, (B.13) is obtained by choosing $t_1 - t_0 \leq \delta(\alpha) := \frac{\delta_{\mathrm{s}}(1)\alpha}{\|A_0\| + \|A_1\|}$ in (B.14). □

B.3 Two Lemmas on Bessel Polynomials

Bessel polynomials are known to satisfy the recurrence relation [114]

$$b_0(z) \equiv 1, \tag{B.15}$$

$$b_1(z) = z + 1, \tag{B.16}$$

$$b_{N+1}(z) = \underbrace{(2N+1)z}_{=:\beta_N(z)} b_N(z) + b_{N-1}(z), \tag{B.17}$$

based on which the following two lemmas establish alternative expressions for $b_N(z)$, laying the foundation for Theorem 7.3.4.

Lemma B.3.1. *For $N \geq 2$, the N-th Bessel polynomial is*

$$b_N(z) = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \beta_0(z) & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} \beta_{N-2}(z) & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \beta_{N-1}(z) \\ 1 \end{bmatrix}, \tag{B.18}$$

where $\beta_j(z) = (2j+1)z$, $j \geq 0$.

Proof. Let $[h_{N,1}(z) \;\; h_{N,2}(z)] = [1\ 1] \prod_{k=0}^{N-2} \left[\begin{smallmatrix} \beta_k(z) & 1 \\ 1 & 0 \end{smallmatrix}\right]$ denote the product from (B.18) without the last vector, such that (B.20) below holds. Then

$$b_{N-1}(z) = \begin{bmatrix} h_{N,1}(z) & h_{N,2}(z) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \tag{B.19}$$

$$b_N(z) = \begin{bmatrix} h_{N,1}(z) & h_{N,2}(z) \end{bmatrix} \begin{bmatrix} \beta_{N-1}(z) \\ 1 \end{bmatrix}, \tag{B.20}$$

$$\begin{aligned} b_{N+1}(z) &= \begin{bmatrix} h_{N,1}(z) & h_{N,2}(z) \end{bmatrix} \begin{bmatrix} \beta_{N-1}(z) & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \beta_N(z) \\ 1 \end{bmatrix} \\ &= \beta_N(z) b_N(z) + b_{N-1}(z) \end{aligned} \tag{B.21}$$

shows that the recurrence relation (B.17) is valid. Moreover, $b_2(z)$ and $b_3(z)$ can be verified to coincide with the above definition. □

Lemma B.3.2. *Assume $N \geq 2$, and let $r_N(z) = [r_{N,0}(z), \ldots, r_{N,(N-1)}(z)]^\top$ solve the system of linear equations*

$$\underbrace{\begin{bmatrix} 1 & -\beta_0(z) & -1 & & \\ & \ddots & \ddots & -1 & \\ & & & -\beta_{N-2}(z) & \\ & & & 1 & \end{bmatrix}}_{M_N} r_N(z) = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \beta_{N-1}(z) \end{bmatrix}, \tag{B.22}$$

with $\beta_j(z) = (2j+1)z$, $j \geq 0$. Then, the N-th Bessel polynomial $b_N(z)$ and $(-1)^N b_N(-z)$ are given by

$$b_N(z) = r_{N,0}(z) + r_{N,1}(z), \tag{B.23}$$

$$(-1)^N b_N(-z) = r_{N,0}(z) - r_{N,1}(z), \tag{B.24}$$

using the first two components of $r_N(z)$ from (B.22).

Proof. The triangular matrix M_N has exclusively ones on the diagonal, and thus

$$M_N^{-1} = \frac{1}{\det(M_N)} \operatorname{adj}(M_N) = \operatorname{adj}(M_N). \tag{B.25}$$

Being only interested in the first two components $r_{N,0}(z)$ and $r_{N,1}(z)$ of the unknown vector $r_N(z)$ in (B.22), only the right upper 2×2 submatrix from (B.25) is required,

$$\begin{bmatrix} r_{N,0}(z) \\ r_{N,1}(z) \end{bmatrix} = \begin{bmatrix} (-1)^{(N-1)+1}\mathfrak{m}_{[N-1,1]} & (-1)^{N+1}\mathfrak{m}_{[N,1]} \\ (-1)^{(N-1)+2}\mathfrak{m}_{[N-1,2]} & (-1)^{N+2}\mathfrak{m}_{[N,2]} \end{bmatrix} \begin{bmatrix} 1 \\ \beta_{N-1}(z) \end{bmatrix}, \tag{B.26}$$

where $\mathfrak{m}_{[j,k]}$ denotes the j-th row, k-th column minor of the $N \times N$ matrix M_N,

$$\mathfrak{m}_{[N-1,1]} = \det \begin{bmatrix} -\beta_0 & -1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & 1 & -\beta_{N-3} & -1 \\ 0 & \cdots & 0 & 1 \end{bmatrix}, \qquad \mathfrak{m}_{[N,1]} = \det \begin{bmatrix} -\beta_0 & -1 & & \\ 1 & \ddots & & \\ & & & -1 \\ & & 1 & -\beta_{N-2} \end{bmatrix},$$

$$\mathfrak{m}_{[N-1,2]} = \det \begin{bmatrix} 1 & -1 & & & \\ 0 & -\beta_1 & & & \\ \vdots & 1 & \ddots & & \\ & & \ddots & 1 & -\beta_{N-3} & -1 \\ 0 & & \cdots & 0 & 1 \end{bmatrix}, \quad \mathfrak{m}_{[N,2]} = \det \begin{bmatrix} 1 & -1 & & & \\ 0 & -\beta_1 & & & \\ \vdots & 1 & \ddots & & -1 \\ 0 & & & 1 & -\beta_{N-2} \end{bmatrix}.$$

The signs in (B.26) simplify by using that $\det(M_{\mathrm{red}}) = (-1)^{N-1} \det(-M_{\mathrm{red}})$ for $M_{\mathrm{red}} \in \mathbb{R}^{(N-1)\times(N-1)}$. Continuants, i.e., determinants of tridiagonal matrices

$$K_{[i,\ldots,j]}(z) = \det \left(\begin{bmatrix} \beta_i(z) & 1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & -1 & \beta_j(z) \end{bmatrix} \right) \tag{B.27}$$

occur in the minors, in terms of which (B.26) can be written as

$$\begin{bmatrix} r_{N,0}(z) \\ r_{N,1}(z) \end{bmatrix} = \begin{bmatrix} K_{[0,\ldots,(N-3)]}(z) & K_{[0,\ldots,(N-2)]}(z) \\ K_{[1,\ldots,(N-3)]}(z) & K_{[1,\ldots,(N-2)]}(z) \end{bmatrix} \begin{bmatrix} 1 \\ \beta_{N-1}(z) \end{bmatrix}. \tag{B.28}$$

From the context of continued fractions (to which continuants are closely related as $\frac{K_{[0,\ldots,j]}(z)}{K_{[1,\ldots,j]}(z)}$ represents a finite continued fraction), the matrix identity

$$\begin{bmatrix} K_{[0,\ldots,j]}(z) & K_{[0,\ldots,j-1]}(z) \\ K_{[1,\ldots,j]}(z) & K_{[1,\ldots,j-1]}(z) \end{bmatrix} = \begin{bmatrix} \beta_0(z) & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} \beta_j(z) & 1 \\ 1 & 0 \end{bmatrix} \tag{B.29}$$

is known [181]. Hence, (B.28) with permuted columns becomes

$$\begin{bmatrix} r_{N,0}(z) \\ r_{N,1}(z) \end{bmatrix} = \begin{bmatrix} \beta_0(z) & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} \beta_{N-2}(z) & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \beta_{N-1}(z) \\ 1 \end{bmatrix}. \tag{B.30}$$

Lemma B.3.1, shows that

$$b_N(z) = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} r_{N,0}(z) \\ r_{N,1}(z) \end{bmatrix}, \tag{B.31}$$

which completes the proof of (B.23). Moreover, (B.24) holds since, if N is even, then $r_{N,0}$ contains only even powers of z, and $r_{N,1}$ contains only odd powers of z, whereas, if N is odd, then $r_{N,0}$ contains only odd powers of z, and $r_{N,1}$ contains only even powers of z. □

List of Figures

List of Tables

List of the Author's Publications

Prepublications of Contributions in this Thesis and Submissions

[S1] *Scholl, T. H.; Gröll, L.:* Stability criteria for time-delay systems from an insightful perspective on the characteristic equation. IEEE Transactions on Automatic Control 68 (2023) 4, 2352–2359.

[S2] *Scholl, T. H.; Hagenmeyer, V.; Gröll, L.:* What ODE-approximation schemes of time-delay systems reveal about Lyapunov–Krasovskii functionals. IEEE Transactions on Automatic Control 69 (2024) 7, 4614–4629.

[S3] *Scholl, T. H.:* Lyapunov–Krasovskii functionals of robust type for the stability analysis in time-delay systems. Submitted for publication, arXiv preprint available (2023). arXiv:2312.16738.

[S4] *Scholl, T. H.; Hagenmeyer, V.; Gröll, L.:* Lyapunov–Krasovskii functionals of robust type and their Legendre-tau-based approximation. IFAC-PapersOnLine (2024), in press.

Further Publications by the Author

[1] *Scholl, T. H.; Gröll, L.; Hagenmeyer, V.:* Time delay in the swing equation: A variety of bifurcations. Chaos 29 (2019) 12, 123118.

[2] *Scholl, T. H.; Hagenmeyer, V.; Gröll, L.:* On norm-based estimations for domains of attraction in nonlinear time-delay systems. Nonlinear Dynamics 100 (2020) 3, 2027–2045.

[3] *Scholl, T. H.; Hagenmeyer, V.; Gröll, L.:* Basic concepts for estimations of domains of attraction in time-delay systems. at - Automatisierungstechnik 68 (2020) 8, 667–686. in German.

[4] *Gröll, L.; Kastner, A.; Scholl, T. H.; Hagenmeyer, V.:* Some notes on two tests for stability in lossy power systems. IFAC-PapersOnLine 55 (2022) 40, 247–252.

Authorship Statement

The author's contributions in [S1], relevant to Chapter 2:

- idea for the presented criteria, i.e., the delay-dependent and the spectral-abscissa-based delay-independent criterion, and the framework of three perspectives;
  (motivated by initial research questions from L. Gröll: The starting point was a proposal by him for an improved logarithmic-norm-based delay-independent criterion. He also posed the question whether the spectral-abscissa-based criterion I came up with is even a necessary one, which in the end turned out to be the decisive question.)
- entire text;
- all theorems and complete development of the proofs;
- all examples and calculations;
- all graphics;
- largest part of the literature review and context research (some important references have been found by L. Gröll).

The author's contributions in [S2], relevant to Chapter 4:

- overall idea;
- entire text;
- all theorems and complete development of the proofs;
- all examples and calculations;
- all graphics/tables;
- entire literature review and context research.

The author's contributions in [S4], relevant to Section 6.1:

- overall idea;
- entire text;
- the outlook on theorems and proofs, which so far are only part of this thesis;
- all examples and calculations;
- all graphics/tables;
- (the context research coincides with [S2] and [S3] since, in [S4], the numerical method from [S2] is applied to the functionals from [S3]).

Bibliography

[5] *Aizerman, M. A.; Gantmacher, F. R.:* Die absolute Stabilität von Regelsystemen. Oldenbourg, München, 1965.

[6] *Alexandrova, I. V.:* On the robustness and estimation of the attraction region for a class of nonlinear time delay systems. Appl. Math. Lett. 106 (2020), 106374.

[7] *Anikushin, M.:* Frequency theorem and inertial manifolds for neutral delay equations. J. Evol. Equ. 23 (2023) 4, 66.

[8] *Appeltans, P.; Silm, H.; Michiels, W.:* TDS-control: a Matlab package for the analysis and controller-design of time-delay systems. IFAC-PapersOnLine 55 (2022) 16, 272–277.

[9] *Arnold, W. F.; Laub, A. J.:* Generalized eigenproblem algorithms and software for algebraic Riccati equations. Proc. IEEE 72 (1984) 12, 1746–1754.

[10] *Bajodek, M.; Gouaisbaut, F.; Seuret, A.:* Necessary and sufficient stability condition for time-delay systems arising from Legendre approximation. IEEE Trans. Autom. Control 68 (2023) 10, 6262–6269.

[11] *Bajodek, M.; Seuret, A.; Gouaisbaut, F.:* Order of Legendre-LMI conditions to assess stability of time-delay systems. IFAC-PapersOnLine 55 (2022) 36, 175–180.

[12] *Bajodek, M.; Seuret, A.; Gouaisbaut, F.:* On the necessity of sufficient LMI conditions for time-delay systems arising from Legendre approximation. Automatica 159 (2024), 111322.

[13] *Banks, H. T.; Fabiano, R. H.:* Approximation issues for applications in optimal control and parameter estimation. In: Modelling and computation for applications in mathematics, science, and engineering (ed.: Jerome, J. W.), Clarendon Press, Oxford 1998, 141–165.

[14] *Bellman, R.; Cooke, K. L.:* Differential-difference equations. Academic Press, London, 1963.

[15] *Bernier, C.; Manitius, A.:* On semigroups in $\mathbb{R}^n \times L^p$ corresponding to differential equations with delays. Canad. J. Math. 30 (1978) 5, 897–914.

[16] *Bernstein, D. S.:* Matrix mathematics: Theory, facts, and formulas. Princeton University Press, Princeton, NJ, 2009.

[17] *Bliman, P.-A.:* Extension of Popov absolute stability criterion to non-autonomous systems with delays. Int. J. Control 73 (2000) 15, 1349–1361.

[18] *Bliman, P.-A.:* LMI characterization of the strong delay-independent stability of linear delay systems via quadratic Lyapunov–Krasovskii functionals. Syst. Control Lett. 43 (2001) 4, 263–274.

[19] *Bliman, P.-A.:* Lyapunov–Krasovskii functionals and frequency domain: delay-independent absolute stability criteria for delay systems. Int. J. Robust Nonlinear Control 11 (2001) 8, 771–788.

[20] *Bliman, P.-A.:* Absolute stability criteria with prescribed decay rate for finite-dimensional and delay systems. Automatica 38 (2002) 11, 2015–2019.

[21] *Bliman, P.-A.:* Lyapunov equation for the stability of linear delay systems of retarded and neutral type. IEEE Trans. Automat. Contr. 47 (2002) 2, 327–335.

[22] *Bliman, P.-A.:* Stability of non-linear delay systems: Delay-independent small gain theorem and frequency domain interpretation of the Lyapunov–Krasovskii method. Int. J. Control 75 (2002) 4, 265–274.

[23] *Boyd, J. P.:* Chebyshev and Fourier spectral methods. Dover Publications, Mineola, NY, 2000.

[24] *Boyd, S.; Desoer, C. A.:* Subharmonic functions and performance bounds on linear time-invariant feedback systems. IMA J. Math. Control Inf. 2 (1985) 2, 153–170.

[25] *Boyd, S. P.:* Linear matrix inequalities in system and control theory. SIAM, Philadelphia, PA, 1994.

[26] *Breda, D.:* Pseudospectral methods for the stability analysis of delay equations. Part I: The infinitesimal generator approach. In: Controlling Delayed Dynamics (ed.: Breda, D.), Springer, Cham 2023, 65–94.

[27] *Breda, D.; Della Schiava, S.:* Pseudospectral reduction to compute Lyapunov exponents of delay differential equations. Discrete Contin. Dyn. Syst.-Ser. B 23 (2018) 7, 2727–2741.

[28] *Breda, D.; Diekmann, O.; Gyllenberg, M.; Scarabel, F.; Vermiglio, R.:* Pseudospectral discretization of nonlinear delay equations: New prospects for numerical bifurcation analysis. SIAM J. Appl. Dyn. Syst. 15 (2016) 1, 1–23.

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

[30] *Breda, D.; Maset, S.; Vermiglio, R.:* Stability of linear delay differential equations: A numerical approach with MATLAB. Springer, New York, NY, 2015.

[31] *Briat, C.:* Linear parameter-varying and time-delay systems: Analysis, observation, filtering & control. Springer, Heidelberg, 2015.

[32] *Brockett, R. W.:* Finite dimensional linear systems. Wiley, New York, NY, 1970.

[33] *Brogliato, B.; Lozano, R.; Maschke, B.; Egeland, O.:* Dissipative systems analysis and control: Theory and applications. Springer, Cham, 2020.

[34] *de Bruin, M. G.; Saff, E. B.; Varga, R. S.:* On the zeros of generalized Bessel polynomials. I. Indagationes Mathematicae (Proceedings) 84 (1981) 1, 1–13.

[35] *Burns, J.; Ito, K.; Propst, G.:* On nonconvergence of adjoint semigroups for control systems with delays. SIAM J. Control Optim. 26 (1988) 6, 1442–1454.

[36] *Burton, T. A.:* Stability and periodic solutions of ordinary and functional differential equations. Academic Press, Orlando, FL, 1985.

[37] *Böhlke, T.:* Skript zur Vorlesung Mathematische Methoden der Festigkeitslehre. Karlsruhe Institute of Technology, Karlsruhe, 2015.

[38] *Canuto, C.; Hussaini, M. Y.; Quarteroni, A.; Zang, T. A.:* Spectral methods: Fundamentals in single domains. Springer, Berlin, 2006.

[39] *Carlson, D.; Schneider, H.:* Inertia theorems for matrices: the semi-definite case. Bull. Amer. Math. Soc. 68 (1962) 5, 479–483.

[40] *Chen, J.:* On computing the maximal delay intervals for stability of linear delay systems. IEEE Trans. Automat. Contr. 40 (1995) 6, 1087–1093.

[41] *Chen, J.; Gu, G.; Nett, C. N.:* A new method for computing delay margins for stability of linear delay systems. Syst. Control Lett. 26 (1995) 2, 107–117.

[42] *Chen, J.; Latchman, H. A.:* Frequency sweeping tests for stability independent of delay. IEEE Trans. Automat. Contr. 40 (1995) 9, 1640–1645.

[43] *Chen, Y.; Wei, Y.:* Numerical radius for the asymptotic stability of delay differential equations. Linear Multilinear Algebra 65 (2017) 11, 2306–2315.

[44] *Chiasson, J.:* A method for computing the interval of delay values for which a differential-delay system is stable. IEEE Trans. Automat. Contr. 33 (1988) 12, 1176–1178.

[45] *Churilov, A. N.:* Solubility of matrix inequalities. Mathematical notes of the Academy of Sciences of the USSR 36 (1984) 5, 862–866.

[46] *Cooke, K. L.; Grossman, Z.:* Discrete delay, distributed delay and stability switches. J. Math. Anal. Appl. 86 (1982) 2, 592–627.

[47] *Curtain, R.; Zwart, H.:* Introduction to infinite-dimensional systems theory. Springer, New York, NY, 2020.

[48] *Cuvas, C.; Mondié, S.:* Necessary stability conditions for delay systems with multiple pointwise and distributed delays. IEEE Trans. Automat. Contr. 61 (2016) 7, 1987–1994.

[49] *Datko, R.:* A procedure for determination of the exponential stability of certain differential-difference equations. Quart. Appl. Math. 36 (1978) 3, 279–292.

[50] *Datko, R.:* Remarks concerning the asymptotic stability and stabilization of linear delay differential equations. J. Math. Anal. Appl. 111 (1985) 2, 571–584.

[51] *Davis, P. J.:* Interpolation and approximation. Dover Publications, New York, NY, 1975.

[52] *Delfour, M.; Mitter, S.:* Hereditary differential systems with constant delays. I. General case. J. Differ. Equ. 12 (1972) 2, 213–235.

[53] *Desoer, C. A.; Vidyasagar, M.:* Feedback systems: Input-output properties. (originally published: Academic Press, 1975). SIAM, Philadelphia, PA, 2009.

[54] *Deutsch, E.:* On matrix norms and logarithmic norms. Numer. Math. 24 (1975) 1, 49–51.

[55] *Deutscher, J.:* Zustandsregelung verteilt-parametrischer Systeme. Springer, Heidelberg, 2012.

[56] *Diekmann, O.; Scarabel, F.; Vermiglio, R.:* Pseudospectral discretization of delay differential equations in sun-star formulation: Results and conjectures. Discret. Contin. Dyn. Syst.-Ser. S 13 (2019) 9.

[57] *Diekmann, O.; Verduyn Lunel, S. M.; Gils, S. A.; Walther, H.-O.:* Delay equations: Functional-, complex-, and nonlinear analysis. Springer, New York, NY, 1995.

[58] *Dong, C.; Yang, S.; Jia, H.; Wang, P.:* Padé-based stability analysis for a modular multilevel converter considering the time delay in the digital control system. IEEE Trans. Ind. Electron. 66 (2019) 7, 5242–5253.

[59] *Driscoll, T. A.; Hale, N.; Trefethen, L. N.:* Chebfun guide. Pafnuty Publications, Oxford, 2014.

[60] *Egorov, A. V.; Cuvas, C.; Mondié, S.:* Necessary and sufficient stability conditions for linear systems with pointwise and distributed delays. Automatica 80 (2017), 218–224.

[61] *Egorov, A. V.; Kharitonov, V. L.:* Approximation of delay Lyapunov matrices. Int. J. Control 91 (2018) 11, 2588–2596.

[62] *Egorov, A. V.; Mondié, S.:* Necessary stability conditions for linear delay systems. Automatica 50 (2014) 12, 3204–3208.

[63] *Engel, K.-J.; Nagel, R.:* One-parameter semigroups for linear evolution equations. Springer, New York, NY, 2000.

[64] *Engelborghs, K.; Luzyanina, T.; Roose, D.:* Numerical bifurcation analysis of delay differential equations using DDE-BIFTOOL. ACM Trans. Math. Softw. 28 (2002) 1, 1–21.

[65] *Fornberg, B.:* A practical guide to pseudospectral methods. Cambridge University Press, Cambridge, 1996.

[66] *Francis, B. A.:* A course in H_∞ control theory. Springer, Berlin, 1987.

[67] *Fridman, E.:* Introduction to time-delay systems: Analysis and control. Springer, Cham, 2014.

[68] *Funaro, D.:* Polynomial approximation of differential equations. Springer, Berlin, 1992.

[69] *Garcia-Lozano, H.; Kharitonov, V. L.:* Lyapunov matrices for time delay system with commensurate delays. IFAC Proc. Vol. 37 (2004) 21, 91–95.

[70] *Gibson, J. S.:* Linear-quadratic optimal control of hereditary differential systems: Infinite dimensional Riccati equations and numerical approximations. SIAM J. Control Optim. 21 (1983) 1, 95–139.

[71] *Glover, K.; Lam, J.; Partington, J. R.:* Rational approximation of a class of infinite-dimensional systems II: Optimal convergence rates of L_∞ approximants. Mathematics of Control, Signals and Systems 4 (1991) 3, 233–246.

[72] *Gomez, M. A.; Egorov, A. V.; Mondié, S.:* Lyapunov matrix based necessary and sufficient stability condition by finite number of mathematical operations for retarded type systems. Automatica 108 (2019), 108475.

[73] *Gottlieb, D.; Orszag, S. A.:* Numerical analysis of spectral methods: Theory and applications. Soc. for Industrial and Applied Mathematics, Philadelphia, PA, 1993.

[74] *Gu, K.:* Discretized LMI set in the stability problem of linear uncertain time-delay systems. Int. J. Control 68 (1997) 4, 923–934.

[75] *Gu, K.:* A further refinement of discretized Lyapunov functional method for the stability of time-delay systems. Int. J. Control 74 (2001) 10, 967–976.

[76] *Gu, K.; Kharitonov, V. L.; Chen, J.:* Stability of time-delay systems. Birkhäuser, Boston, 2003.

[77] *Guiver, C.; Logemann, H.; Opmeer, M. R.:* Transfer functions of infinite-dimensional systems: positive realness and stabilization. Mathematics of Control, Signals and Systems 29 (2017) 4, 20.

[78] *Gusev, S. V.; Likhtarnikov, A. L.:* Kalman–Popov–Yakubovich lemma and the S-procedure: A historical essay. Autom. Remote Control 67 (2006) 11, 1768–1810.

[79] *Haag, T.; Münz, U.; Allgöwer, F.:* Comparison of different stability conditions for linear time-delay systems with incommensurate delays. IFAC Proc. Vol. 42 (2009) 14, 136–141.

[80] *Halanay, A.:* Differential equations: Stability, oscillations, time lags. Academic Press, New York, NY, 1966.

[81] *Hale, J. K.; Infante, E. F.; Tsen, F.-S. P.:* Stability in linear delay equations. J. Math. Anal. Appl. 105 (1985) 2, 533–555.

[82] *Hale, J. K.; Verduyn Lunel, S. M.:* Introduction to functional differential equations. Springer, New York, NY, 1993.

[83] *Han, Q.-L.:* Absolute stability of time-delay systems with sector-bounded nonlinearity. Automatica 41 (2005) 12, 2171–2176.

[84] *Han, Q.-L.:* A new delay-dependent absolute stability criterion for a class of nonlinear neutral systems. Automatica 44 (2008) 1, 272–277.

[85] *Hayes, N. D.:* Roots of the transcendental equation associated with a certain difference-differential equation. J. Lond. Math. Soc. 1 (1950) 3, 226–232.

[86] *Hermann, R.:* Ricci and Levi–Civita's tensor analysis paper: Translation, comments, and additional material. Number 2. Math Sci Press, Brookline, MA, 1975.

[87] *Hertz, D.; Jury, E. I.; Zeheb, E.:* Stability independent and dependent of delay for delay differential systems. J. Frankl. Inst. 318 (1984) 3, 143–150.

[88] *Hesthaven, J. S.; Gottlieb, S.; Gottlieb, D.:* Spectral methods for time-dependent problems. Cambridge University Press, Cambridge, 2007.

[89] *Hinrichsen, D.; Pritchard, A. J.:* Mathematical systems theory I: Modelling, state space analysis, stability and robustness. Springer, Berlin, 2005.

[90] *Hmamed, A.:* Further results on the robust stability of uncertain time-delay systems. Int. J. Syst. Sci. 22 (1991) 3, 605–614.

[91] *Horn, R. A.; Zhang, F.:* Basic properties of the Schur complement. In: The Schur complement and its applications (ed.: Zhang, F.), Springer, New York, NY 2005.

[92] *Huang, T.; Yang, F.; Zhang, D.; Chen, X.:* High-frequency stability analysis and impedance optimization for an MMC-HVDC integrated system considering delay effects. IEEE Journal on Emerging and Selected Topics in Circuits and Systems 12 (2022) 1, 59–72.

[93] *Huang, W.:* Generalization of Liapunov's theorem in a linear delay system. J. Math. Anal. Appl. 142 (1989) 1, 83–94.

[94] *Huesca, E.; Mondié, S.; Santos, O.:* Polynomial approximations of the Lyapunov matrix of a class of time delay systems. IFAC Proc. Vol. 42 (2009) 14, 261–266.

[95] *ed.: Iannelli, M.* Mathematics of biology. Springer, Berlin, 2011.

[96] *in't Hout, K. J.:* The stability of θ-methods for systems of delay differential equations. Annals Numer. Math. 1 (1994), 323–334.

[97] *Ito, K.:* Legendre-tau approximation for functional differential equations part III: Eigenvalue approximations and uniform stability. Proc. Distributed Parameter Systems (ed.: Kappel, F.; Kunisch, K.; Schappacher, W.), Springer, Berlin 1985, 191–212.

[98] *Ito, K.; Teglas, R.:* Legendre-tau approximations for functional-differential equations. SIAM J. Control Optim. 24 (1986) 4, 737–759.

[99] *Ito, K.; Teglas, R.:* Legendre-tau approximation for functional differential equations part II: The linear quadratic optimal control problem. SIAM J. Control Optim. 25 (1987) 6, 1379–1408.

[100] *Jarlebring, E.; Hochstenbach, M. E.:* Polynomial two-parameter eigenvalue problems and matrix pencil methods for stability of delay-differential equations. Linear Algebra Appl. 431 (2009) 3, 369–380.

[101] *Jarlebring, E.; Vanbiervliet, J.; Michiels, W.:* Characterizing and computing the $\mathcal{H}_2$ norm of time-delay systems by solving the delay Lyapunov equation. IEEE Trans. Automat. Contr. 56 (2011) 4, 814–825.

[102] *Juárez, L.; Alexandrova, I. V.; Mondié, S.:* Robust stability analysis for linear systems with distributed delays: A time-domain approach. Int. J. Robust Nonlinear Control 30 (2020) 18, 8299–8312.

[103] *Juárez, L.; Mondié, S.; Kharitonov, V. L.:* Dynamic predictor for systems with state and input delay: A time-domain robust stability analysis. Int. J. Robust Nonlinear Control 30 (2020) 6, 2204–2218.

[104] *Kamen, E.:* On the relationship between zero criteria for two-variable polynomials and asymptotic stability of delay differential equations. IEEE Trans. Automat. Contr. 25 (1980) 5, 983–984.

[105] *Kamen, E.:* Linear systems with commensurate time delays: stability and stabilization independent of delay. IEEE Trans. Automat. Contr. 27 (1982) 2, 367–375.

[106] *Kamen, E.:* Correction to "Linear systems with commensurate time delays: Stability and stabilization independent of delay". IEEE Trans. Automat. Contr. 28 (1983) 2, 248–249.

[107] *Kato, J.:* Absolute stability of control systems with multiple feedback. Proc. Seminar on Differential Equations and Dynamical Systems, II, Seminar Lectures at the University of Maryland 1969 (ed.: Yorke, J. A.), Springer, Berlin 1970, 128–133.

[108] *Khalil, H. K.:* Nonlinear systems. Prentice Hall, Upper Saddle River, NJ, 3rd edition, 2002.

[109] *Kharitonov, V. L.:* On the determination of the maximum admissible delay in a stability problem. Differentsial'nye Uravneniya 18 (1982) 4, 723–724. (in Russian).

[110] *Kharitonov, V. L.:* Time-delay systems: Lyapunov functionals and matrices. Birkhäuser Springer, New York, NY, 2013.

[111] *Kharitonov, V. L.; Zhabko, A. P.:* Lyapunov–Krasovskii approach to the robust stability analysis of time-delay systems. Automatica 39 (2003) 1, 15–20.

[112] *Kolmanovskii, V.; Myshkis, A.:* Applied theory of functional differential equations. Springer, Dordrecht, 1992.

[113] *Kolmanovskii, V. B.; Nosov, V. R.:* Stability of functional differential equations. Academic Press, London, 1986.

[114] *Krall, H. L.; Frink, O.:* A new class of orthogonal polynomials: The Bessel polynomials. Trans. Amer. Math. Soc. 65 (1949) 1, 100–115.

[115] *Krasovskii, N. N.:* On the analytic construction of an optimal control in a system with time lags. Pmm-J. Appl. Math. Mech. 26 (1962) 1, 50–67.

[116] *Krasovskii, N. N.:* Stability of motion: Applications of Lyapunov's second method to differential systems and equations with delay. Stanford University Press, Stanford, 1963.

[117] *Kudryakov, D. A.; Alexandrova, I. V.:* A new stability criterion and its application to robust stability analysis for linear systems with distributed delays. Automatica 152 (2023), 110973.

[118] *Lakshmikantham, V.; Leela, S.:* Differential and integral inequalities: Theory and applications; volume I: ordinary differential equations. Academic Press, New York, NY, 1969.

[119] *Lakshmikantham, V.; Liu, X. Z.:* Stability analysis in terms of two measures. World Scientific Pub. Co, Singapore, 1993.

[120] *Lancaster, P.; Rodman, L.:* Algebraic Riccati equations. Clarendon Press, Oxford, 1995.

[121] *Lancaster, P.; Tismenetsky, M.:* The theory of matrices. Academic Press, San Diego, CA, 1985.

[122] *Lanczos, C.:* Trigonometric interpolation of empirical and analytical functions. Journal of Mathematics and Physics 17 (1938) 1-4, 123–199.

[123] *Lanczos, C.:* Applied analysis. Prentice Hall, Englewood Cliffs, NJ, 1956.

[124] *LaSalle, J. P.; Artstein, Z.:* The stability of dynamical systems. SIAM, Philadelphia, PA, 1976.

[125] *Lee, S. M.; Park, J. H.:* Delay-dependent criteria for absolute stability of uncertain time-delayed Lur'e dynamical systems. J. Frankl. Inst. 347 (2010) 1, 146–153.

[126] *Liang, H.; Choi, B. J.; Zhuang, W.; Shen, X.:* Stability enhancement of decentralized inverter control through wireless communications in microgrids. IEEE Trans. Smart Grid 4 (2013) 1, 321–331.

[127] *Liberzon, M. R.:* Essays on the absolute stability theory. Automation and Remote Control 67 (2006) 10, 1610–1644.

[128] *Lichnerowicz, A.:* Elements of tensor calculus. Wiley, New York, NY, 2nd edition, 1962.

[129] *Likhtarnikov, A. L.:* Absolute stability criteria for nonlinear operator equations. Math. USSR Izv. 11 (1977) 5, 1011–1029.

[130] *Likhtarnikov, A. L.; Yakubovich, V. A.:* The frequency theorem for equations of evolutionary type. Sib. Math. J. 17 (1976) 5, 790–803.

[131] *Likhtarnikov, A. L.; Yakubovich, V. A.:* The frequency theorem for continuous one-parameter semigroups. Mathematics of the USSR-Izvestiya 11 (1977) 4, 849.

[132] *Louisell, J.:* A matrix method for determining the imaginary axis eigenvalues of a delay system. IEEE Trans. Automat. Contr. 46 (2001) 12, 2008–2012.

[133] *Lumer, G.:* Semi-inner-product spaces. Trans. Amer. Math. Soc. 100 (1961) 1, 29–43.

[134] *Medvedeva, I. V.; Zhabko, A. P.:* Constructive method of linear systems with delay stability analysis. IFAC Proc. Vol. 46 (2013) 3, 1–6.

[135] *Medvedeva, I. V.; Zhabko, A. P.:* A novel approach to robust stability analysis of linear time-delay systems. IFAC-PapersOnLine 48 (2015) 12, 233–238.

[136] *Medvedeva, I. V.; Zhabko, A. P.:* Synthesis of Razumikhin and Lyapunov–Krasovskii approaches to stability analysis of time-delay systems. Automatica 51 (2015), 372–377.

[137] *Melchor-Aguilar, D.; Niculescu, S.-I.:* Estimates of the attraction region for a class of nonlinear time-delay systems. IMA J. Math. Control Inf. 24 (2007) 4, 523–550.

[138] *Michiels, W.; Gumussoy, S.:* Characterization and computation of H_∞ norms for time-delay systems. SIAM J. Matrix Anal. & Appl. 31 (2010) 4, 2093–2115.

[139] *Michiels, W.; Niculescu, S.-I.:* Characterization of delay-independent stability and delay interference phenomena. SIAM J. Control Optim. 45 (2007) 6, 2138–2155.

[140] *Michiels, W.; Zhou, B.:* Computing delay Lyapunov matrices and $\mathcal{H}_2$ norms for large-scale problems. SIAM J. Matrix Anal. & Appl. 40 (2019) 3, 845–869.

[141] *Milano, F.; Anghel, M.:* Impact of time delays on power system stability. IEEE Trans. Circuits Syst. I 59 (2012) 4, 889–900.

[142] *Mondié, S.; Egorov, A.; Gomez, M. A.:* Lyapunov stability tests for linear time-delay systems. Annual Reviews in Control 54 (2022), 68–80.

[143] *Mori, T.; Fukuma, N.; Kuwahara, M.:* Simple stability criteria for single and composite linear systems with time delays. Int. J. Control 34 (1981) 6, 1175–1184.

[144] *Morris, K. A.:* Controller design for distributed-parameter systems. Springer, Cham, 2020.

[145] *Mou, Q.; Ye, H.; Liu, Y.:* Enabling highly efficient eigen-analysis of large delayed cyber-physical power systems by partial spectral discretization. IEEE Trans. Power Syst. 35 (2020) 2, 1499–1508.

[146] *Muhič, A.; Plestenjak, B.:* On the quadratic two-parameter eigenvalue problem and its linearization. Linear Algebra Appl. 432 (2010) 10, 2529–2542.

[147] *Musleh, A. S.; Muyeen, S. M.; Al-Durra, A.; Kamwa, I.; Masoum, M. A. S.; Islam, S.:* Time-delay analysis of wide-area voltage control considering smart grid contingences in a real-time environment. IEEE Trans. Ind. Inform. 14 (2018) 3, 1242–1252.

[148] *Niculescu, S.-I.:* Stability and hyperbolicity of linear systems with delayed state: a matrix-pencil approach. IMA J. Math. Control Inf. 15 (1998) 4, 331–347.

[149] *Niculescu, S.-I.; Verriest, E. I.; Dugard, L.; Dion, J.-M.:* Stability of linear systems with delayed state: A guided tour. IFAC Proc. Vol. 31 (1998) 19, 31–38.

[150] *Nussbaumer, T.; Heldwein, M. L.; Gong, G.; Round, S. D.; Kolar, J. W.:* Comparison of prediction techniques to compensate time delays caused by digital control of a three-phase buck-type PWM rectifier system. IEEE Trans. Ind. Electron. 55 (2008) 2, 791–799.

[151] *Peet, M.; Papachristodoulou, A.; Lall, S.:* Positive forms and stability of linear time-delay systems. SIAM J. Control Optim. 47 (2009) 6, 3237–3258.

[152] *Peiffer, K.; Rouche, N.:* Liapunov's second method applied to partial stability. Journal de mécanique 8 (1969) 2, 323–334.

[153] *Perron, O.:* Die Lehre von den Kettenbrüchen. Teubner, Leipzig, 1913.

[154] *Peyret, R.:* Spectral methods for incompressible viscous flow. Springer, New York, NY, 2002.

[155] *Plestenjak, B.:* MultiParEig. Version 2.5. MATLAB Central File Exchange, 2020. Available online at https://www.mathworks.com/matlabcentral/fileexchange/47844-multipareig.

[156] *Plischke, E.:* Transient Effects of Linear Dynamical Systems. PhD thesis, Universität Bremen, Bremen, 2005.

[157] *Pollard, H.:* The convergence almost everywhere of Legendre series. Proc. Amer. Math. Soc. 35 (1972) 2, 442–444.

[158] *Popov, V. M.; Halanay, A.:* On the stability of nonlinear automatic control systems with lagging argument. Autom. Remote Control 23 (1962) 7, 7.

[159] *Pritchard, A. J.; Townley, S.:* Robustness of linear systems. J. Differ. Equ. 77 (1989) 2, 254–286.

[160] *Rekasius, Z. V.:* A stability test for systems with delays. Proc. Proc. of the 1980 Joint Automatic Control Conference, volume 2, 1980, TP9–A.

[161] *Runge, C.:* Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten. Zeitschrift für Mathematik und Physik 46 (1901), 224–243.

[162] *Răsvan, V.; Niculescu, S.-I.:* Oscillations in lossless propagation models: a Liapunov–Krasovskii approach. IMA J. Math. Control Inf. 19 (2002), 157–172.

[163] *Scherer, C. W.:* The algebraic Riccati equation and inequality for systems with uncontrollable modes on the imaginary axis. SIAM J. Matrix Anal. & Appl. 16 (1995) 4, 1308–1327.

[164] *Schiffer, J.; Fridman, E.; Ortega, R.; Raisch, J.:* Stability of a class of delayed port-Hamiltonian systems with application to microgrids with distributed rotational and electronic generation. Automatica 74 (2016), 71–79.

[165] *Sepulchre, R.; Janković, M.; Kokotović, P. V.:* Constructive nonlinear control. Springer, London, 1997.

[166] *Seuret, A.; Gouaisbaut, F.:* Hierarchy of LMI conditions for the stability analysis of time-delay systems. Syst. Control Lett. 81 (2015), 1–7.

[167] *Seuret, A.; Gouaisbaut, F.; Baudouin, L.:* D1.1 - Overview of Lyapunov methods for time-delay systems: Rapport laas no. 16308. HAL archives-ouvertes.fr (2016) hal-01369516.

[168] *Shayman, M. A.:* Geometry of the algebraic Riccati equation, part I. SIAM J. Control Optim. 21 (1983) 3, 375–394.

[169] *Sipahi, R.:* Delay-margin design for the general class of single-delay retarded-type LTI systems. Int. J. Dyn. Control 2 (2014) 2, 198–209.

[170] *Solomonoff, A.; Turkel, E.:* Global properties of pseudospectral methods. J. Comput. Phys. 81 (1989) 2, 239–276.

[171] *Souza, F. O.; de Oliveira, M. C.; Palhares, R. M.:* Stability independent of delay using rational functions. Automatica 45 (2009) 9, 2128–2133.

[172] *Stépán, G.:* Retarded dynamical systems: Stability and characteristic functions. Longman Scientif. & Techn, Harlow, 1989.

[173] *Sun, J.; Vieto, I.; Larsen, E. V.; Buchhagen, C.:* Impedance-based characterization of digital control delay and its effects on system stability. Proc. 2019 20th Workshop on Control and Modeling for Power Electronics (COMPEL), 2019, 1–8.

[174] *Talitckii, A.; Peet, M. M.; Seiler, P.:* Integral quadratic constraints with infinite-dimensional channels. Proc. 2023 American Control Conference (ACC), 2023, 1576–1583.

[175] *Thowsen, A.:* The Routh–Hurwitz method for stability determination of linear differential-difference systems. Int. J. Control 33 (1981) 5, 991–995.

[176] *Townsend, A.; Webb, M.; Olver, S.:* Fast polynomial transforms based on Toeplitz and Hankel matrices. Math. Comp. 87 (2018) 312, 1913–1934.

[177] *Trefethen, L. N.:* Spectral methods in MATLAB. SIAM, Philadelphia,PA, 2000.

[178] *Trefethen, L. N.:* Is Gauss quadrature better than Clenshaw–Curtis? SIAM Rev. 50 (2008) 1, 67–87.

[179] *Trefethen, L. N.:* Approximation theory and approximation practice, extended edition. Society for Industrial and Applied Mathematics, 2019.

[180] *Truesdell, C.; Noll, W.; Antman, S. S.:* The non-linear field theories of mechanics. Springer, Berlin, 2004.

[181] *van der Poorten, A. J.:* An introduction to continued fractions. In: Diophantine analysis (ed.: Loxton, J. H.; van der Poorten, A. J.), Cambridge University Press, Cambridge 1986, 99–138.

[182] *Villafuerte, R.; Mondié, S.:* On improving estimate of the region of attraction of a class of nonlinear time delay system. IFAC Proc. Vol. 40 (2007) 23, 227–232.

[183] *Vinter, R. B.:* On the evolution of the state of linear differential delay equations in M_2: Properties of the generator. IMA J. Appl. Math. 21 (1978) 1, 13–23.

[184] *Vorotnikov, V. I.:* Partial stability and control. Birkhäuser, Boston, MA, 1998.

[185] *Vyhlídal, T.; Zítek, P.:* QPmR - quasi-polynomial root-finder: Algorithm update and examples. In: Delay Systems (ed.: Vyhlídal, T.; Lafay, J.-F.; Sipahi, R.), Springer, Cham 2014.

[186] *Walker, J. A.:* Stability of feedback systems involving time-delays and a time-varying nonlinearity. Int. J. Control 6 (1967) 4, 365–372.

[187] *Walton, K.; Marshall, J. E.:* Direct method for TDS stability analysis. IEE Proc. D - Control Theory Appl. 134 (1987) 2, 101–107.

[188] *Wang, J.; Yan, J. D.; Jiang, L.; Zou, J.:* Delay-dependent stability of single-loop controlled grid-connected inverters with LCL filters. IEEE Trans. Power Electron. 31 (2016) 1, 743–757.

[189] *Wang, X.; Blaabjerg, F.; Loh, P. C.:* Passivity-based stability analysis and damping injection for multiparalleled VSCs with LCL filters. IEEE Trans. Power Electron. 32 (2017) 11, 8922–8935.

[190] *Willems, J.:* Least squares stationary optimal control and the algebraic Riccati equation. IEEE Trans. Automat. Contr. 16 (1971) 6, 621–634.

[191] *Willems, J. C.:* Dissipative dynamical systems part I: General theory. Arch. Ration. Mech. Anal. 45 (1972) 5, 321–351.

[192] *de Wolff, B. A. J.; Scarabel, F.; Verduyn Lunel, S. M.; Diekmann, O.:* Pseudospectral approximation of Hopf bifurcation for delay differential equations. SIAM J. Appl. Dyn. Syst. 20 (2021) 1, 333–370.

[193] *Wu, H.; Tsakalis, K. S.; Heydt, G. T.:* Evaluation of time delay effects to wide-area power system stabilizer design. IEEE Trans. Power Syst. 19 (2004) 4, 1935–1941.

[194] *Wu, Z.; Michiels, W.:* Reliably computing all characteristic roots of delay differential equations in a given right half plane using a spectral method. J. Comput. Appl. Math. 236 (2012) 9, 2499–2514.

[195] *Xu, L.; Guo, Q.; Wang, Z.; Sun, H.:* Modeling of time-delayed distributed cyber-physical power systems for small-signal stability analysis. IEEE Trans. Smart Grid 12 (2021) 4, 3425–3437.

[196] *Yakubovich, V. A.:* S-procedure in nonlinear control theory. Vestnik Leningrad University (1971), 62–77. Russian.

[197] *Yakubovich, V. A.:* A frequency theorem in control theory. Sib. Math. J. 14 (1973) 2, 265–289.

[198] *Yakubovich, V. A.:* A frequency theorem for the case in which the state and control spaces are Hilbert spaces, with an application to some problems in the synthesis of optimal controls. I. Sib. Math. J. 15 (1974) 3, 457–476.

[199] *Yakubovich, V. A.:* Nonconvex optimization problem: The infinite-horizon linear-quadratic control problem with quadratic constraints. Syst. Control Lett. 19 (1992) 1, 13–22.

[200] *Yakubovich, V. A.; Leonov, G. A.; Gelig, A. C.:* Stability of stationary sets in control systems with discontinuous nonlinearities. World Scientific, Singapore, 2004.

[201] *Yao, W.; Jiang, L.; Wu, Q. H.; Wen, J. Y.; Cheng, S. J.:* Delay-dependent stability analysis of the power system with a wide-area damping controller embedded. IEEE Trans. Power Syst. 26 (2011) 1, 233–240.

[202] *Ye, H.; Liu, Y.; Zhang, P.:* Efficient eigen-analysis for large delayed cyber-physical power system using explicit infinitesimal generator discretization. IEEE Trans. Power Syst. 31 (2016) 3, 2361–2370.

[203] *Zakeri, H.; Antsaklis, P. J.:* Passivity measures in cyberphysical systems design: An overview of recent results and applications. IEEE Control Syst. Mag. 42 (2022) 2, 118–130.

[204] *Zhabko, A. P.; Alexandrova, I. V.:* Complete type functionals for homogeneous time delay systems. Automatica 125 (2021), 109456.

[205] *Zhu, J.; Hu, J.; Lin, L.; Wang, Y.; Wei, C.:* High-frequency oscillation mechanism analysis and suppression method of VSC-HVDC. IEEE Trans. Power Electron. 35 (2020) 9, 8892–8896.

[206] *Zou, C.; Liu, B.; Duan, S.; Li, R.:* Influence of delay on system stability and delay optimization of grid-connected inverters with LCL filter. IEEE Trans. Ind. Inform. 10 (2014) 3, 1775–1784.

[207] *Zwart, H. J.; Kuiper, C. R.:* Relation between invariant subspaces of the Hamiltonian and the algebraic Riccati equation. In: Modelling and Optimization of Distributed Parameter Systems (ed.: Malanowski, K.; Nahorski, Z.; Peszyńska, M.), Springer, Boston, MA 1996, 183–190.